USGS
science for a changing world

In cooperation with the Federal Emergency Management Agency

Estimates of the Magnitude and Frequency of Flood Flows in the Connecticut River in Connecticut

Open-File Report 2005-1369

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

Estimates of the Magnitude and Frequency of Flood Flows in the Connecticut River in Connecticut

By Elizabeth A. Ahearn

In cooperation with the Federal Emergency Management Agency

Open-File 2005-1369

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

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

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

U.S. Geological Survey, Reston, Virginia: 2005

For sale by U.S. Geological Survey, Information Services
Box 25286, Denver Federal Center
Denver, CO 80225

For more information about the USGS and its products:
Telephone: 1-888-ASK-USGS
World Wide Web: http://www.usgs.gov/

Contents

Figures

Tables

Conversion Factors and Datums

Multiply	By	To obtain
	Length	
foot (ft)	0.3048	meter
mile (mi)	1.609	kilometer
	Area	
square mile (mi^2)	2.590	square kilometer (km^2)
	Flow rate	
cubic foot per second (ft^3/s)	0.02832	cubic meter per second (m^3/s)

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

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

Estimates of the Magnitude and Frequency of Flood Flows in the Connecticut River in Connecticut

By Elizabeth A. Ahearn

Abstract

Annual peak-flow data from three U.S. Geological Survey (USGS) gaging stations on the Connecticut River—at Thompsonville, Conn. (USGS station 01184000), at Hartford, Conn. (USGS station 01190070), and near Middletown, Conn. (USGS station 01193000)—were used to estimate flood flows for annual exceedance probabilities of 0.1, 0.04, 0.02, 0.01, and 0.002 (10-, 25-, 50-, 100-, and 500-year recurrence intervals, respectively). Flood flows range from 198,000 cubic feet per second (ft^3/s) at Thompsonville to 211,000 ft^3/s at Middletown at the 100-year recurrence interval.

Flood-frequency analysis of annual peak flow through September 30, 2004 was performed using the procedures described in the publication "Guidelines for Determining Flood-Flow Frequency," commonly referred to as Bulletin 17B, published by the Interagency Advisory Committee on Water Data in 1982. The 100-year flood estimates at all three stations have large uncertainties, as represented by the 90-percent confidence interval (about 50,000 ft^3/s). Long-term records (1929-2004) at the three stations represent relatively constant watershed conditions based on results of statistical trend tests and analysis of frequency curves in hydrologically similar basins. Further study is needed to estimate the effects of flood-plain storage, channel storage, and flow regulation by dams on peak flows in the Connecticut River.

Introduction

The U.S. Geological Survey (USGS) collects and interprets flood data for the Nation. Estimates of the magnitude of floods for selected recurrence intervals (such as the 100-year recurrence interval) are needed by federal, state, and local officials for effective flood-plain management. This report, prepared by the USGS, in cooperation with the Federal Emergency Management Agency (FEMA), updates the flood flows for the 10-, 25-, 50-, 100-, and 500-year recurrence intervals for the Connecticut River in Connecticut from long-term data (tens of years) collected at three gaging stations: Thompsonville (USGS station 01184000), Hartford (USGS station 01190070), and Middletown (USGS station 01193000) (fig.1). Standard procedures for performing statistical flood-frequency analysis of annual peak flow, as described in Bulletin 17B (Interagency Advisory Committee on Water Data, 1982), were used to compute the magnitude and frequency of the floods at the three gaging stations. Information about the annual peak-flow data and assumptions of frequency analysis is included in the report. The report provides the flood-frequency estimates, frequency curves, and relevant information used in the analyses.

Data Used to Determine the Magnitude and Frequency of Flood Flows

The annual peak-flow data from the three gaging stations were analyzed to provide flood-frequency estimates to be used in updating the Flood Insurance Study (FIS) for the Connecticut River. The computer program PEAKFQ (Thomas and others, 1998) was used for the statistical flood-frequency analyses. PEAKFQ automates many of the analysis procedures recommended in Bulletin 17B, including identifying outliers, adjusting for historical periods, weighting skew coefficients, and fitting a log-Pearson Type III distribution to the streamflow data. The parameters of the Pearson Type III frequency curve are estimated by the logarithmic sample moments (mean, standard deviation, and skew coefficient), with adjustments for low outliers, high outliers, and historic peaks. The following sections discuss the annual peak-flow records used in the analyses, high-flow measurements used to verify the stage-discharge relations, and the results of the frequency analyses for the three stations.

Annual Peak-Flow Records

A 76-year record (1929-2004) at Thompsonville, a 162-year record (1843-2004) at Hartford, and a 58-year record (1947-2004) at Middletown (fig. 1) are available for flood-frequency analysis. Historic flood data for unusually large flood events on the Connecticut River at Hartford date back to 1683.

Figure 1. Location of U.S. Geological Survey gaging stations used to estimate the magnitude and frequency of floods on the Connecticut River in Connecticut. The entire Connecticut River Basin is shown on the index map; only the parts of the basin discussed in the report are shown on the larger map.

Time-series plots of the annual peak flow over time for Thompsonville, Hartford, and Middletown are shown in figure 2. The records at Thompsonville and Hartford encompass a wide range of water years[1]. The record for Middletown is limited, as it does not include some of the larger floods in the 20th century (most notably the floods of 1936 and 1938). The annual peak-flow data are stored in the USGS National Water-Information System (NWIS) peak-flow file [http://water.usgs.gov/nwis/sw].

The accuracy of the peak-flow records depends primarily on the accuracy of the stage-discharge relation, the frequency and range of the streamflow measurements used to define the stage-discharge relation, and the accuracy of the stage and streamflow measurements. A stage-discharge relation is developed by measuring streamflow at a variety of stages, plotting the stage against streamflow points, and drawing a best-fit curvilinear line through the points. Issues of data accuracy are most likely to affect the largest floods. These flood events occur infrequently and present fewer opportunities to measure these events and define the stage-discharge relation at the upper end of the curve. The stage-discharge relations for Hartford and Middletown are not well-defined by streamflow measurements, particularly for high stages at Hartford. Consequently, the stage-discharge relation and record of annual peak flow at Hartford and Middletown have a fair degree of uncertainty. (Gaging stations at Hartford and Middletown currently are operated as "stage-only" stations for flood-forecasting by the National Weather Service and do not require streamflow measurements.) Conversely, the stage-discharge relation at Thompsonville is well-defined for high stages, and the record of annual peak flow is reasonably accurate. The streamflow-gaging station at Thompsonville covers the runoff from about 92 percent of the drainage area above Hartford. The total drainage of the Connecticut River is 9,660 square miles (mi^2) at Thompsonville, 10,487 mi^2 at Hartford, and 10,887 mi^2 at Middletown. Because of the accuracy in the stage-discharge relation at Thompsonville and the long length of record, the data from Thompsonville were used to help verify the magnitude and frequency of peak flows at Hartford and Middletown.

High-Flow Measurements, April 2005

High-flow measurements were made April 4-5, 2005 and were used to verify the stage-discharge relations at Thompsonville, Hartford, and Middletown. High-flow measurements were taken at four locations, including one measurement in Holyoke, Massachusetts (fig. 1). The April 4-5 flows had a 2- to 5-year recurrence interval. The measurements, made using ADCP (acoustic doppler current profiler) technology, show that peak flow on the Connecticut River increases as the drainage area increases: Holyoke, Massachusetts (drainage area 8,332 mi$^{2)}$ peak flow 91,900 ft^3/s; Thompsonville (drainage area 9,660 mi^2) peak flow 99,000 ft^3/s; Hartford (drainage area 10,487 mi^2) peak flow 102,000 ft^3/s; Middletown (drainage area 10,887 mi^2) peak flow 110,000 ft^3/s. [The peak flows estimated for April 4-5 are provisional (as of August 9, 2005)]. Channel storage between the Thompsonville and Middletown gaging stations was recognized as a factor that might appreciably affect the peak flows (Kinnison, 1938). High-flow measurements using ADCP technology show that the effects of channel storage in the reach between Thompsonville and Middletown appear to be minor or negligible; however, additional measurements at higher flows (greater than 10-year recurrence interval) are needed to verify the stage-discharge relations and evaluate the effects of flood-plain storage.

Estimates of the Magnitude and Frequency of Flood Flows in the Connecticut River in Connecticut

Estimates of flood flows for Thompsonville, Hartford, and Middletown having annual exceedance probabilities of 0.10, 0.04, 0.02, 0.01, and 0.002 (recurrence intervals of 10, 25, 50, 100, and 500 years, respectively) are presented in table 1. Flood flows at the 100-year recurrence interval range from 198,000 ft^3/s at Thompsonville to 211,000 ft^3/s at Middletown. At the 10-year recurrence interval, the flow is about 140,000 ft^3/s from Thompsonville to Middletown. The years of systematic record and years of the historic peaks outside the systematic record are included in table 1. Station locations are shown in figure 1.

Flood-Frequency Analysis for Thompsonville, Hartford, and Middletown

Frequency curves for Thompsonville, Hartford, and Middletown are shown in figure 3. The frequency curves were computed by fitting a Pearson type-III distribution to the logarithms of the annual peak flows for each of the three stations. Because the record length that is used to compute the flood flows has a substantial effect on the frequency curve, a common period of record (1929-2004) was used in the frequency analysis for Thompsonville, Hartford, and Middletown. The 1929-2004 record is considered a relatively long record and includes floods of high recurrence intervals. Generally, a longer record reduces the variance in the estimated probabilities.

[1]A water year is defined as the 12-month period October 1 through September 30. The water year is designated by the calendar year in which it ends and which includes 9 of the 12 months. Thus, the year ending September 30, 2004 is called the 2004 water year.

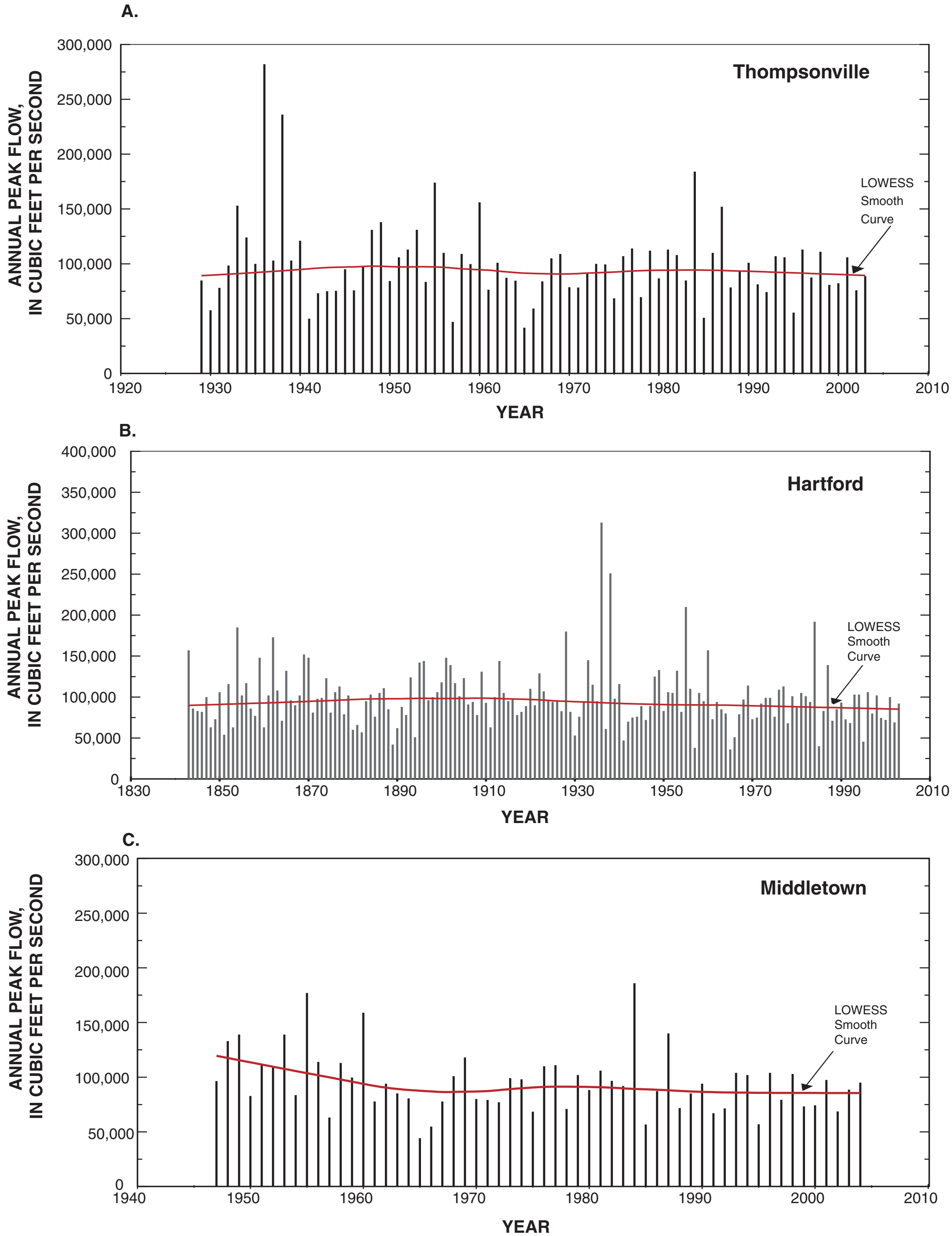

Figure 2. Time-series plots of the annual peak flow with a LOWESS smooth curve for trends in the annual peak-flow data for
(A) Connecticut River at Thompsonville, Conn. (USGS station 01184000),
(B) Connecticut River at Hartford, Conn. (USGS station 01190070), and
(C) Connecticut River near Middletown, Conn. (USGS station 01193000).

Table 1. Summary of the magnitude and frequency of flood flows for Connecticut River at Thompsonville, Conn. (USGS station 01184000), Connecticut River at Hartford, Conn. (USGS station 01190070), and Connecticut River near Middletown, Conn. (USGS station 01193000), for selected annual exceedance probabilities.

U.S. Geological Survey gaging station		Drainage area (square miles)	Period of record	Historic peaks outside period of record	Annual exceedance probability (percent)	Recurrence interval (years)	Estimates of flood flow, in cubic feet per second, derived using methods described in Bulletin 17B		
Number	Name						Flood-flow estimate	5-percent confidence limit	95-percent confidence limit
01184000	Connecticut River at Thompsonville	9,660	1929-2004	None	10	10	140,000	130,000	153,000
					4	25	163,000	150,000	181,000
					2	50	180,000	164,000	203,000
					1	100	198,000	178,000	225,000
					.2	500	239,000	211,000	279,000
01190070	Connecticut River at Hartford	10,487	[1]1929-2004	1683, 1692, 1801, 1828, 1838, 1839	10	10	138,000	128,000	153,000
			[2]1841-2004		4	25	163,000	148,000	183,000
					2	50	181,000	163,000	207,000
					1	100	200,000	178,000	230,000
					.2	500	243,000	212,000	288,000
01193000	Connecticut River at Middletown	10, 887	[3]1947-2004	1814, 1854, 1860, 1862, 1927, 1936, 1938	10	10	142,000	131,000	156,000
					4	25	169,000	154,000	189,000
					2	50	190,000	171,000	216,000
					1	100	211,000	189,000	244,000
					.2	500	265,000	231,000	315,000

[1]1929-2004 data used for flood-flow frequency analysis.

[2]1841-2004 data used for stage frequency analysis.

[3]1929-2004 data used for frequency analysis using the “Two-station comparison” method with records at Thompsonville.

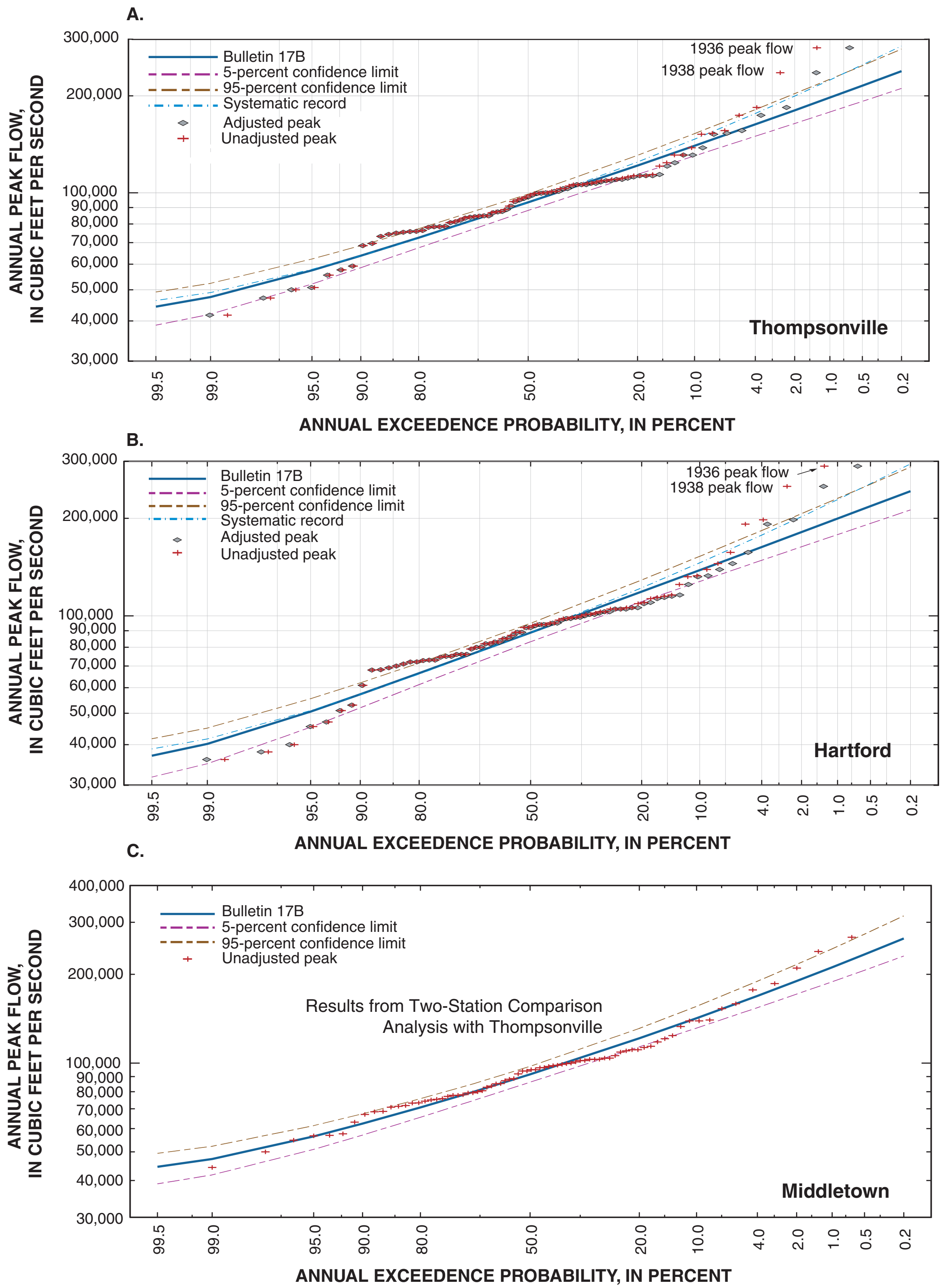

Figure 3. Flood-frequency curves for (A) Connecticut River at Thompsonville, Conn. (USGS station 01184000), (B) Connecticut River at Hartford, Conn. (USGS station 01190070), and (C) Connecticut River near Middletown, Conn. (USGS station 01193000).

From the analysis of the systematic record of 76 years (1929-2004) with an historic record adjustment (discussed below), the flood flow at the 1-percent annual exceedance probability (100-year flood) is 198,000 ft^3/s at Thompsonville and 200,000 ft^3/s at Hartford. The 90-percent confidence interval associated with the 100-year flood ranges from 178,000 to 225,000 ft^3/s at Thompsonville and 178,000 to 230,000 ft^3/s at Hartford. The systematic record at Thompsonville and Hartford contains two peaks (1936 and 1938 flood events) that are high outliers (fig. 3). Outliers depart significantly from the trend of the remaining data when plotted as a frequency curve on magnitude-probability coordinates. The 1936 and 1938 floods are considered extraordinarily large floods. Historic flood records indicate that the 1936 flood on the Connecticut River at Hartford was the largest event since 1854 (Thomson and others, 1964). The frequency statistics were calculated using an historic record adjustment that was based on a historical period of 150 years (back to 1854) and historical threshold of 230,000 ft^3/s at Thompsonville and 250,000 ft^3/s at Hartford. The historic record can be used to supplement the systematic record provided that all the historical peaks that are above some historical threshold have been documented. Peaks above these historic thresholds in the systematic annual peak-flow record are adjusted down on the basis that these floods are expected to occur less frequently. The historic record adjustment indicates that the unrecorded portion of the historic record contains only peaks below the historic threshold. For example, the record at Thompsonville is known to be complete for all flood events exceeding 230,000 ft^3/s (historic-adjustment threshold) between 1854 and 2004 (historical period). Historic-record adjustments for high outliers were applied in accordance with procedures described in Bulletin 17B (Interagency Advisory Committee on Water Data, 1982).

The flood-frequency estimates for Middletown were determined using the “Two Station Comparison” procedure (Interagency Advisory Committee on Water Data, 1982, appendix 7). The procedure is based on adjusting the mean and standard deviation of annual peaks using regression techniques with the long-term record at Thompsonville. The reliability of the adjusted data depends on the correlation between the short-term station (Middletown) data and the long-term station (Thompsonville) data and the length of the concurrent record. The standard error of estimate of the flood estimate at the 100-year recurrence interval can be reduced by 10 percent when the criterion for this procedure is met (correlation coefficient is greater than 0.90 and the record length calculation [described in Bulletin 17B, appendix 7] is greater than 2.0). The criterion for this procedure was met: correlation coefficient is 0.96 and the record length calculation is 2.31. The concurrent record is 58 years. From the two-station comparison procedure, the flood flow at Middletown at the 1-percent annual exceedance probability is 211,000 ft^3/s, and the 90-percent confidence interval ranges from 189,000 to 244,000 ft^3/s.

From the analysis of the systematic record at Middletown (1947-2004) prior to extending the record length or applying an historic record adjustment, the flood flow at the 1-percent annual exceedance probability (100-year flood) is 181,000 ft^3/s. The 90-percent confidence interval ranges from 163,000 to 208,000 ft^3/s. Adjusting the systematic record at Middletown with the historic record (1854, 1936, and 1938 flood events), the frequency estimate at the 1-percent annual exceedance probability is 215,000 ft^3/s, and the 90-percent confidence interval ranges from 189,000 to 254,000 ft^3/s. The flood flows from the analysis of the systematic record with historic-record adjustment and the results from two-station comparison procedure are similar (within 2 percent at the 100-year recurrence interval) and comparable with the upstream flood estimates.

Stage-Frequency Analysis for Hartford

Stage-based flood-frequency analysis can be used as an alternative approach in computing flood probabilities. A benefit of stage-based frequency analysis is that this analysis reduces the error in translating to a discharge and back to a flood elevation. Stage, however, is site specific and cannot be extrapolated upstream or downstream. Because the Hartford station has an accurate and long stage record (1843-2004), a stage-frequency analysis was performed on the record of stages (flood elevations). The probability plotting positions were computed using the Bulletin 17B formula (Weibull formula). The probability positions were plotted against the untransformed stage on an arithmetic-normal probability scale (fig. 4). The results of the stage-frequency analysis show that, at the 1-percent annual-exceedance probability (100-year flood), the stage at Hartford is 31.6 ft, and the 90-percent confidence interval ranges from 30.2 to 33.3 ft (table 2). Because of the stream slope and variability of channel geometry, the computed stage is only appropriate for the location where the stage data were collected (gaging station at Hartford). In flat topography, the range associated with the 90-percent confidence interval at the 100-year flood elevation can translate to a large difference in the delineation of the regulatory 100-year flood plain.

Table 2. Summary of the magnitude and frequency of stage for Connecticut River at Hartford, Conn. (USGS station 01190070).

Annual exceedance probability	Recurrence interval (years)	Estimates of stage, in feet, derived using methods described in Bulletin 17B		
		Stage estimate	5-percent confidence limit	95-percent confidence limit
10	10	25.5	24.7	26.4
4	25	28.1	27.0	29.3
2	50	29.9	28.7	31.3
1	100	31.6	30.2	33.3
.2	500	35.5	33.7	37.8

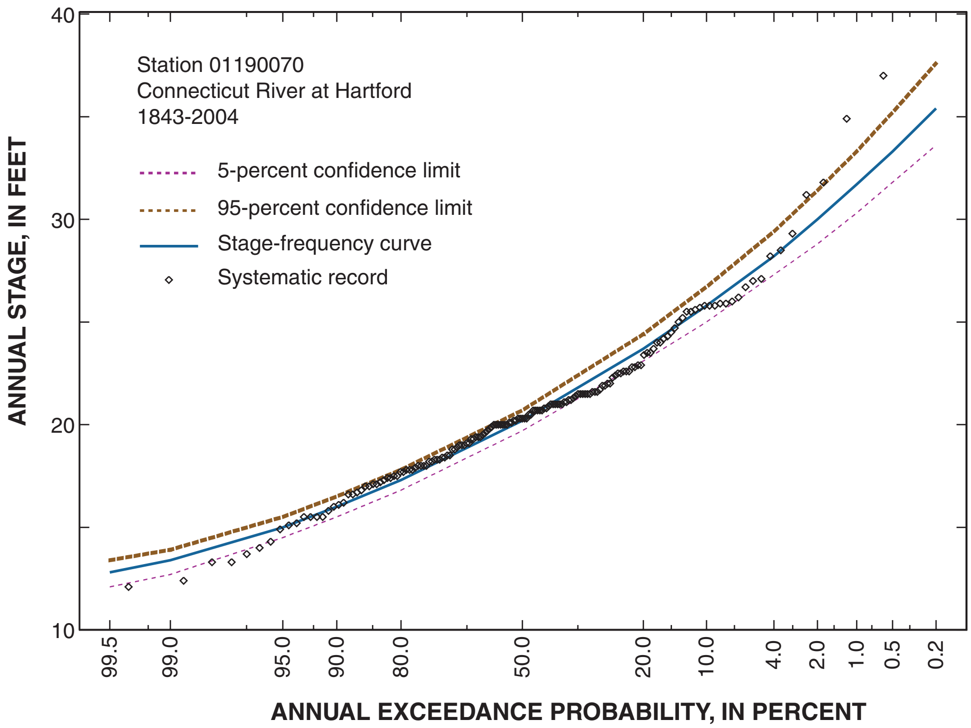

Figure 4. Stage-frequency curve for Connecticut River at Hartford, Conn. (USGS station 01190070).

Confidence Limits

Confidence limits provide a measure of the uncertainty of the flow at a selected exceedance probability. The limits are farther apart, representing greater uncertainty, in the tail of the distribution than in the center (near the 50-percent annual exceedance probability, 2-year flood). In practice, the record length or sample size usually is small (less than 60 years) in relation to the annual exceedance probabilities or recurrence intervals of interest (100 or 500 years). The 5- and 95-percent confidence limits, (also referred to as the 90-percent confidence interval) for Connecticut River at Thompsonville, Hartford, and Middletown are shown in table 1 and in figure 3. At the 1-percent annual exceedance probability (100-year flood), the 90-percent confidence interval is 47,000 ft^3/s at Thompsonville, 52,000 ft^3/s at Hartford, and 55,000 ft^3/s at Middletown (table 1). At Hartford, the 90-percent confidence interval for the stage at the 1-percent annual exceedance probability is 3.1 ft (table 2).

The updated flood-flow estimates for Thompsonville and Hartford do not significantly differ from the established values determined in previous flood insurance studies (about 4- and 8-percent larger, respectively). The previously established values at Thompsonville (191,000 ft^3/s) and Hartford (185,500 ft^3/s) are within the 90-percent confidence interval of the updated flood flows (from the FIS for the City of Hartford, December 1986, and the City of Suffield, February 1979). The previous established value for Middletown (186,000 ft^3/s) is outside the 90-percent confidence intervals of the updated flood flows (FIS City of Middletown, March 2001). The updated flood flows for Middletown are about 13 percent larger than the previously established values. As more streamflow and stage data become available, the estimates of flood flows can be improved and the confidence limits narrowed.

Evaluation and Statistical Analysis of the Flood-Flow Estimates

The accuracy and reliability of flood-frequency analysis depends on three assumptions about the data (Interagency Advisory Committee on Water Data, 1982):

- Annual peak flows are independent and are randomly sampled from a population of all possible flood flows.
- Flood-generating mechanism is time-stationary; the hydrologic and the hydraulic conditions of the river and flood plain remain relatively constant through time.
- Distribution of the logarithms of the annual peak flows can be approximated by the log-Pearson Type III.

Factors, such as the construction of flood-control dams and the effects of development (urbanization) in the basin, have the potential to make all or part of the flood record unrepresentative of future flood risks. Physical changes to the basin that affect the magnitude of the annual flood flows can invalidate the assumptions of flood-frequency analysis. The past flood record is considered a sample of the total statistical population consisting of past and future floods in the frequency analysis. The assumptions necessary for the flood-frequency analysis completed as part of this study were evaluated by assessing changes in the basin and doing statistical trend tests to ensure useful and accurate statistical results from the flood-frequency analysis.

Flood Control in the Connecticut River Basin

Between 1940 and 1971, the U.S. Army Corps of Engineers (USACE) built 16 flood-control dams in the Connecticut River Basin. The majority of the dams were in operation by 1965. The 16 dams, which are located on headwater streams, control storm runoff from about 15 percent of the basin's drainage area. Conversely, storm runoff from 85 percent of the Connecticut River drainage area at Thompsonville (about 8,200 mi^2) is uncontrolled. The effects of the flood-control dams on the magnitudes of the peak flows in the Lower Connecticut River may not persist as far downstream as Thompsonville and Hartford depending on the spatial and temporal source of the runoff generating the peak flows.

A study of streams in humid areas by the USGS found that the magnitudes of the annual peak flows are affected by less than 10 percent when the storage in a basin is less than 103 acre-feet per square mile (acre-ft/mi^2) (Benson, 1962). At Thompsonville, the flood-control storage associated with the USACE dams is about 50 acre-ft/mi^2 (appendix 1). Because the flood-control storage at Thompsonville is considerably less than 103 acre-ft/mi^2, it is assumed for the frequency analyses that the effects of flood-control dams do not appreciably alter the magnitudes of the peak flows on the Connecticut River at Thompsonville, Hartford, and Middletown.

Comparison of Frequency Curves in Hydrologically Similar Basins

Comparisons of frequency curves derived from an analysis of the entire record and the 1965-2004 record (post-flood control) at Thompsonville and Hartford and those from a hydrologically similar region were used to test the reasonableness of frequency determinations. The frequency curves derived from an analysis of data collected from 1965 to 2004 are less steep and their annual exceedance probability discharges are considerably less than the frequency curves derived from long-term records (greater than 70 years). For example, the 100-year flood flow at Thompsonville is 164,000 ft^3/s based on 1965-2004 annual peak discharges, and 198,000 ft^3/s based on the annual peak discharges of the entire record (1929-2004). With the 1965-2004 period, errors of sampling introduce large errors in judging the magnitude of the greater floods. Because the 1965-2004 record excludes some of the greater floods of the 20th century, such as the 1936, 1938, and 1955 flood peaks, the flood probability at the 100-year recurrence interval for the 1965-2004 period is biased and under estimates the flood probabilities for higher recurrence intervals.

The pattern of frequency curves being less steep for the record 1965-2004 as compared to long-term record (70 or more years) also was observed on the Housatonic and Pomperaug Rivers in Connecticut and the Hudson River in New York, indicating that the lower flood probabilities from an analysis of the 1965-2004 record is a result of a short-term climatic phenomena rather than the result of the flood-control dams. The Housatonic and Pomperaug Rivers are not regulated by flood-control dams, and the Hudson River in New York has some minor regulation. The frequency curve comparison may provide some evidence that the short record (1965-2004) is biased. Floods studied by use of glacial geology and tree rings also show both long-term and short-term variations in time. In the absence of any definitive proof that the recent period characterized by lower peak flows represents a climatic shift that will continue in the future, the flood-frequency curves based on the entire record (76 years) with historic data incorporated in the analysis provide conservatively interpreted estimates of flood probabilities.

By adjusting the 1936 and 1938 peaks to reflect flood control using USACE-modified peak flow values for 1936 and 1938 peaks in the frequency analysis, the 100-year flood flow at Thompsonville is similar (within 1-percent) to the estimate based on the entire record. From the analysis of the systematic record of 42 years (1965-2004) and a historic record adjustment applied to 3 historic peaks (1936-USACE modified, 1938-USACE modified, and 1955), and a historic record adjustment, the flood flow at the 1-percent annual exceedance probability (100-year flood) is 196,000 ft^3/s at Thompsonville. This estimate (196,000 ft^3/s) is much closer to the estimate derived from the entire record (198,000 ft^3/s) than the estimate from the 1965-2004 record (164,000 ft^3/s).

Trend Analysis

Changes in the watershed conditions, resulting in corresponding changes in the magnitude of peak flows, can show up as trends or discrete jumps in time series. To check for trends and discrete jumps in the annual series of maximum peak flows (largest flow from each year's streamflow record), statistical tests and graphical methods were performed on the annual series of peak-flow data from the gaging stations at Thompsonville, Hartford, and Middletown. Time-series plots of the annual peak flow were visually inspected and used to make preliminary inferences concerning possible changes in the magnitudes over time. To illustrate central patterns in annual peak discharges as a function of time, a smooth line is added to the time-series plots. The smoothing procedure used is LOWESS, locally weighted scatterplot smoothing (Cleveland, 1979). The smooth curve is derived by the pattern of the data and indicates trend directions over time. A smoothing factor is used to control the fit of the curve to the data and ranges from 0 to 1. Smaller smoothing factors result in smoothing of the curve to the data. A smoothing factor of 0.5 was used in the final fit. Time-series plots of the annual peak flow and LOWESS smooth curves show no discrete jumps or significant trends in the data (fig. 2).

In addition to visual inspection of time-series plots, the Mann-Kendall trend test and the Kolmogorov-Smirnov test were used to test for changes in the central tendency and distribution of the annual peak flow. The Mann-Kendall test (Helsel and Hirsch, 1992) uses a rank-based procedure that tests for one-directional changes over time (whether the peaks tend to increase or decrease over time). This test can be used to evaluate subtle changes in the watershed (such as increased urbanization). The Kolmogorov-Smirnov test (Haan, 1977) compares the distribution of two samples (pre- and post-flood control periods) for determining if the two samples differ significantly. The Kolmogorov-Smirnov test was used to assess the effects of the flood-control dams. The Mann-Kendall trend test results show no statistically significant trends or changes in the annual peak flow over time. The Kolmogorov-Smirnov test results show no statistically significant differences in the distribution of the annual peak flow for the pre- and post-flood control periods. Consequently, there is no strong statistical evidence indicating that watershed conditions have significantly changed to invalidate the assumptions for flood-frequency analysis. Development in the Connecticut River Basin, particularly along the lower Connecticut River Valley, is relatively minor in comparison to the size of the undeveloped or rural part of the basin. The results of the statistical trend tests indicate that the long-term records (1929-2004) at Thompsonville, Hartford, and Middletown represent relatively constant watershed conditions. From the results of the trend test, comparison of the frequency curves, and the basin's usable storage being less than the criterion for affecting the magnitude of the peak flows, it was justifiable to treat the entire record of annual peak flows as a single population for flood-frequency analysis.

Skew Analysis

The skew of the frequency distribution has a large effect on the shape of the frequency curve and the resulting flood probabilities. The accuracy of the skew coefficient for stations with short records with extreme events can be improved by weighting the station skew with a generalized skew. For records that are considered long, the station skew is given more weight than the generalized skew. The skew map in Bulletin 17B provides estimates of generalized skew. The Bulletin 17B skew map is based on data for essentially unregulated basins. The generalized skew coefficients in Bulletin 17B are not representative of regional conditions for the Connecticut River in Connecticut. The physiographic factors controlling the skew coefficients of large basins are different, at least in magnitude, from those in smaller basins. The development of an average regional coefficient of skew for the Connecticut River is impractical to attain (few basins that are similar in size in the region), as well as unnecessary (the record length exceeds 70 years). The difference between frequency estimates derived using the station skew and the weighted skew is negligible—less than 0.5 percent at the 1-percent annual exceedance probability, primarily because of the long record length. The station skew was used in defining the final frequency curves, because (1) the Bulletin 17B skew map is based on data that are not consistent with the Lower Connecticut River, and (2) the annual peak flows likely differ from the natural flow to some (unknown) extent.

Mixed Population Analysis

Peak flows associated with different climatic processes often fit different frequency distributions. If the distributions are appreciably different, the composite frequency distribution may have a sharp curvature that cannot be fit by a log-Pearson Type III distribution. A mixed-population method requires that peak flows be separated by cause (for example, tropical storm, snowmelt, or rainfall) prior to analysis and a frequency curve be developed, for example, for each cause with 10 or more years of data (10 values). A final flood-frequency curve is developed by combining the single-population curves for each cause.

Frequency curves were developed for Thompsonville, Hartford, and Middletown from the entire population (all the peaks) and evaluated for goodness of fit. Visual inspection of the frequency curves derived from the entire population did not show a break in the curve that would indicate the presence of a mixed population. Of the 162 years of peak flow record (1843-2004) for the Connecticut River at Hartford, three peaks depart significantly from the trend of the remaining data and are high outliers. The high outliers are from two meteorological conditions: the 1936 flood caused by rain and snowmelt, and the 1938 and 1955 floods caused by hurricanes. The outliers depart from the fitted (log-Pearson Type III) frequency curve. Of the 76 years of peak flow record (1929-2004) for the Connecticut River at Thompsonville, two peaks (1936 and 1938 flood events) depart from the trend of the data and are high outliers. The two outliers for Thompsonville also depart from the fitted frequency curve and are outside the 90-percent confidence interval. Because of two factors—the small number of peaks that depart from the frequency distribution, and the reasonableness of the fit of the entire population to the log-Pearson Type III distribution—a mixed-population analysis was not warranted for the stations. A single-population analysis (not separated by event type) was used to compute the final flood-frequency estimates.

Summary

Estimates of the magnitude of floods for selected recurrence intervals (for example, the 100-year recurrence interval) are needed by federal, state, and local officials for effective flood-plain management. Because of this need, the U.S. Geological Survey (USGS), in cooperation with the Federal Management Agency (FEMA), completed an analysis updating the flood flows for the 10-, 25-, 50-, 100-, and 500-year recurrence intervals for the Connecticut River in Connecticut. Flood flows were updated from long-term streamflow and stage data collected at three USGS gaging stations: Thompsonville, Conn. (station 01184000), Hartford, Conn., (station 01190070), and Middletown, Conn. (station 01193000).

Flood-frequency estimates were derived for Thompsonville, Hartford, and Middletown from data through September 30, 2004 (76-, 158-, and 58-years, respectively) by fitting the annual series of peak-flow data to a log-Pearson Type III frequency distribution. The final curves were based on an analysis of a common time period (1929-2004) to provide continuity in the frequency estimates between sites. Frequency estimates were adjusted for historic flood information and high outliers. The station skew was used to derive the final frequency curves. A single-population analysis (not separated by event type) was used to compute the final flood-frequency estimates. The frequencies of peak flows over a range of magnitudes were computed following the guidelines recommended in Bulletin 17B by the Interagency Advisory Committee on Water Data in 1982. The 100-year flood flows range from 198,000 ft^3/s at Thompsonville to 211,000 ft^3/s at Middletown. The 100-year flood estimates have large uncertainties, as represented by the 90-percent confidence interval (95- and 5-percent confidence limits). The 100-year flows are 4- and 8-percent larger at Thompsonville and Hartford and about 13 percent larger at Middletown than the values in previous flood insurance studies.

An analysis of changes in the Connecticut River Basin and statistical trend test results indicate that the long-term records (1929-2004) at Thompsonville, Hartford, and Middletown represent relatively constant watershed conditions. High-flow measurements above a 10-year recurrence interval are needed to verify the stage-discharge relation at Hartford and Middletown. Further investigations also are needed to estimate the effects of flood-plain storage, channel storage, and flow regulation on peak flows.

References

Ahearn, E.A., 2003, Peak-flow frequency estimates for U.S. Geological Survey streamflow-gaging stations in Connecticut: U.S. Geological Survey Water-Resources Investigations Report 03-4196, 14 p.

Ahearn, E.A., 2004, Regression equations for estimating flood flows for the 2-, 10-, 25-, 50-, 100-, and 500-year recurrence intervals in Connecticut: U.S. Geological Survey Scientific Investigations Report 2004-5160, 62 p.

Benson, M.A., 1962, Factors influencing the occurrence of floods in a humid region of diverse terrain: U.S. Geological Survey Water-Supply Paper 1580-B, 64 p.

Cleveland, W.S., 1979, Robust locally weighted regression and smoothing scatterplots: Journal of American Statistical Association, v. 74, no. 368, p. 829-836.

Federal Emergency Management Agency, Federal Insurance Administration, March 1983, Guidelines and specifications for study contractors: FEMA 37.

Federal Emergency Management Agency, Federal Insurance Study—City of Hartford, Connecticut, Hartford County, December 14, 1986.

Federal Emergency Management Agency, Federal Insurance Study—City of Middletown, Connecticut, Middlesex County, revised March 7, 2001.

Federal Emergency Management Agency, Federal Insurance Study—City of Suffield, Connecticut, Hartford County, February 1979.

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

Helsel, D.R., and Hirsch, R.M., 1992, Statistical methods in water resources: New York, Elsevier, 522 p.

Interagency Advisory Committee on Water Data, 1982, Guidelines for determining flood flow frequency—Bulletin 17B of the Hydrology Subcommittee: U.S. Geological Survey, Office of Water-Data Collection, 183 p.

Kinnison, H.B., Conover, L.F., and Bigwood, B.L., 1938, Stages and flood discharges of the Connecticut River at Hartford, Connecticut: U.S. Geological Survey Water-Supply Paper 836-A, 18 p.

Thomas, W.O., Jr., Lumb, A.M., Flynn, K.M., and Kirby, W.H., 1998, User's manual for program PEAKFQ, annual flood frequency analysis using Bulletin 17B guidelines: U.S. Geological Survey unpublished document, 89 p.

Thomson, M.T., Gannon, W.B., Thomas, M.P., Hayes, G.S., and others, 1964, Historical floods in New England: U.S. Geological Survey Water-Supply Paper 1779-M, 105 p.

Appendix 1. Data for U.S. Army Corps of Engineers flood-control dams in the Connecticut River Basin.

[Data compiled from information on the U.S. Army Corps of Engineers website: http://www.nae.usace.army.mil/recreati/damlist.htm]

U.S. Army Corps of Engineers flood-control dam	U.S. Geological gaging station	Start of operation	Drainage area (square miles)	Percentage of drainage area at the gaging station that is controlled[1]	Usable storage (acre-feet)	Flood-control storage in the basin [2] (acre feet per square mile)
Union Village Dam, Vt.		March 1950	126		38,000	
North Hartland Lake, Vt.		January 1961	220		71,000	
North Springfield Lake, Vt.		September 1960	158		50,200	
Ball Mountain Lake, Vt.		October 1961	172		52,450	
Townshend Lake, Vt.		June 1961	106		32,900	
Subtotal			782		244,550	
	Connecticut River at Vernon, Vt.		6,260	12.5		39.1
Otter Brook Lake, N.H.		April 1958	47		17,600	
Surry Mountain Lake, N.H.		October 1941	100		31,680	
Birch Hill Dam, Mass.		December 1941	175		49,400	
Tully Dam, Mass.		September 1949	50		22,000	
Cumulative subtotal			1,154		365,230	
	Connecticut River at Turner Falls, Mass.		7,163	16.1		51.0
	Connecticut River at Montague City, Mass.		7,860	14.7		46.5
	Connecticut River at Holyoke, Mass.		8,332	13.8		43.8
Barre Falls Dam, Mass.		July 1958	55		24,000	
Conant Brook Dam, Mass.		June 1966	7.8		3,740	
Knightville Dam, Mass.		December 1941	162		49,000	
Littleville Dam, Mass.		October 1965	52.3		23,000	
Cumulative subtotal			1,431.1		464,970	
	Connecticut River at Thompsonville, Conn.		9,660	14.8		48.1
Colebrook River Dam, Conn.		June 1969	118		50,200	
Sucker Brook Dam, Conn.		June 1971	3.43		1,480	
Mad River Dam, Conn.		June 1963	18.2		9,510	
Cumulative total			1,570.7		526,160	
	Connecticut River at Hartford, Conn.		10,487.0	15.0		50.2
	Connecticut River at Middletown, Conn.		10,897.0	14.4		48.3

[1]Percentage of drainage area at the gaging station that is controlled is equal to the drainage area of the flood-control dam divided by the drainage area at the gaging station.

[2]Flood-control storage in the basin is equal to the usable storage divided by the drainage area at the gaging station.