

**In cooperation with the Illinois Center for Transportation, Illinois Department of Transportation,
and the Illinois Department of Natural Resources—Office of Water Resources**

Implementation and Evaluation of the Streamflow Statistics (StreamStats) Web Application for Computing Basin Characteristics and Flood Peaks in Illinois

Scientific Investigations Report 2009–5197

Implementation and Evaluation of the Streamflow Statistics (StreamStats) Web Application for Computing Basin Characteristics and Flood Peaks in Illinois

By Audrey L. Ishii, David T. Soong, and Jennifer B. Sharpe

In cooperation with the Illinois Center for Transportation, Illinois Department of Transportation, and the Illinois Department of Natural Resources—Office of Water Resources

Scientific Investigations Report 2009–5197

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

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

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

U.S. Geological Survey, Reston, Virginia: 2010

For more information on the USGS—the Federal source for science about the Earth, its natural and living resources, natural hazards, and the environment, visit <http://www.usgs.gov> or call 1-888-ASK-USGS

For an overview of USGS information products, including maps, imagery, and publications, visit <http://www.usgs.gov/pubprod>

To order this and other USGS information products, visit <http://store.usgs.gov>

Any use of trade, product, or firm names is for descriptive purposes only and does not imply endorsement by the U.S. Government.

Although this report is in the public domain, permission must be secured from the individual copyright owners to reproduce any copyrighted materials contained within this report.

Suggested citation:

Ishii, A.L., Soong, D.T., and Sharpe, J.B., 2010, Implementation and evaluation of the Streamflow Statistics (Stream-Stats) Web application for computing basin characteristics and flood peaks in Illinois: U.S. Geological Survey Scientific Investigations Report 2009–5197, 25 p.

Also published as:

Ishii, A.L., Soong, D.T., and Sharpe, J.B., 2010, Implementation and evaluation of the Streamflow Statistics (Stream-Stats) Web application for computing basin characteristics and flood peaks in Illinois: Illinois Center for Transportation Report FHWA-ICT-10-63, 39 p.

Technical Report Documentation Page

1. Report No. FHWA-ICT-10-063	2. Government Accession No.	3. Recipient's Catalog No.	
4. Title and Subtitle Implementation and evaluation of the Streamflow Statistics (StreamStats) Web application for computing basin characteristics and flood peaks in Illinois		5. Report Date	
		6. Performing Organization Code	
7. Author(s) Audrey L. Ishii, David T. Soong, Jennifer B. Sharpe		8. Performing Organization Report No. U.S. Geological Survey Scientific Investigations Report 2010-5197	
9. Performing Organization Name and Address U.S. Geological Survey Illinois Water Science Center 1201 W. University Avenue Urbana, IL 61801		10. Work Unit (TRAIS)	
		11. Contract or Grant No.	
12. Sponsoring Agency Name and Address Illinois Center for Transportation		13. Type of Report and Period Covered Final Report	
		14. Sponsoring Agency Code	
15. Supplementary Notes Prepared in cooperation with the Illinois Department of Natural Resources, Office of Water Resources, and the Illinois Department of Transportation			
16. Abstract Illinois StreamStats (ILSS) is a Web-based application for computing selected basin characteristics and flood-peak quantiles based on the most recently (2010) published (Soong and others, 2004) regional flood-frequency equations at any rural stream location in Illinois. Limited streamflow statistics including general statistics, flow durations, and base flows also are available for U.S. Geological Survey (USGS) streamflow-gaging stations. ILSS can be accessed on the Web at http://streamstats.usgs.gov/ by selecting the <i>State Applications</i> hyperlink and choosing <i>Illinois</i> from the pull-down menu. ILSS was implemented for Illinois by obtaining and projecting ancillary geographic information system (GIS) coverages; populating the StreamStats database (StreamStatsDB) with streamflow-gaging station data; hydroprocessing the 30-meter digital elevation model (DEM) for Illinois to conform to streams represented in the National Hydrography Dataset 1:100,000 stream coverage; and customizing the Web-based Extensible Markup Language (XML) programs for computing basin characteristics for Illinois. The basin characteristics computed by ILSS then were compared to the basin characteristics used in the published study, and adjustments were applied to the XML algorithms for slope and basin length. Testing of ILSS was accomplished by comparing flood quantiles computed by ILSS at an approximately random sample of 170 streamflow-gaging stations computed by ILSS with the published flood-quantile estimates. Differences between the log-transformed flood quantiles were not statistically significant at the 95-percent confidence level for the State as a whole, nor by the regions determined by each equation, except for region 1, in the northwest corner of the State. In region 1, the average difference in flood-quantile estimates ranged from 3.76 percent for the 2-year flood quantile to 4.27 percent for the 500-year flood quantile. The total number of stations tested in region 1 was small (21) and the mean difference is not large (less than one-tenth of the average prediction error for the regression-equation estimates). The sensitivity of the flood-quantile estimates to differences in the computed basin characteristics are determined and presented in tables. A test of usage consistency was conducted by having at least 7 new users compute flood-quantile estimates at 27 locations. The average maximum deviation of the flood-quantile estimate from the mode value at each site was 1.31 percent for the 100-year flood quantile after four mislocated sites were removed. A comparison of manual 100-year flood-quantile computations with ILSS computations at 34 sites indicated no statistically significant difference. ILSS appears to be an accurate, reliable, and effective tool for flood-quantile estimates.			
17. Key Words Flood frequency, flood-peak estimation, flood-peak discharge, streamflow statistics, automated watershed delineation, basin characteristics, Web application, StreamStats		18. Distribution Statement No restrictions. This document is available to the public through the National Technical Information Service, Springfield, Virginia 22161.	
19. Security Classif. (of this report) Unclassified	20. Security Classif. (of this page) Unclassified	21. No. of Pages 25	22. Price

Contents

Abstract	1
Introduction.....	1
Implementing Illinois StreamStats (ILSS)	2
Data Preparation.....	2
Computer Code Development.....	3
Illinois StreamStats Database	3
Evaluation and Adjustment of Basin Characteristics	5
Drainage Basin Area, Percent Water, Average Soil Permeability.....	5
Basin Length	6
Stream Slope	10
Evaluation of Flood-Peak Quantiles.....	12
Test of Illinois StreamStats Usage Consistency.....	17
Sensitivity of Flood-Peak Quantiles	18
Conclusions and Limitations	23
Summary.....	24
Acknowledgments	25
References Cited.....	25

Figures

1. Map showing hydrologic regions for flood-frequency regression equations of rural streams in Illinois (Soong and others, 2004)	4
2–13. Graphs showing—	
2. Relation between Illinois StreamStats and published drainage areas—logarithmically scaled axes	7
3. Relation between Illinois StreamStats and published drainage areas—arithmetically scaled axes	7
4–6. Relation between percent differences in published and Illinois StreamStats (ILSS) values of drainage areas and published drainage areas:	
4. Full scale	8
5. Scale truncated to 100 square miles	8
6. Scale truncated to 10 square miles	9
7. Relation between Illinois StreamStats basin length using 100-meter grid and the published BasinSoft basin length	9
8. Relation between Illinois StreamStats and published BasinSoft slope	10
9. Relation between adjusted slope residual (published BasinSoft-computed stream slope (BS_SL) minus adjusted slope (SLadj), expressed as a percent $[(BS_SL-SLadj)*100/BS_SL]$), and published drainage area	11
10. Relation between adjusted slope residual (published BasinSoft-computed stream slope (BS_SL) minus adjusted slope (SLadj), expressed as a percent $[(BS_SL-SLadj)*100/BS_SL]$), and BasinSoft slope	11

11. Relation between StreamStats and Soong and others (2004) 100-year flood quantiles.....	17
12. Relation between Illinois StreamStats (ILSS) and manually determined 100-year flood quantiles.....	18
13. Distribution of differences in basin characteristics between the published basin characteristics (2004_BC) and the Illinois StreamStats basin characteristics (ILSS_BC) for (a) drainage area, (b) average permeability, (c) percentage of open water and herbaceous wetland, (d) slope, and (e) basin length.....	19

Tables

1. Differences between the published basin characteristics and the Illinois StreamStats (ILSS) basin characteristics (published value minus ILSS value).....	6
2. Comparisons of published Q_T^{2004} and Illinois StreamStats (ILSS) Q_T^{ILSS} flood quantiles as proportional differences— $([Q_T^{2004} - Q_T^{ILSS}] / Q_T^{2004})$ and differences of the log-transformed quantiles ($\log [Q_T^{2004}] - \log [Q_T^{ILSS}]$)—for selected streamflow-gaging stations, all regions.....	13
3–9. Comparisons of published Q_T^{2004} and Illinois StreamStats (ILSS) Q_T^{ILSS} flood quantiles as proportional differences— $([Q_T^{2004} - Q_T^{ILSS}] / Q_T^{2004})$ and differences of the log-transformed quantiles ($\log [Q_T^{2004}] - \log [Q_T^{ILSS}]$)—for streamflow-gaging stations, sorted by region:	
3. Region 1.....	13
4. Region 2.....	14
5. Region 3.....	14
6. Region 4.....	15
7. Region 5.....	15
8. Region 6.....	16
9. Region 7.....	16
10–14. Sensitivity of flood quantiles based on the published regression equations for specified return-intervals, Q_T , to differences in drainage area for (a) regions 1, 3, and 5; (b) regions 2, 6, and 7; and (c) region 4, expressed as the ratio of the tested Q_T to the published Q_T , corresponding to percent of published—	
10. Drainage area.....	20
11. Stream slope.....	21
12. Average permeability.....	22
13. %Water.....	22
14. Basin length.....	23

Conversion Factors

Multiply	By	To obtain
Length		
inch (in.)	25.4	millimeter (mm)
foot (ft)	0.3048	meter (m)
mile (mi)	1.609	kilometer (km)
meter (m)	3.281	foot (ft)
Area		
square foot (ft ²)	0.09290	square meter (m ²)
square mile (mi ²)	2.590	square kilometer (km ²)
acre	0.4047	square hectometer
Flow rate		
cubic foot per second (ft ³ /s)	0.02832	cubic meter per second
Slope		
foot per mile (ft/mi)	0.1894	meter per kilometer

Abbreviations, Acronyms, and Symbols

BCF	basin characteristic file
BL	basin length
BS	BasinSoft
DA	drainage area
DEM	digital elevation model
DRGs	digital raster graphics
ESRI	Environmental Systems Research Institute, Inc.
GIS	geographic information system
ICT	Illinois Center for Transportation
IDNR–OWR	Illinois Department of Natural Resources–Office of Water Resources
IDOT	Illinois Department of Transportation–Bureau of Bridges and Structures
ILSS	Illinois StreamStats
LFP	longest flow path
MCL	main channel length
NED	National Elevation Dataset
NHD	National Hydrography Dataset
NLCD	National Land Cover Data
NRCS	Natural Resources Conservation Service
PermAvg	average soil permeability
σ	standard deviation
SL	slope
StreamStatsDB	StreamStats database
USGS	U.S. Geological Survey
USGS–ILWSC	USGS–Illinois Water Science Center
%Water	percentage of open water and herbaceous wetland
WATSTORE	Water Storage Retrieval system
WBD	Watershed Boundary Dataset
WDNR	Wisconsin Department of Natural Resources
WSRT	Wilcoxon Signed Rank Test
XML	extensible markup language

Implementation and Evaluation of the Streamflow Statistics (StreamStats) Web Application for Computing Basin Characteristics and Flood Peaks in Illinois

By Audrey L. Ishii, David T. Soong, and Jennifer B. Sharpe

Abstract

Illinois StreamStats (ILSS) is a Web-based application for computing selected basin characteristics and flood-peak quantiles based on the most recently (2010) published (Soong and others, 2004) regional flood-frequency equations at any rural stream location in Illinois. Limited streamflow statistics including general statistics, flow durations, and base flows also are available for U.S. Geological Survey (USGS) streamflow-gaging stations. ILSS can be accessed on the Web at <http://streamstats.usgs.gov/> by selecting the State Applications hyperlink and choosing Illinois from the pull-down menu.

ILSS was implemented for Illinois by obtaining and projecting ancillary geographic information system (GIS) coverages; populating the StreamStats database with streamflow-gaging station data; hydroprocessing the 30-meter digital elevation model (DEM) for Illinois to conform to streams represented in the National Hydrographic Dataset 1:100,000 stream coverage; and customizing the Web-based Extensible Markup Language (XML) programs for computing basin characteristics for Illinois. The basin characteristics computed by ILSS then were compared to the basin characteristics used in the published study, and adjustments were applied to the XML algorithms for slope and basin length. Testing of ILSS was accomplished by comparing flood quantiles computed by ILSS at an approximately random sample of 170 streamflow-gaging stations computed by ILSS with the published flood quantile estimates. Differences between the log-transformed flood quantiles were not statistically significant at the 95-percent confidence level for the State as a whole, nor by the regions determined by each equation, except for region 1, in the northwest corner of the State. In region 1, the average difference in flood quantile estimates ranged from 3.76 percent for the 2-year flood quantile to 4.27 percent for the 500-year flood quantile. The total number of stations in region 1 was small (21) and the mean difference is not large (less than one-tenth of the average prediction error for the regression-equation estimates). The sensitivity of the flood-quantile estimates to differences in the computed basin characteristics are determined and presented in tables. A test of usage consistency

was conducted by having at least 7 new users compute flood quantile estimates at 27 locations. The average maximum deviation of the estimate from the mode value at each site was 1.31 percent after four mislocated sites were removed. A comparison of manual 100-year flood-quantile computations with ILSS at 34 sites indicated no statistically significant difference. ILSS appears to be an accurate, reliable, and effective tool for flood-quantile estimates.

Introduction

Streamflow statistics such as peak-discharge estimates for floods of various frequencies (flood quantiles) are used widely in engineering and scientific applications such as determining flood plains, designing hydraulic structures such as bridges and culverts, and the planning and management of the State's water resources to protect water quality and supply. While representative streamflow records are essential for deriving reliable flow statistics, streamflow records are site-specific information, whereas the need for such information is region-wide. The U.S. Geological Survey (USGS) has developed regional regression equations for estimating statistical streamflow characteristics at ungaged sites, which are used to transfer site-based information, such as streamflow statistics, to those sites. The equations were developed by use of regression-analysis techniques to relate streamflow characteristics to basin characteristics, which can be determined through a variety of methods (Jennings and others, 1994). To apply the regional regression equations, the user must determine the same basin characteristics for ungaged sites that were determined for gaged sites and used in the regional regression equations. These characteristics are determined by a variety of methods such as the manual or digital measurement of maps; from field surveys; from paper records; or by other methods, such as geographic information system (GIS) software. Such steps are time-consuming or require considerable user expertise, and the results can be inconsistent. A single, integrated application that provides an automated determination of the needed basin characteristics and solves the regional regression

2 Implementation and Evaluation of the Streamflow Statistics (StreamStats) Web Application for Illinois

equations to provide the estimated flood quantiles can reduce the time required and the potential inconsistencies in the results. To meet this need, the USGS, in cooperation with the Environmental Systems Research Institute (ESRI), Inc., has developed StreamStats, a national Web-based GIS application that serves streamflow statistics and determines basin characteristics and flood quantiles based on consistently processed data sets and methods in utilizing the flood-frequency regional regression equations (Ries and others, 2008).

The USGS–Illinois Water Science Center (USGS–ILWSC) used basin characteristics that were derived from GIS data layers and from the application of an Arc INFO-based program, BasinSoft (version 1.1, Harvey and Eash, 1996), to determine the current (2010) regional regression equations for estimates of flood quantiles for rural streams in Illinois. The regional analysis included the investigation of functional relations with more numerous, consistently determined basin characteristics than was possible in earlier investigations (Soong and others, 2004). However, the use of GIS-based data and methods also created difficulties for users who do not have the resources to access the GIS databases and (or) software for determining these selected basin characteristics. The availability of StreamStats provided the opportunity to satisfy the public need for the utilization of GIS techniques without extensive software or user expertise; consequently, the USGS–ILWSC, in cooperation with the Illinois Center for Transportation (ICT); the Illinois Department of Transportation–Bureau of Bridges and Structures (IDOT); and the Illinois Department of Natural Resources–Office of Water Resources (IDNR–OWR), conducted this study to implement and evaluate the StreamStats application for determining basin characteristics and flood-peak quantiles for rural streams in Illinois.

The purpose of this report is to describe the procedures used to develop Illinois StreamStats (ILSS) and the analyses performed to evaluate ILSS and present the results. This includes the preparation of the GIS-data layers and Web-based Extensible Markup Language (XML) programming, the development of the streamflow-statistics database, the evaluation and adjustment of the basin characteristics determined from ILSS, and the evaluation of the flood-peak quantiles from ILSS. The limitations of the application and the sensitivity to basin characteristics differences also are described.

Implementing Illinois StreamStats (ILSS)

StreamStats is a USGS Web-based application that makes the process of computing streamflow statistics faster and more consistent than previously used manual methods. StreamStats can be accessed on the Web at <http://streamstats.usgs.gov/>. StreamStats includes five major components: (1) the user interface, which displays the maps and enables users to select the stream locations for which information is desired; (2) the database, which contains previously published streamflow statistics

and other descriptive information for streamflow-gaging stations; (3) the automated GIS processes, which determine the drainage boundaries and other drainage-basin characteristics by utilizing the underlying preprocessed GIS database; (4) the GIS database, which stores the base-map data; and (5) the implementation of the National Streamflow Statistics program (Ries, 2006), which uses the regional regression equations along with the basin-characteristics input to compute and output the various streamflow statistics to the user display.

The implementation of ILSS required the collection and processing of base GIS-data layers for consistency in projection, hydroprocessing of the digital elevation model (DEM), development of ancillary data layers, programming algorithms for computing selected basin characteristics, comparing ILSS results—both basin characteristics and flood-peak quantiles—with those obtained in the 2004 analysis, and providing adjustments where required.

Data Preparation

Processing of three primary GIS-data layers was needed to produce the ILSS data layers. In the present study, the 1:100,000-scale USGS National Hydrography Dataset (NHD) (<http://nhd.usgs.gov/>) was used to develop a dendritic stream network. This processing involved the removal of braided streams and reconnecting or removing disconnected stream segments. All elevation information used in ILSS, including flow direction and flow accumulation, was derived from the USGS 1-arc-second National Elevation Dataset (NED). The 1-arc-second NED is a national seamless DEM with a resolution of 30 m (<http://seamless.usgs.gov/viewer.htm>). The downloaded NED blocks were merged and reprojected to the Albers Equal-Area Conic projection. After tile edges were examined to make sure elevation values were consistent, the NED was resampled to a 10-m resolution grid for the ILSS project. A hydro-corrected DEM was developed from the NED. This was done by first filling depressions or sinks in the NED (areas surrounded by areas of higher elevation values). Next, the NHD streams were “burned” into the NED to create well-defined flow paths through the elevation data. The “burning” process involves artificially reducing the elevation of DEM cells that are co-located with the NHD stream lines. The processing was done by the 8-digit Hydrologic Unit Code watershed. These data layers, along with the Natural Resources Conservation Service (NRCS) Watershed Boundary Dataset (WBD) (<http://www.nrcs.usda.gov/products/datasets/watershed/>), were processed through the ESRI ArcHydro Tools (ESRI, Inc., 2005) interface to produce all the data layers used in ILSS.

Using ArcHydro Tools, Version 1.1—a set of utilities developed to operate in the ArcGIS environment (<http://www.crwr.utexas.edu/gis/gishydro06/ArcHydro/ArcHydroTools/Doc/Arc%20Hydro%20Tools%20-%20Overview.pdf>)—49 processing units, based on the WBD 8-digit hydrologic units, were created and additional data layers were generated for each processing unit. These layers were developed to calculate basin characteristics used in the Illinois flood-peak regional

equations. Primary base-grid data layers that were created include catchments, flow accumulation, flow direction, and an artificial flow-path grid used to delineate drainage basins in the ILSS application. These layers then were used to create layers that control the StreamStats delineation, including AdjointCatchment, Catchment, DrainageLine, DrainagePoint, LongestFlowPathCat, and LongestFlowPathAdjCat. After all 49 processing units were processed, the global geodatabase was created. This database directs StreamStats as to how all the units interact. In addition, the NED was resampled to 100 m for use in the basin-length calculations (see programming for basin length (BL) in the Computer Code Development section).

GIS-data layers for average soil permeability (PermAvg), open water and herbaceous wetland (from which %Water is calculated), and hydrologic regions also are used for computing basin characteristics. The PermAvg grid was obtained by taking the arithmetic average of the high and low soil-permeability values from the STATSGO soil database (Natural Resources Conservation Service, 1993, <http://www.ncgc.nrcs.usda.gov/products/datasets/statsgo/>). The open-water and herbaceous-wetland grid was derived from the 1992 National Land Cover Data (NLCD) (<http://www.epa.gov/mrlc/nlcd.html>). The regions grid defines which regression equations to use after a user has delineated a drainage basin in ILSS; the regions used in ILSS are shown in figure 1.

The majority of Illinois Digital Raster Graphics (DRGs) used as base maps for this evaluation of the ILSS were purchased in 1998 from the EROS Data Center. Other DRGs not included in this purchase were obtained from several different Web sites, including the ISGS (<http://www.isgs.uiuc.edu/nsd/home/webdocs/drg/>), the Indiana Spatial Data Portal at Indiana University (<http://www.indiana.edu/~gisdata/>), and the Wisconsin Department of Natural Resources (WDNR) (<ftp://gomapout.dnr.state.wi.us/>). DRGs at scales of 1:24,000, 1:100,000, and 1:250,000 are displayed in ILSS to assist users with locating their sites.

Computer Code Development

XML algorithms are the computer codes implemented in the ArcHydro Tools parameter-configuration software to direct the computation of a basin characteristic when it is selected by a user. XML is understood by all modern Web browsers. The Web software used in StreamStats—ArcIMS—has many predefined XML algorithms for computing common parameters. For those basin characteristics for which XML algorithms were already defined, such as drainage area and stream slope, all that was needed was to make the fields specific to Illinois. XML algorithms also were customized for computing PermAvg, %Water, and the placeholder variable for the portion of the regression equation indicating the hydrologic region factor. For both PermAvg and %Water, the XML calculates an area-weighted value based on the delineated drainage basin.

A new XML algorithm was coded by ESRI to replicate the BL parameter derived by the BasinSoft program (Harvey

and Eash, 1996) and used in the Illinois regional regression equations. This algorithm is discussed below.

Several definitions of BL exist in the literature. The BasinSoft definition of BL states that it is measured, in miles, along a line areally centered through the basin polygon from the basin outlet to where the main-channel extension meets the basin divide (Harvey and Eash, 1996). These two end points, the basin-outlet and the basin-divide point, are located on the perimeter of the basin polygon and BasinSoft calculates the least-cost path through the polygon connecting the points to measure BL. BasinSoft prompts the user to manually digitize the main-channel extension to the divide, based on contours displayed on screen. This manual method of extending the main channel may result in different extensions by different users. For example, one user may identify the main channel as the major named stream for a given basin, whereas another user may identify the main channel as the longest flow path (LFP) for the basin. Similarly the basin divide (the point where the main channel, if extended, would cross the drainage basin boundary) could be interpreted as the nearest saddle point (a local low point on the basin boundary) or as the highest point on the upper basin boundary. For consistency, the endpoint used in the ILSS is based on the LFP extended to the intersection with the highest adjacent point on the basin boundary.

The least-cost path in the ILSS is determined by first creating a cost-surface wherein each grid cell is assigned a cost based on the inverse of the Euclidean distance from the basin boundary. The XML algorithm then computes the path resulting in the least-cost path from the basin divide LFP endpoint to the outlet using this cost surface. The grid-cell size used in BasinSoft was 100.12 ft, whereas in StreamStats, computer processing limitations required that the cell-size be increased to 100 m.

The final step in the XML programming was to incorporate adjustments to the StreamStats-computed basin characteristics to make them more comparable with the original basin characteristics determined from BasinSoft. These adjustments consisted of a power equation for the slope and a linear equation for the basin length. The determination of the coefficients for these equations is discussed in the Evaluation and Adjustment of Basin Characteristics section.

Illinois StreamStats Database

Information from the Illinois StreamStats database (StreamStatsDB) is available through the StreamStats GageInfo tool. The database was designed to be populated with site-specific information for streamflow-gaging stations and other data-collection stations, such as basin and climatic characteristics (Physical Characteristics) and streamflow statistics. The current (2010) StreamStatsDB implementation in ILSS was developed by the USGS StreamStats team by importing station descriptions and streamflow statistics from an obsolete database known as the Basin Characteristic File (BCF) of the USGS Water Storage Retrieval (WATSTORE) System, the

4 Implementation and Evaluation of the Streamflow Statistics (StreamStats) Web Application for Illinois

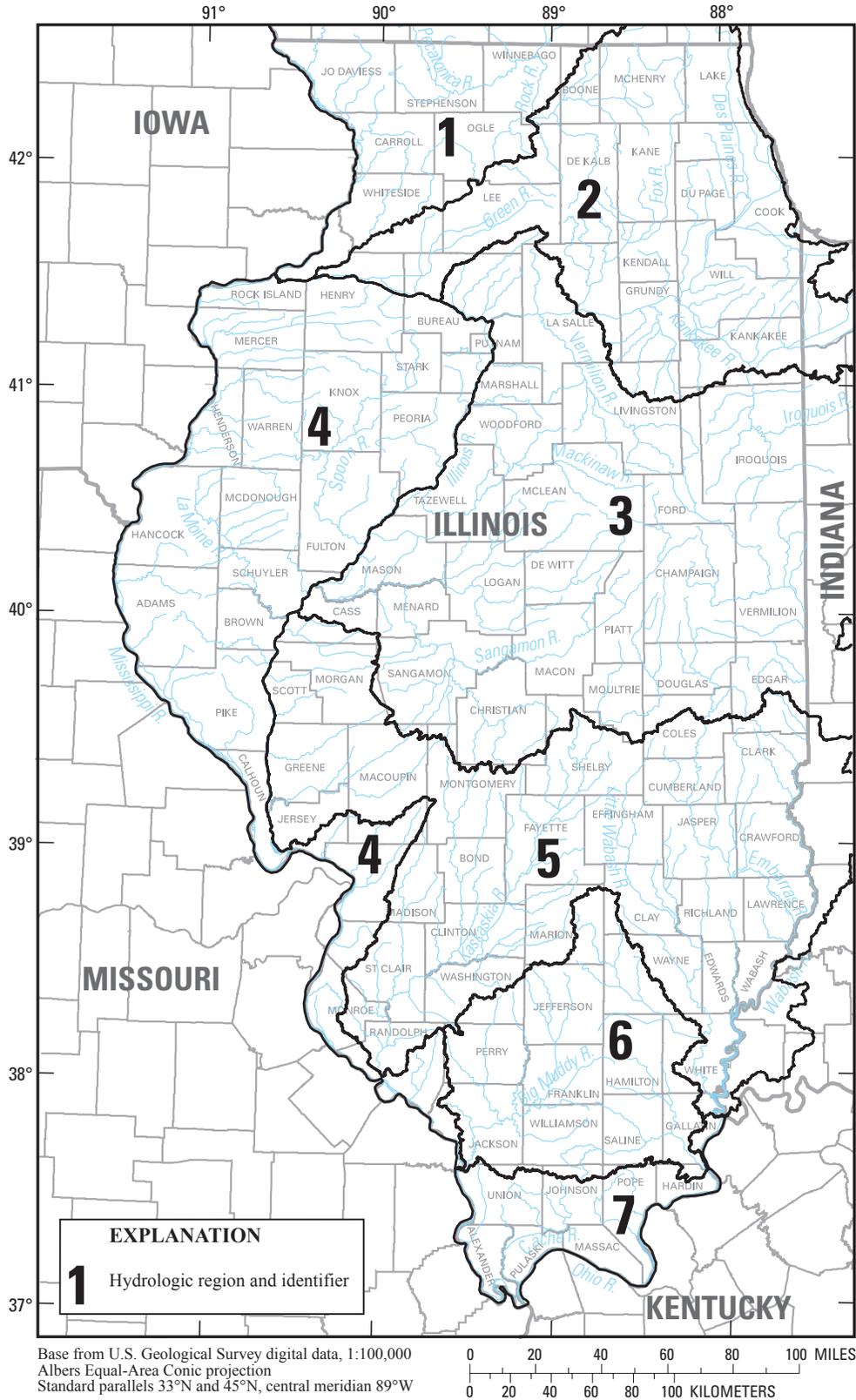


Figure 1. Hydrologic regions for flood-frequency regression equations of rural streams in Illinois (Soong and others, 2004).

contemporaneous water-information database. The BCF has not been updated since the 1990s, and no time tags were associated with any of the data entered into the BCF. As a result, the information from the BCF is likely to be at least 10 years old. Consequently, the values do not necessarily agree with more recently published values, or with the values computed by ILSS. The flow-duration and general flow statistics by StreamStatsDB were obtained from Wolock (2003a), and base-flow statistics were obtained from Wolock (2003b). The flood quantiles are from a variety of sources, including Soong and others (2004). The sources are documented on the StreamStats Web pages.

Evaluation and Adjustment of Basin Characteristics

The sets of BasinSoft- and GIS-map-derived basin characteristics used in the regression equations for determining flood quantiles published in Soong and others (2004) were compared with the approximately equivalent basin characteristics available from ILSS. To determine whether adjustments to ILSS computations of basin characteristics were required, basin characteristics were obtained from ILSS before the quantile regression-equation computations were implemented in ILSS. The ILSS values were obtained by identifying streamflow-gaging stations on the digital/base maps and selecting the nearest road crossing as the starting point for the ILSS watershed delineation. The percent and absolute differences between the two data sets for each basin characteristic were statistically analyzed by use of parametric and non-parametric significance tests. The distribution of the differences data set generally was not perfectly normal; therefore, non-parametric tests may be preferred to detect significant differences in the two data sets. Both parametric and non-parametric significance tests were computed, because the paired t-test is often considered sufficiently robust to detect differences even where normality in the data sets is moderately violated (Berk and Carey, 2004). The distributions of the differences are shown in the Sensitivity of Flood-Peak Quantiles section. The distributions generally were mildly skewed and (or) too leptokurtic (most values near the mean with few extreme values) to be considered normal.

All basin characteristics were tested with both the parametric paired t-test and the non-parametric Wilcoxon signed-rank test (WSRT). Because of the large range in values, the differences of the log-transformed basin characteristics were also tested for drainage area, slope, %Water, and PermAvg. Of the original 288 streamflow-gaging stations used in the flood-frequency regression analysis, two drainage basins were in the Lake Michigan watershed, which is not implemented in ILSS, and three additional drainage basins were found to have errors in the original data set, so the tests utilized 283 streamflow-gaging stations. Basin Length, which is applied only in the

region 4 part of the State, was tested at 47 streamflow-gaging stations. The differences of the log-transformed values are not reported because the transformation was not suitable for the relatively small range and linear relation between the ILSS and published values for Basin Length. The results of the tests are shown in table 1.

Paired t-test p-values greater than 0.05 indicate that the mean difference between the data sets is not statistically significant at the 0.05 significance level (referred to hereafter as the “95-percent confidence level”) (Schlotzhauer and Littel, 1987). Wilcoxon signed-rank test p-values greater than 0.05 indicate that the median difference between the data sets is not statistically significant at the 95-percent confidence level (Helsel and Hirsch, 2002). Results of the tests and subsequent adjustments (where performed) are described below.

Drainage Basin Area, Percent Water, Average Soil Permeability

Mean differences between the mean values published in Soong and others (2004) and those obtained from ILSS for drainage area (DA) and percentage of open water and herbaceous wetland (%Water) were not statistically significant. The scatter plot in figure 2 demonstrates the relation between the DA values published in Soong and others (2004) and those available from ILSS on logarithmically scaled axes. The same relation is shown on arithmetically scaled axes in figure 3. The scatter plots for PermAvg and %Water are not shown, because these basin characteristics are derived directly from the watershed drainage area. Differences between the Soong and others (2004) published values and the values computed by ILSS as percent differences are shown in figure 4. Percent difference is computed as $100 * (2004_DA - ILSS_DA) / 2004_DA$, where 2004_DA is the published value in Soong and others (2004) and ILSS_DA is the ILSS value. The same data are plotted on truncated scales in figures 5 and 6 to better illustrate the relative size of the difference for smaller watersheds. The percent differences are larger for smaller watersheds, especially those under 0.5 mi². A major reason for this occurrence is that, for a given difference between the values, the percent difference increases as the magnitude of the initial value decreases. For example, a difference of 0.1 mi² is 1 percent where the initial drainage area is 10 mi², but the difference is 10 percent where the drainage area is 1 mi².

It should be noted that the correctness of the drainage areas was not determined. In some cases, ILSS may compute a more correct drainage area than was determined by manual methods. These values were obtained from ILSS without any user intervention. In practice, the user can use the EditBasin tool to ensure that the watershed is properly represented with consideration of local hydraulic flow controls such as roads and culverts. The computed basin characteristics and flood quantiles then will be produced for the edited watershed.

6 Implementation and Evaluation of the Streamflow Statistics (StreamStats) Web Application for Illinois

Table 1. Differences between the published basin characteristics (Soong and others, 2004) and the Illinois StreamStats (ILSS) basin characteristics (published value minus ILSS value).

[n, number of paired observations; diffs, differences; percent diffs, 100 times differences divided by published value; WSRT, Wilcoxon signed-rank test; DA, drainage area in square miles; SL, slope; SLadj, the adjusted StreamStats SL; BL, basin length; BLadj, adjusted BL; %Water, percent open water and herbaceous wetland; PermAvg, areally weighted average of permeability; -, not applicable]

Statistic	DA	SL	SLadj	BL	BLadj	%Water	PermAvg
Statistics for drainage-basin characteristic							
<i>n</i>	283	283	283	47	47	283	283
Mean of diffs	-0.12357	-01.56415	0.782558	0.4897	0.00037	-0.02103	0.006958
Mean of percent diffs	.042	-8.569	-4.225	4.036	-6.783	-486	.253
Median of diffs	-.008	.047	.121	.130	-0.235	.000	.000
Median of percent diffs	-.019	.783	2.542	1.716	-2.547	.000	.000
t-test p-value ^a	.2022	.0222	.2436	.0464	.9988	.4708	.0436
WSRT p-value ^b	.0897	.2882	.0057	.0089	.1109	.9030	.8578
Test statistics for log-transformed drainage-basin characteristic							
t-test p-value ^a	0.3415	0.0240	0.9982	-	-	0.3816	0.0724
WSRT p-value ^b	.8978	.5861	.0328	-	-	.9737	.7965

^a p-values greater than 0.05 indicate that there is not a statistically significant mean difference between the published (Soong and others, 2004) values and the ILSS values.

^b p-values greater than 0.05 indicate that there is not a statistically significant median difference between the published (Soong and others, 2004) values and the ILSS values.

The paired t-test p-value (0.0436) indicated that the difference between the 2004 PermAvg and the ILSS-computed PermAvg was statistically significant, but the Wilcoxon signed-rank test p-value (0.8578) indicated that the difference was not statistically significant. The average mean difference between PermAvg determined using BasinSoft (BS) and ILSS was very small (0.25 percent), resulting in a potential adjustment-equation slope for PermAvg that was very close to 1.0 (1.016) and the intercept that was close to zero (-0.0156). The mean difference of the log-transformed values was not statistically significant. Furthermore, the sensitivity of the computed flood quantiles to differences in PermAvg also was very small, as discussed in the Sensitivity of Flood-Peak Quantiles section; therefore, no adjustment equation was applied to the PermAvg basin characteristics.

The paired t-test p-values indicated that differences in the mean values of BL (0.0464) and slope (SL) (0.0222) were significant. The Wilcoxon signed-rank test p-values indicated that the differences between the median values of BL (p-value=0.0089) were significant, but differences for SL (p-value=0.2882) were not significant. These characteristics were further analyzed to see whether an adjustment factor could be determined using simple linear regression as described in the next two sections.

Basin Length

BasinSoft (version 1.1, Harvey and Eash, 1996) was used to determine the BL values that were used in the current (2010) update of techniques for estimating flood quantiles for

rural streams in Illinois (Soong and others, 2004). The definitions of BL differ in two regards between ILSS and BasinSoft: (1) the intersection of the LFP with the watershed divide is used as the second endpoint of the least-cost path from the watershed outlet in ILSS, rather than the intersection of the user-determined main channel with the watershed divide that is used in BasinSoft; and (2) the computational-grid spacing was changed in the ILSS to from 10 to 100 m because of computer-processing limitations.

The BL differences computed using the original 10-m computation grid in ILSS were not statistically significant from those computed by BasinSoft, but the computational time was prohibitive; therefore, the 100-m grid values with an applied-adjustment equation is used in ILSS.

A linear equation was found to adjust the set of ILSS values of BL computed with the 100-m computational grid so that the differences between the adjusted BL values (BLadj) and the BasinSoft computed BL values (BS_BL) were not statistically significant. The equation was

$$BLadj = 1.0164 * ILSS_BL + 0.2364$$

$$R\text{-squared} = 0.9915$$

where ILSS_BL is the value of BL computed by ILSS using the 100-m grid as discussed in the section on BL (see fig. 7).

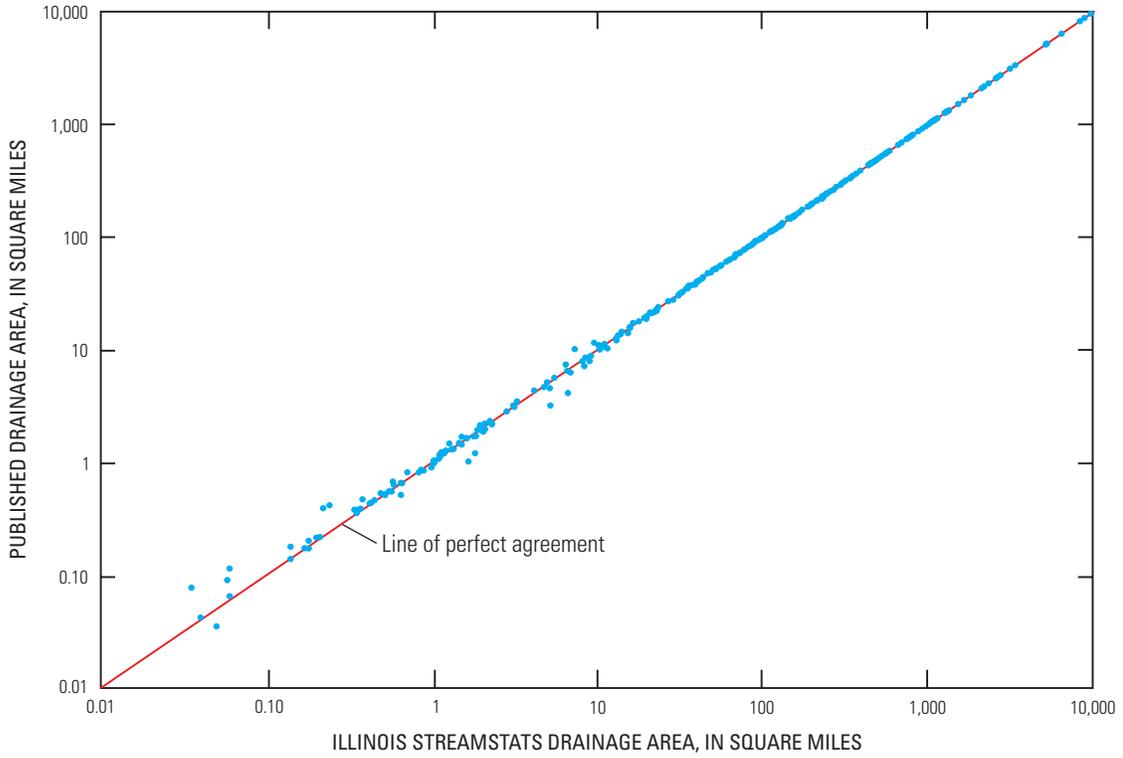


Figure 2. Relation between Illinois StreamStats and published drainage areas (Soong and others, 2004)—logarithmically scaled axes.

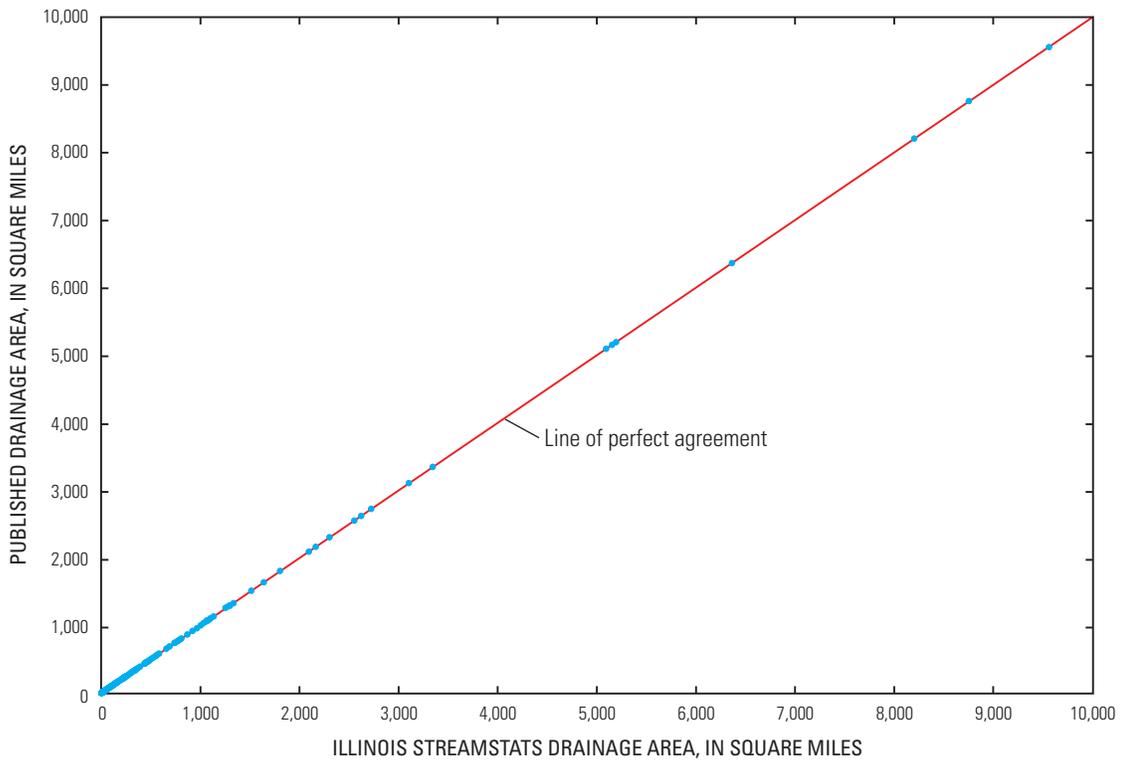


Figure 3. Relation between Illinois StreamStats and published drainage areas (Soong and others, 2004)—arithmetically scaled axes.

8 Implementation and Evaluation of the Streamflow Statistics (StreamStats) Web Application for Illinois

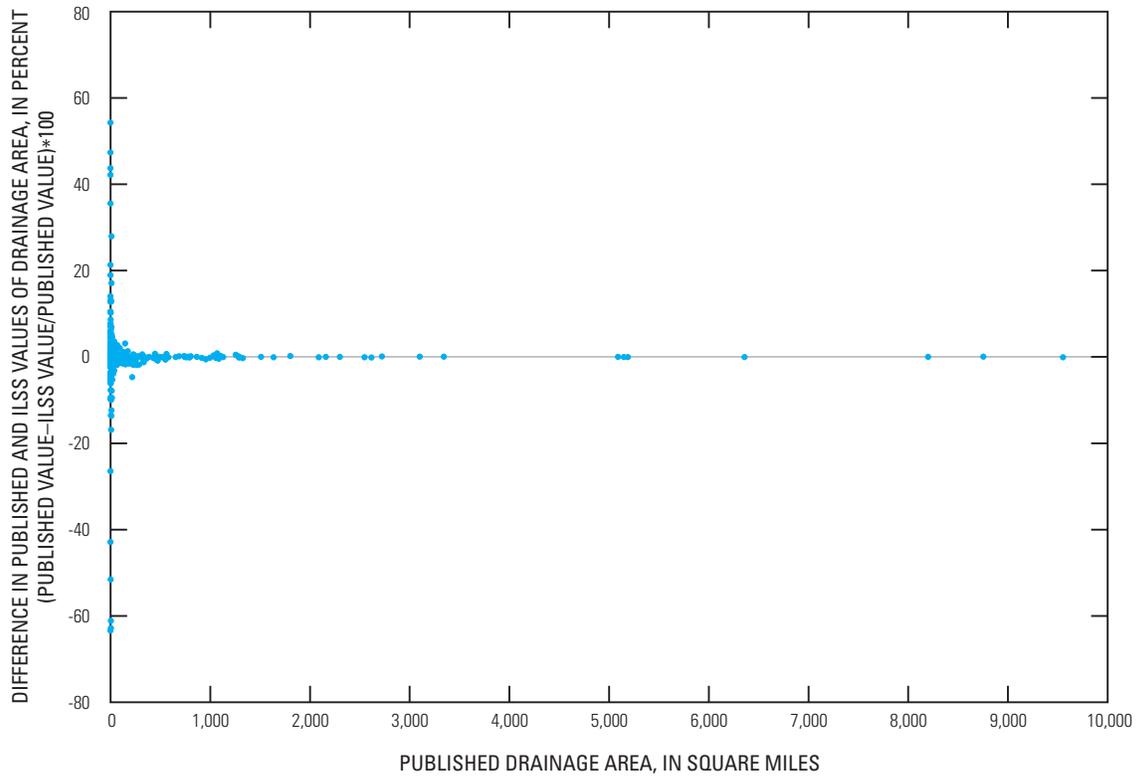


Figure 4. Relation between percent differences in published and Illinois StreamStats (ILSS) values of drainage areas and published drainage areas (Soong and others, 2004)—full scale.

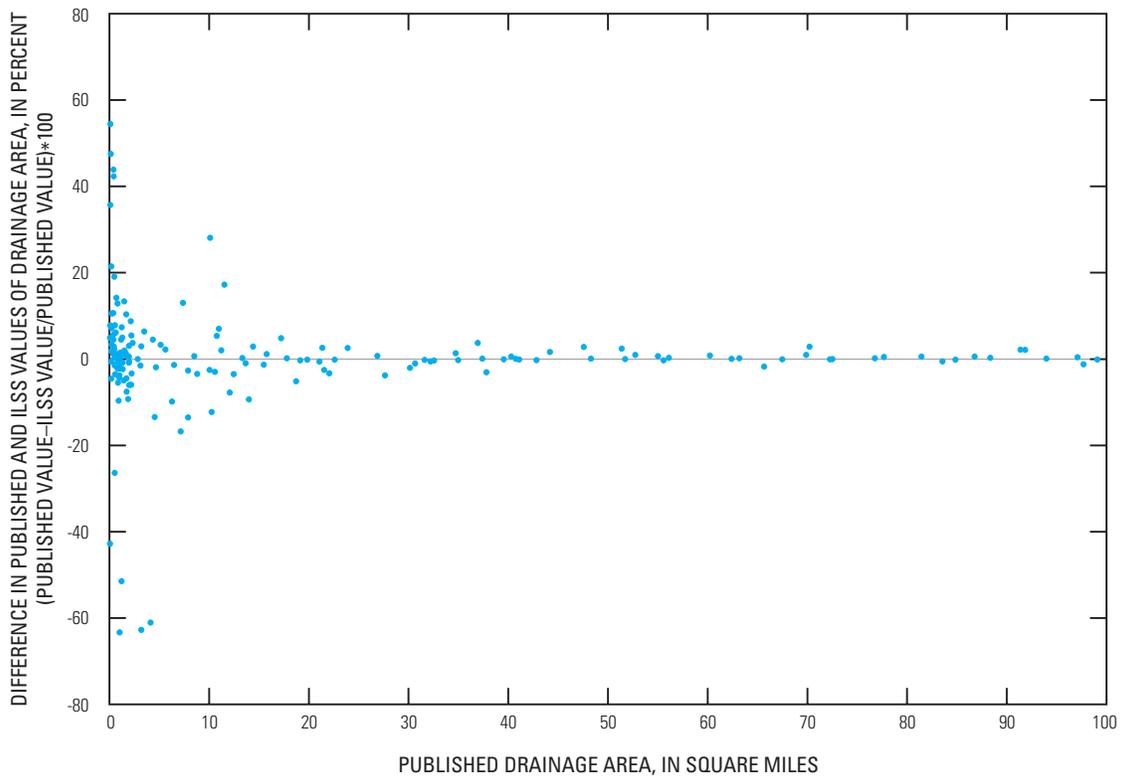


Figure 5. Relation between percent differences in published and Illinois StreamStats (ILSS) values of drainage areas and published drainage areas (Soong and others, 2004)—scale truncated to 100 square miles.

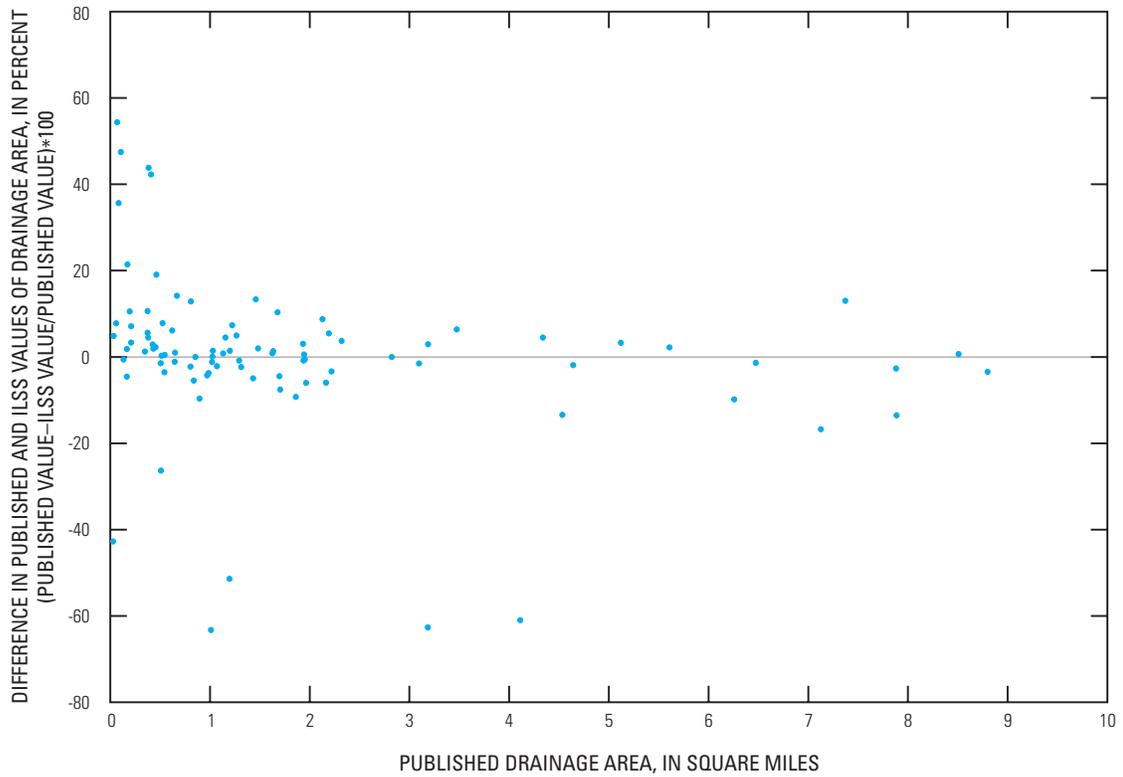


Figure 6. Relation between percent difference in published and Illinois StreamStats (ILSS) values of drainage areas and published drainage areas (Soong and others, 2004)—scale truncated to 10 square miles.

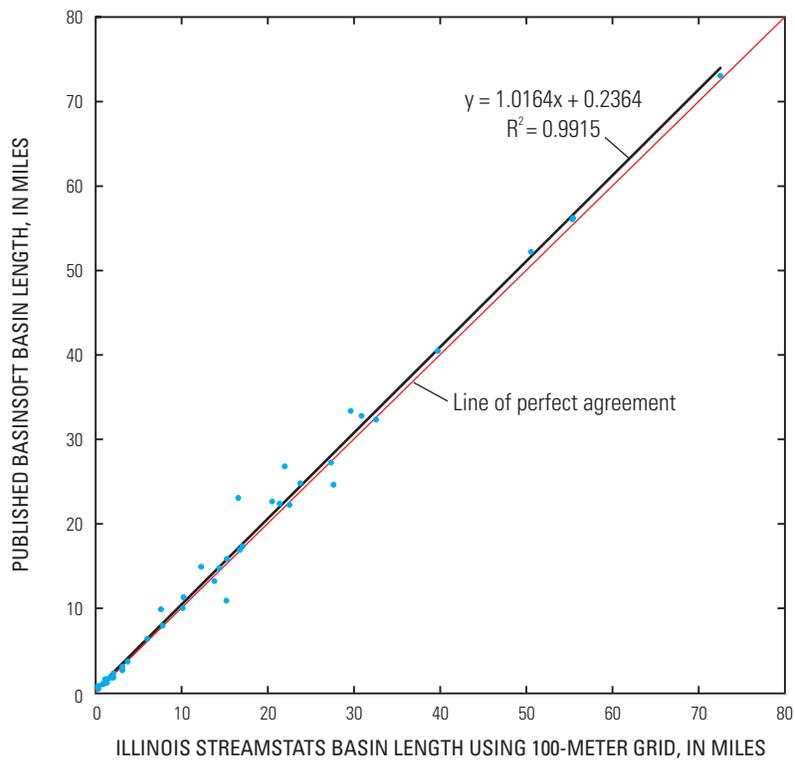


Figure 7. Relation between Illinois StreamStats basin length using 100-meter grid and the published BasinSoft basin length (Soong and others, 2004).

Stream Slope

BasinSoft (version 1.1, Harvey and Eash, 1996) was used to determine the slope basin characteristic that was used in the current (2010) update of techniques for estimating flood quantiles for rural streams in Illinois (Soong and others, 2004). The main channel was determined by using the 1:100,000-scale NHD and manually extending the apparent main channel to an intersection with the watershed divide, an adjacent saddle point between peaks. The slope was determined by determining the elevation, in feet, at points of 10 and 85 percent along the main channel from the outlet, and dividing the difference in elevation by 75 percent of the total main-channel length (which is the distance between the 10 and 85 percent points), in miles. The StreamStats method is based on an ArcIMS XML algorithm, the SL_{10-85} , which determines the longest flow path from the outlet to the divide by using the DEM to determine the flow path from the watershed outlet, extending up the longest continuous flow path to the adjacent peak point on the watershed divide. The slope is determined by dividing the difference in elevation, in feet, at 10 and 85 percent of the distance from the outlet to the intersection of the LFP with the watershed divide, by 75 percent of the total LFP length (the distance between the 10 and 85 percent points), in miles.

The distribution of the differences in SL displayed a skew, with larger values of SL determined by ILSS than those determined using BasinSoft. The t-test for differences in SL (table 1) indicated that differences in the means for both the values and log-transformed values were statistically significant (p-value of 0.0222 and 0.0240, respectively); thus, a correction was tried. The best-fit linear-regression equation (determined

by linear regression on log-transformed variables—see fig. 8) was determined to be

$$SL_{adj} = 1.0767 * (ILSS_SL) ** 0.9486$$

$$R\text{-squared} = 0.9529$$

where

- SL_{adj} is ILSS slope adjusted to be closer to the BasinSoft slope, and
- ILSS_SL is the ILSS slope.

The residuals from the adjusted values as percent differences were plotted in relation to the 2004 values for drainage area (2004_DA) and BasinSoft slope (BS_SL) and no pattern was apparent (figs. 9 and 10); therefore, no regression relation utilizing these basin characteristics was considered. The adjustment equation removed the statistical significance of the differences determined by the paired t-test, but the adjusted values were significantly different from the published values under the Wilcoxon signed-rank test. Similar results were found for the log-transformed values. The mean percent difference between the published values and the adjusted values was -4.22, compared to -8.57 for the mean difference for the unadjusted values; however, the median percent difference decreased from -0.783 to -2.54, indicating a slightly more skewed distribution after adjustment. The adjusted distribution was found to be preferable according to the t-test assumption of a normal distribution. Figure 8 demonstrates that there is a skew in the distribution of the ILSS slopes compared to the BasinSoft slopes, but the effect of this on the overall distribution and relevance of the test was considered nominal.

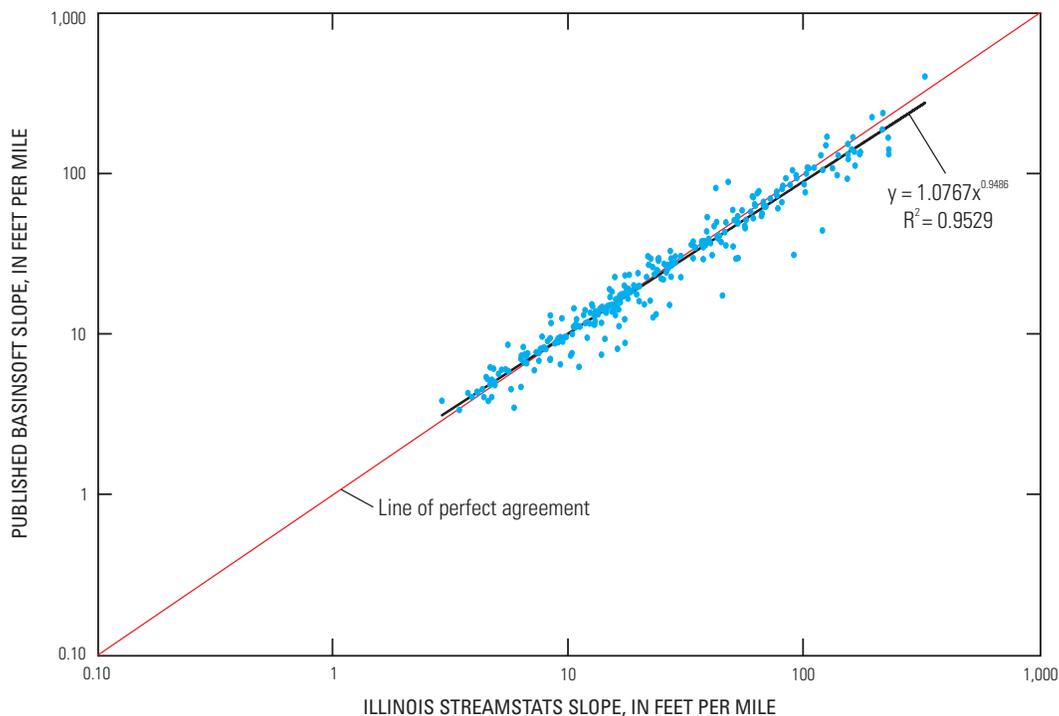


Figure 8. Relation between Illinois StreamStats and published BasinSoft slope (Soong and others, 2004).

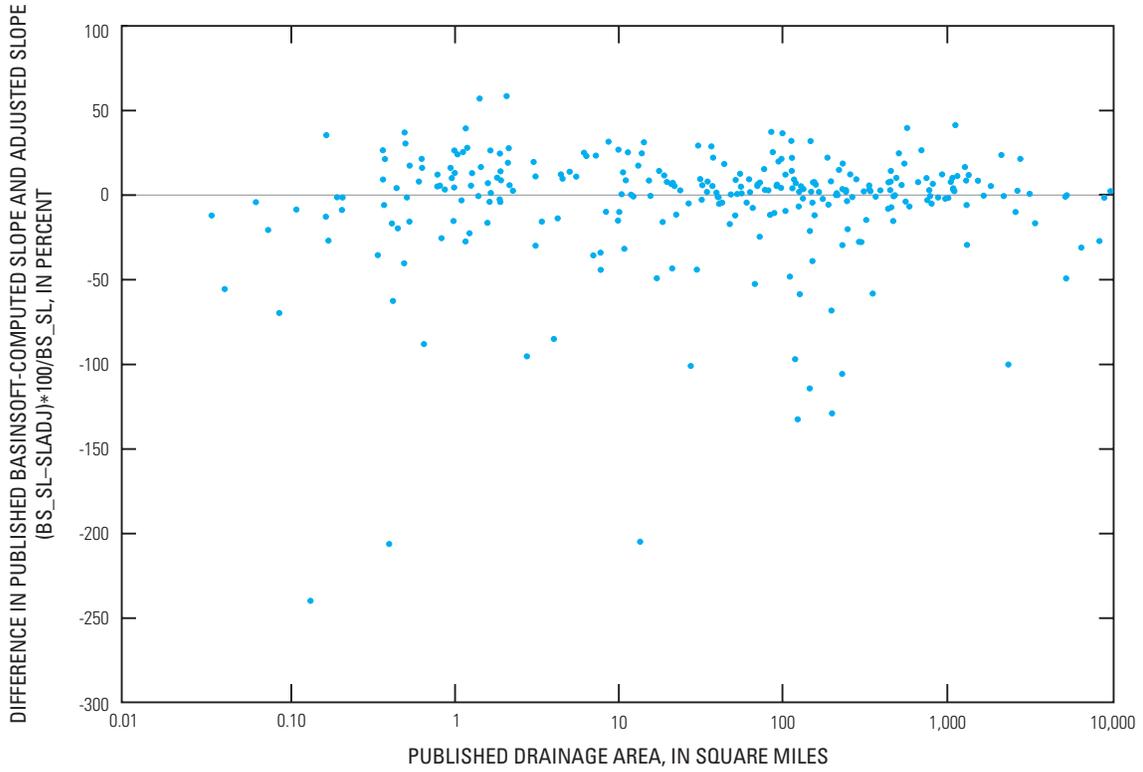


Figure 9. Relation between adjusted slope residual (published BasinSoft-computed stream slope (BS_SL) minus adjusted slope (SLadj), expressed as a percent $[(BS_SL-SLadj)*100/BS_SL]$, and published drainage area (Soong and others, 2004).

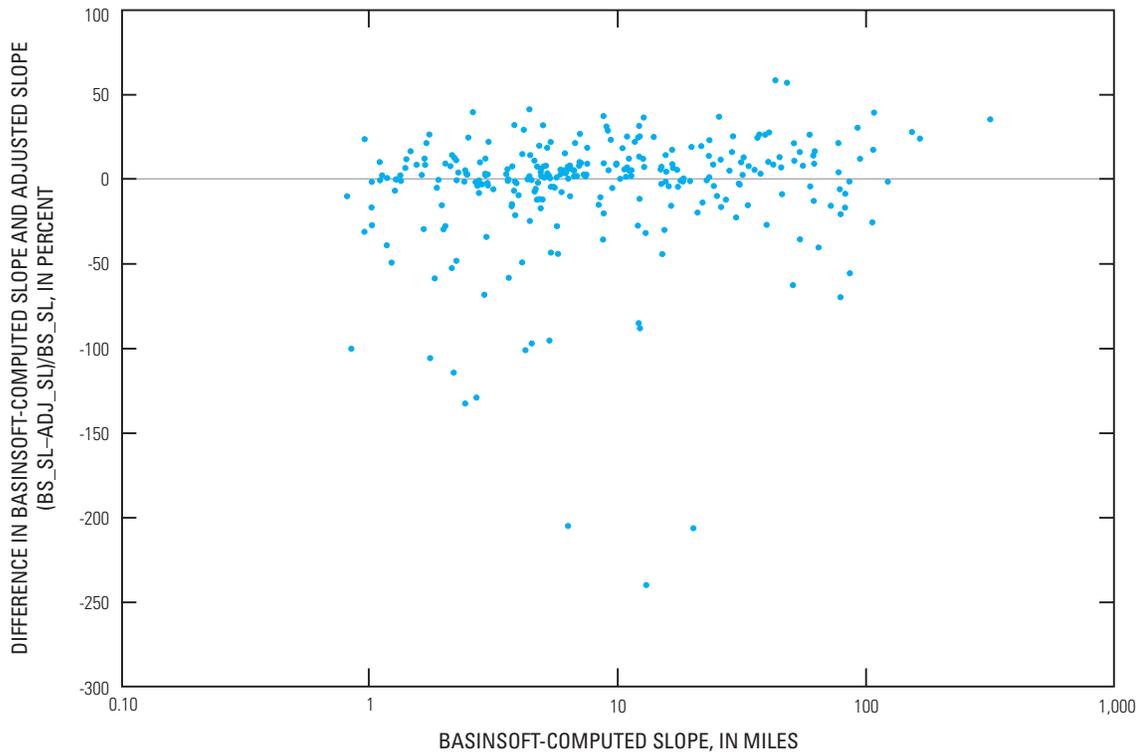


Figure 10. Relation between adjusted slope residual (published BasinSoft-computed stream slope (BS_SL) minus adjusted slope (SLadj), expressed as a percent $[(BS_SL-SLadj)*100/BS_SL]$, and BasinSoft slope.

The channel lengths of the BasinSoft main channel (MCL) and the ILSS LFP were available for 47 drainage basins in region 4. Their magnitudes did not differ significantly, although the LFP generally was longer than the MCL, indicating that the larger slopes in ILSS may not be a direct result of the length difference, but rather the elevation differences as computed at points 10 and 85 percent along the LFP from the outlet. The automated method of selecting the intersection of the LFP with the basin divide may tend to seek the high point on the divide, whereas the manual determination of the main-channel extension to the basin divide for Illinois tended to select the saddle; therefore, the ILSS may result in a higher 85-percent elevation and a larger value of SL.

The regression equation did not completely remove the statistically significant difference in the two data sets according to the Wilcoxon signed-rank test; however, the sample of 170 adjusted slopes obtained from ILSS for the quantile test reported in the next section was not significantly different according to the Wilcoxon signed-rank test nor the paired t-test and also was not significantly different when sorted by region and tested.

Evaluation of Flood-Peak Quantiles

In general, the user cannot assume that the flood-peak quantiles computed by ILSS and those published in Soong and others (2004) are identical; however, a reasonable sample of the flood-peak quantiles computed by ILSS should be unbiased compared with the published regression equation flood-peak quantiles, and the mean of the differences should not be significantly different from zero at the 95-percent confidence level.

An approximately random sample of 170 streamflow-gaging stations was tested for all quantiles to verify the ILSS application. The sample was selected by numbering the observations and using a random number set to select the sample. The sample was increased by adding all locations in regions 1 and 7, because of the small sample size in those regions. All regions also had at least one station crossing the 8-digit Hydrologic Unit Code processing units. All stations that had drainage basins crossing more than one hydrologic region were removed from this comparison, because the quantiles published in Soong and others (2004) used only the regional regression equation applicable at the streamflow station and did not weight by percent area in separate hydrologic regions. ILSS computes the quantiles in each region separately and provides a weighted-quantile estimate as well as each regional quantile estimate in the output. For the test of quantile estimates, the selection of starting points was determined by obtaining a list of the latitude and longitude to the nearest second of the intersection of the NHD stream network with the digitized drainage divide. The nearest road crossing to this intersection was selected in ILSS to determine the basin characteristics and compute the flood quantiles.

Because of the very large range in quantile estimates, the significance testing was done on three datasets: the differences between the published quantile and the ILSS quantile (published quantile–ILSS quantile) referred to hereafter as the simple differences; the proportional differences resulting from dividing the simple differences by the published quantiles ($[\text{published quantile} - \text{ILSS quantile}] / \text{published quantile}$); and the differences of the log-transformed quantiles ($\log[\text{published quantile}] - \log[\text{ILSS quantile}]$). The results of the significance testing using the proportional differences and the differences of the log-transformed quantiles are presented in tables 2–9, along with the mean and median statistics for the proportional differences. The significance testing utilized both the t-test and the Wilcoxon Signed Rank Test. For the data set consisting of simple differences, neither test indicated any significant difference between the means in either in the overall sample or the individual regions. However, this test was not considered strictly valid because of the large range in quantiles values; therefore, the simple differences results are not included in the following tables. The logarithmic transformations address the problem of the large range of values observed in the simple differences, whereas the statistics for the proportional differences provide a practical measure of the differences. The proportional differences are related to the differences of the log-transformed quantiles (A) by the following equation:

$$\frac{(Q_T^{2004} - Q_T^{ILSS})}{Q_T^{2004}} = 1 - e^{-A}$$

where

$$A = \log\left(\frac{Q_T^{2004}}{Q_T^{ILSS}}\right) = \log(Q_T^{2004}) - \log(Q_T^{ILSS})$$

Q	is the flood quantile,
T	is the t-year recurrence interval
<i>ILSS</i>	is the ILSS-computed value, and
2004	is the value published in Soong and others (2004).

Table 2 displays the results of the tests on the proportional and log-transformed quantiles for the entire sample. It was found that there was no statistically significant difference for any quantile when all the regions were grouped together. A scatter plot of the published 100-year flood quantiles with those obtained from ILSS is illustrated in figure 11. Analyzing the regions separately resulted in a small but significant difference in the flood quantiles for region 1 only (table 3); the tests for the other regions did not indicate a statistically significant difference (tables 4–9).

Table 2. Comparison of published (Soong and others, 2004), Q_T^{2004} , and Illinois StreamStats (ILSS), Q_T^{ILSS} , flood quantiles as proportional differences— $[(Q_T^{2004} - Q_T^{ILSS}) / Q_T^{2004}]$ and differences of the log-transformed quantiles ($\log [Q_T^{2004}] - \log [Q_T^{ILSS}]$)—for selected streamflow-gaging stations, all regions.

[*n*, number of paired observations; WSRT, Wilcoxon signed-rank test]

Statistic	Recurrence interval						
	2-year	5-year	10-year	25-year	50-year	100-year	500-year
Statistics for proportional differences							
<i>n</i>	170	170	170	170	170	170	170
Mean	-0.01260	-0.01455	-0.01666	-0.01707	-0.01923	-0.01742	-0.01962
Median	.00632	.00863	.00835	.00803	.00819	.00942	.00908
t-test p-value ^a	.2563	.2272	.1829	.1909	.1536	.2035	.1739
WSRT p-value ^b	.2053	.2195	.3244	.2825	.3678	.2518	.2619
Test statistics for differences of the log-transformed quantiles							
t-test p-value ^a	0.6732	0.6574	0.5804	0.6263	0.5474	0.6930	0.6482
WSRT p-value ^b	.1531	.1759	.2459	.2044	.2739	.1835	.1919

^a p-values greater than 0.05 indicate that there is not a statistically significant mean difference between the published (Soong and others, 2004) values and the ILSS values.

^b p-values greater than 0.05 indicate that there is not a statistically significant median difference between the published (Soong and others, 2004) values and the ILSS values.

Table 3. Comparison of published (Soong and others, 2004), Q_T^{2004} , and Illinois StreamStats (ILSS), Q_T^{ILSS} , flood quantiles as proportional differences— $[(Q_T^{2004} - Q_T^{ILSS}) / Q_T^{2004}]$ and differences of the log-transformed quantiles ($\log [Q_T^{2004}] - \log [Q_T^{ILSS}]$)—for streamflow-gaging stations, sorted by region: Region 1.

[*n*, number of paired observations; WSRT, Wilcoxon signed-rank test]

Statistic	Recurrence interval						
	2-year	5-year	10-year	25-year	50-year	100-year	500-year
Statistics for proportional differences							
<i>n</i>	21	21	21	21	21	21	21
Mean	0.037587	0.039758	0.038617	0.037923	0.038068	0.040158	0.042711
Median	.02905	.02723	.02716	.02899	.0299	.03137	.03636
t-test p-value ^a	.0228	.025	.0381	.0487	.0525	.0473	.0422
WSRT p-value ^b	.0149	.0187	.0209	.035	.0425	.0317	.0286
Test statistics for differences of the log-transformed quantiles							
t-test p-value ^a	0.0182	0.02	0.029	0.0366	0.0393	0.035	0.0308
WSRT p-value ^b	.0149	.0187	.0187	.0167	.0258	.0187	.0209

^a p-values greater than 0.05 indicate that there is not a statistically significant mean difference between the published (Soong and others, 2004) values and the ILSS values.

^b p-values greater than 0.05 indicate that there is not a statistically significant median difference between the published (Soong and others, 2004) values and the ILSS values.

Table 4. Comparison of published (Soong and others, 2004), Q_T^{2004} , and Illinois StreamStats (ILSS), Q_T^{ILSS} , flood quantiles as proportional differences— $[(Q_T^{2004} - Q_T^{ILSS}) / Q_T^{2004}]$ and differences of the log-transformed quantiles ($\log [Q_T^{2004}] - \log [Q_T^{ILSS}]$)—for selected streamflow-gaging stations, sorted by region: Region 2.

[*n*, number of paired observations; WSRT, Wilcoxon signed-rank test]

Statistic	Recurrence interval						
	2-year	5-year	10-year	25-year	50-year	100-year	500-year
Statistics for proportional differences							
<i>n</i>	40	40	40	40	40	40	40
Mean	-0.06117	-0.06990	-0.07504	-0.07625	-0.08452	-0.08074	-0.09006
Median	.01508	.01743	.01511	.02321	.01903	.02566	.02334
t-test p-value ^a	.108	.0938	.0817	.0912	.0708	.0887	.0733
WSRT p-value ^b	.9055	.9579	.8373	.979	.9369	.9673	.9474
Test statistics for differences of the log-transformed quantiles							
t-test p-value ^a	0.1948	0.1788	0.1601	0.1891	0.1471	0.1948	0.1662
WSRT p-value ^b	.9579	.9895	.902	.9055	.979	.902	.9055

^a p-values greater than 0.05 indicate that there is not a statistically significant mean difference between the published (Soong and others, 2004) values and the ILSS values.

^b p-values greater than 0.05 indicate that there is not a statistically significant median difference between the published (Soong and others, 2004) values and the ILSS values.

Table 5. Comparison of published (Soong and others, 2004), Q_T^{2004} , and Illinois StreamStats (ILSS), Q_T^{ILSS} , flood quantiles as proportional differences— $[(Q_T^{2004} - Q_T^{ILSS}) / Q_T^{2004}]$ and differences of the log-transformed quantiles ($\log [Q_T^{2004}] - \log [Q_T^{ILSS}]$)—for selected streamflow-gaging stations, sorted by region: Region 3.

[*n*, number of paired observations; WSRT, Wilcoxon signed-rank test]

Statistic	Recurrence interval						
	2-year	5-year	10-year	25-year	50-year	100-year	500-year
Statistics for proportional differences							
<i>n</i>	34	34	34	34	34	34	34
Mean	-0.00761	-0.01171	-0.01259	-0.01438	-0.01595	-0.01472	-0.01481
Median	.00314	-.00052	.00212	.00102	-.00199	.00289	.00162
t-test p-value ^a	.6672	.5491	.5304	.4927	.4606	.5052	.5168
WSRT p-value ^b	.8123	.7065	.6684	.6094	.5997	.6557	.7442
Test statistics for differences of the log-transformed quantiles							
t-test p-value ^a	0.9094	0.8042	0.7915	0.7593	0.7296	0.7874	0.8089
WSRT p-value ^b	.8336	.7442	.6684	.6153	.6181	.6684	.757

^a p-values greater than 0.05 indicate that there is not a statistically significant mean difference between the published (Soong and others, 2004) values and the ILSS values.

^b p-values greater than 0.05 indicate that there is not a statistically significant median difference between the published (Soong and others, 2004) values and the ILSS values.

Table 6. Comparison of published (Soong and others, 2004), Q_T^{2004} , and Illinois StreamStats (ILSS), Q_T^{ILSS} , flood quantiles as proportional differences— $[(Q_T^{2004} - Q_T^{ILSS}) / Q_T^{2004}]$ and differences of the log-transformed quantiles ($\log [Q_T^{2004}] - \log [Q_T^{ILSS}]$)—for selected streamflow-gaging stations, sorted by region: Region 4.

[*n*, number of paired observations; WSRT, Wilcoxon signed-rank test]

Statistic	Recurrence interval						
	2-year	5-year	10-year	25-year	50-year	100-year	500-year
Statistics for proportional differences							
<i>n</i>	32	32	32	32	32	32	32
Mean	0.014143	0.01712	0.018984	0.01889	0.022862	0.021554	0.022185
Median	.006028	.008519	.008731	.007102	.00985	.007138	.006338
t-test p-value ^a	.2553	.2365	.2255	.2634	.1937	.2405	.2619
WSRT p-value ^a	.1892	.2605	.2851	.3452	.246	.3067	.3388
Test statistics for differences of the log-transformed quantiles							
t-test p-value ^a	0.1914	0.1675	0.1552	0.1768	0.1256	0.1548	0.1639
WSRT p-value ^b	.1755	.2373	.2685	.3067	.2079	.2709	.3203

^a p-values greater than 0.05 indicate that there is not a statistically significant mean difference between the published (Soong and others, 2004) values and the ILSS values.

^b p-values greater than 0.05 indicate that there is not a statistically significant median difference between the published (Soong and others, 2004) values and the ILSS values.

Table 7. Comparison of published (Soong and others, 2004), Q_T^{2004} , and Illinois StreamStats (ILSS), Q_T^{ILSS} , flood quantiles as proportional differences— $[(Q_T^{2004} - Q_T^{ILSS}) / Q_T^{2004}]$ and differences of the log-transformed quantiles ($\log [Q_T^{2004}] - \log [Q_T^{ILSS}]$)—for selected streamflow-gaging stations, sorted by region: Region 5.

[*n*, number of paired observations; WSRT, Wilcoxon signed-rank test]

Statistic	Recurrence interval						
	2-year	5-year	10-year	25-year	50-year	100-year	500-year
Statistics for proportional differences							
<i>n</i>	20	20	20	20	20	20	20
Mean	-0.05530	-0.05519	-0.06051	-0.06277	-0.06303	-0.06281	-0.06382
Median	-.00446	-.00677	-.00852	-.00895	-.00985	-.00923	-.00997
t-test p-value ^a	.0983	.0966	.0832	.0752	.0748	.079	.0788
WSRT p-value ^b	.2069	.2837	.2121	.1925	.1564	.165	.1819
Test statistics for differences of the log-transformed quantiles							
t-test p-value ^a	0.1	0.098	0.0832	0.0745	0.0741	0.08	0.0811
WSRT p-value ^b	.2247	.2837	.2121	.1992	.1819	.165	.1819

^a p-values greater than 0.05 indicate that there is not a statistically significant mean difference between the published (Soong and others, 2004) values and the ILSS values.

^b p-values greater than 0.05 indicate that there is not a statistically significant median difference between the published (Soong and others, 2004) values and the ILSS values.

Table 8. Comparison of published (Soong and others, 2004), Q_T^{2004} , and Illinois StreamStats (ILSS), Q_T^{ILSS} , flood quantiles as proportional differences— $([Q_T^{2004} - Q_T^{ILSS}] / Q_T^{2004})$ and differences of the log-transformed quantiles ($\log [Q_T^{2004}] - \log [Q_T^{ILSS}]$)—for selected streamflow-gaging stations, sorted by region: Region 6.

[*n*, number of paired observations; WSRT, Wilcoxon signed-rank test]

Statistic	Recurrence interval						
	2-year	5-year	10-year	25-year	50-year	100-year	500-year
Statistics for proportional differences							
<i>n</i>	15	15	15	15	15	15	15
Mean	0.014886	0.016165	0.014695	0.019418	0.015057	0.019857	0.017840
Median	.022989	.026059	.02451	.030965	.027231	.033419	.028037
t-test p-value ^a	.5268	.5245	.5811	.4846	.6006	.5022	.5634
WSRT p-value ^b	.3591	.3591	.391	.3258	.3303	.3258	.3591
Test statistics for differences of the log-transformed quantiles							
t-test p-value ^a	0.4279	0.4192	0.4633	0.375	0.4709	0.3832	0.4301
WSRT p-value ^b	.2769	.3028	.2958	.2676	.2769	.2676	.2769

^a p-values greater than 0.05 indicate that there is not a statistically significant mean difference between the published (Soong and others, 2004) values and the ILSS values.

^b p-values greater than 0.05 indicate that there is not a statistically significant median difference between the published (Soong and others, 2004) values and the ILSS values.

Table 9. Comparison of published (Soong and others, 2004), Q_T^{2004} , and Illinois StreamStats (ILSS), Q_T^{ILSS} , flood quantiles as proportional differences— $([Q_T^{2004} - Q_T^{ILSS}] / Q_T^{2004})$ and differences of the log-transformed quantiles ($\log [Q_T^{2004}] - \log [Q_T^{ILSS}]$)—for streamflow-gaging stations, sorted by region: Region 7.

[*n*, number of paired observations; WSRT, Wilcoxon signed-rank test]

Statistic	Recurrence interval						
	2-year	5-year	10-year	25-year	50-year	100-year	500-year
Statistics for proportional differences							
<i>n</i>	8	8	8	8	8	8	8
Mean	0.02548	0.02479	0.02104	0.02506	0.01963	0.024224	0.021489
Median	-.01248	-.01393	-.01675	-.01317	-.01878	-.01443	-.01801
t-test p-value ^a	.3994	.4024	.4656	.3804	.4877	.3966	.4435
WSRT p-value ^b	.8438	.8438	.8438	.7422	.7422	.7422	.7422
Test statistics for differences of the log-transformed quantiles							
t-test p-value ^a	0.3778	0.3782	0.4332	0.356	0.4515	0.3685	0.4108
WSRT p-value ^b	.7422	.7422	.7422	.7422	.7422	.7422	.7422

^a p-values greater than 0.05 indicate that there is not a statistically significant mean difference between the published (Soong and others, 2004) values and the ILSS values.

^b p-values greater than 0.05 indicate that there is not a statistically significant median difference between the published (Soong and others, 2004) values and the ILSS values.

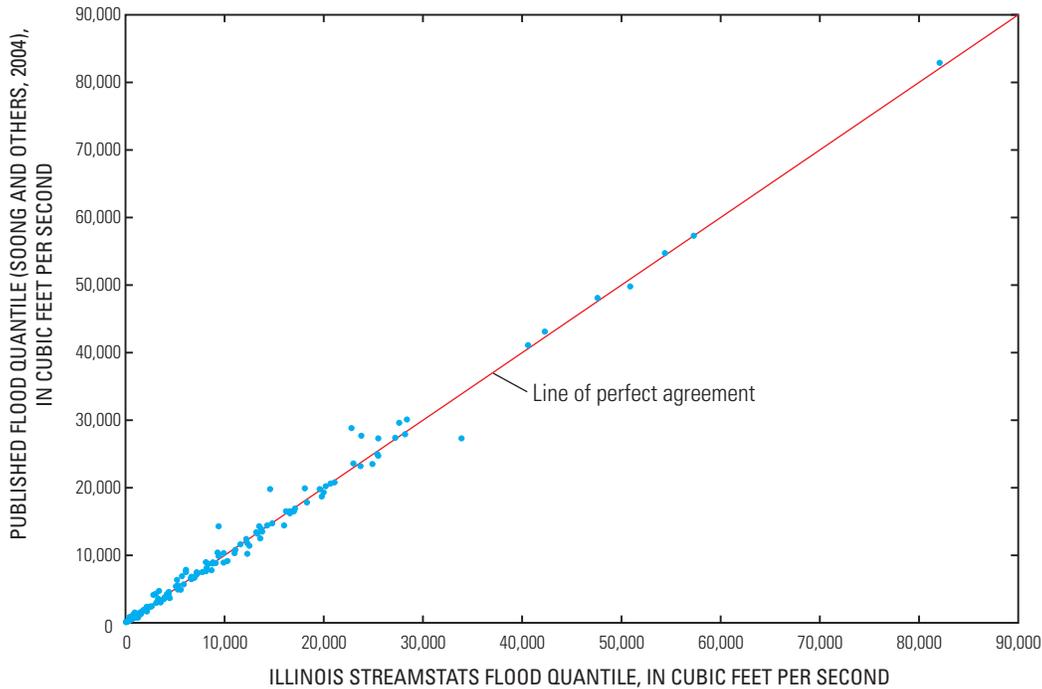


Figure 11. Relation between StreamStats and Soong and others (2004) 100-year flood quantiles.

Region 1 had a small statistically significant difference ranging from 3.8 percent for the 2-year flood-quantile estimate to 4.3 percent for the 500-year flood-quantile estimate at the 95-percent confidence level. All 21 streamflow-gaging stations were considered in the analysis, because of the few streamflow-gaging stations available in region 1. The statistical significance of the difference could be by chance, in light of the small size of the sample. In any case, the difference is small (less than one-tenth) compared to the average prediction errors of the 2- to 500-year regression-equation estimates in this region, which range from 39.5 to 54.9 percent respectively.

Test of Illinois StreamStats Usage Consistency

Typical road-crossing design-site locations were collected from the IDOT district offices and compiled into a test data set to test the consistency of ILSS usage. A list of 28 structures was compiled with latitude and longitude, description, and structure-identification number. Additional structures were distributed (not repeated) among the districts included for optional extra testing; the list was distributed to the IDOT

district offices. At least seven users returned completed tests, and the results were compiled. One structure-site description was found to be ambiguous and was removed from the results. Additionally, at least two users had some difficulty finding the correct site and selected locations that obviously were incorrect. These results also were removed from the final results. The final results included 6 or 7 results at 27 locations. The average maximum deviation from the mode value of the 100-year flood quantile result at each site was 1.31 percent. As a result of the rare difficulty in identifying the correct location, the IDOT Structure Identification Management System shapefile was added to the ILSS implementation. This enables the user to select the coverage for display and use the ILSS Identify tool to determine whether the structure site has been correctly identified.

Some of the sites tested had completed manually computed quantile estimates available. Figure 12 shows the comparison of the 34 sites with manual analyses that were tested with ILSS. The accuracy of the manually computed quantile estimates cannot be evaluated because they were not repeated tests; however, a paired t-test between the manual computations and ILSS 100-year flood quantile estimates indicated no statistically significant difference at the 95-percent confidence level.

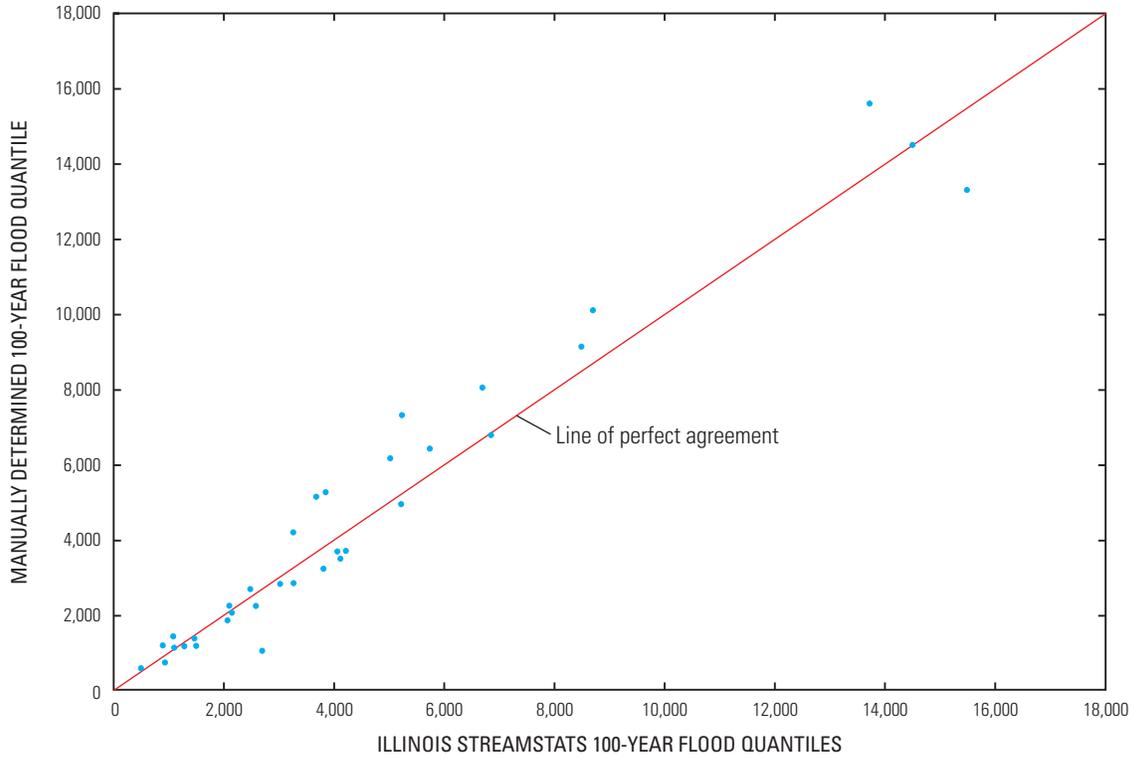


Figure 12. Relation between Illinois StreamStats (ILSS) and manually determined 100-year flood quantiles.

Sensitivity of Flood-Peak Quantiles

A sensitivity analysis was performed to demonstrate the deviation from the 2004 flood quantiles that may be expected for a range of differences in basin characteristics. To illustrate a range of potential differences, the proportional differences between the basin characteristics determined from ILSS and those from Soong and others (2004) for a sample of the rural streamflow-gaging stations used in the flood-frequency regression equations were determined. The sample size was 271 stations, except for basin length, which used 47 stations.

The proportional difference in basin characteristic is expressed as

$$\Delta(BC) = \frac{(BC_{2004} - BC_{ILSS})}{BC_{2004}} \tag{1}$$

where

- BC* is a basin characteristic parameter,
- Δ is the difference,
- ILSS* represents the ILSS-computed value, and
- 2004 is the value published in Soong and others (2004).

Figures 13A through 13E are histograms of the proportional differences in DA, SL, PermAvg, %Water, and BL. Interval of the histogram is set to 0.05, and three lines in each plot indicate the mean (center) and the range of ± 1 standard deviation, σ . The range of $\pm \sigma$ includes more than 67 percent of the differences, because the differences are not distributed

normally. The percent of values included in the $\pm \sigma$ are shown on each histogram.

The sensitivity of the estimated flood quantiles, Q_T , to the total drainage area is shown in table 10. The range of drainage areas tested was from 40 to 160 percent of the published (BC_{2004}) values, with a corresponding range in the flood quantiles change, expressed as the ratio of the tested Q_T to the published Q_T (Q_T^{TEST}/Q_T^{2004}) from 0.46 to 1.49 for Q_{100} and Q_{500} in region 4 to as little as 0.52 to 1.40 for Q_5 through Q_{500} in regions 2, 6, and 7. Drainage area is used in all regional regression equations.

The sensitivity of the estimated flood quantiles, Q_T , to the value of stream or main-channel slope is shown in table 11. The range of slopes tested was from 60 to 140 percent of the published (BC_{2004}) values, with a corresponding range in the flood quantile changes, expressed as the ratio Q_T^{TEST}/Q_T^{2004} from 0.76 to 1.19 for Q_{500} in region 1 to as little as 0.84 to 1.12 for Q_2 in regions 2, 6, and 7. The stream slope characteristic used in all regional regression equations was the main-channel slope, determined by dividing the elevation difference, in feet, by the distance, in miles, at points 10 and 85 percent from the outlet to the intersection of the main channel and the basin divide. The main-channel slope was determined using Basin-Soft, which is used as the published (Soong and others, 2004) slope value for the sensitivity analysis.

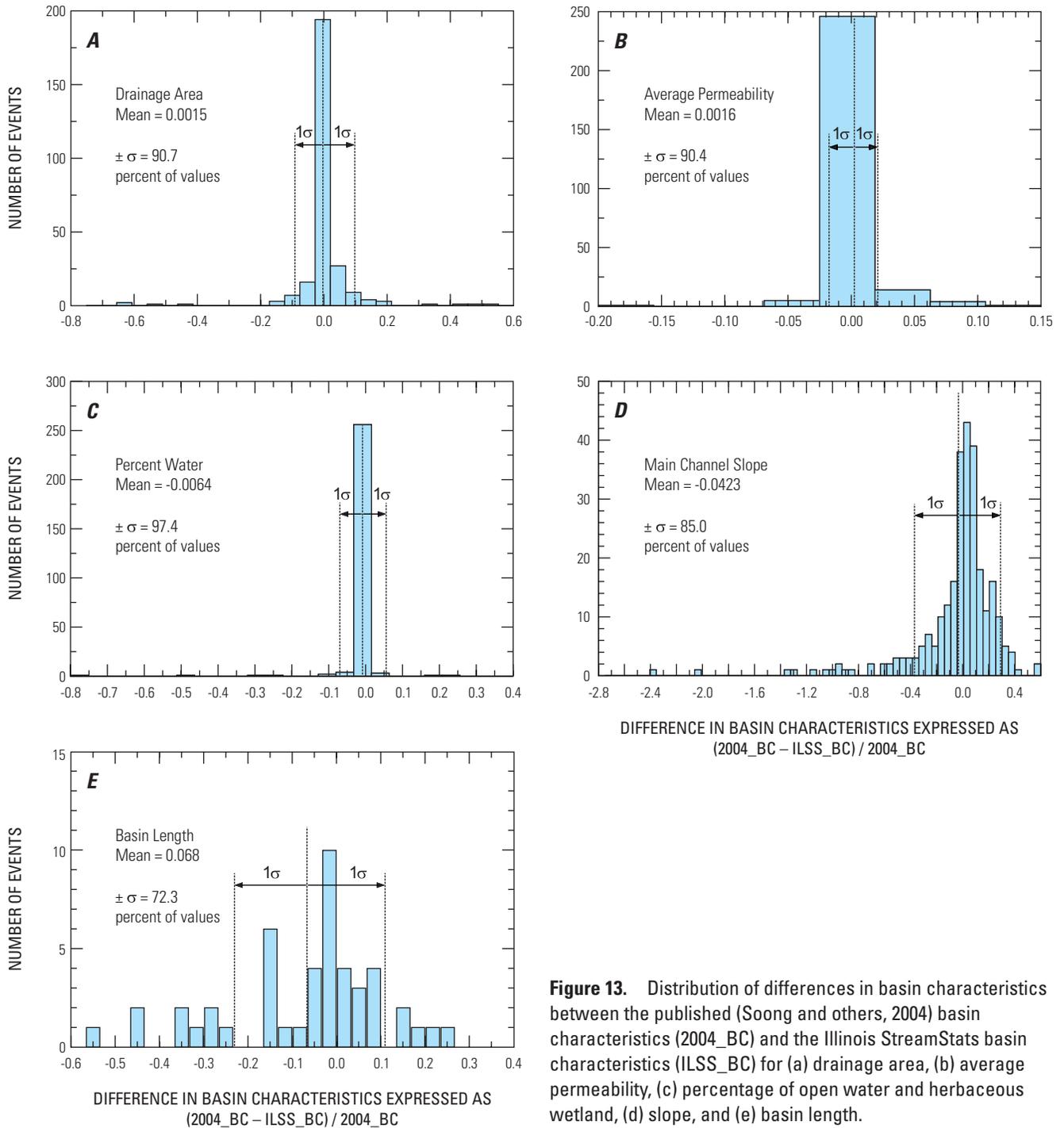


Figure 13. Distribution of differences in basin characteristics between the published (Soong and others, 2004) basin characteristics (2004_BC) and the Illinois StreamStats basin characteristics (ILSS_BC) for (a) drainage area, (b) average permeability, (c) percentage of open water and herbaceous wetland, (d) slope, and (e) basin length.

Table 10. Sensitivity of flood quantiles based on the published (Soong and others, 2004) regression equations for specified return-intervals, Q_T , to differences in drainage area for (a) regions 1, 3, and 5; (b) regions 2, 6, and 7; and (c) region 4, expressed as the ratio of the tested Q_T to the published Q_T , corresponding to percent of published drainage area.

[DA, drainage area; Q_T , flood quantile; Q_2 , 2-year flood quantile; Q_5 , 5-year flood quantile; Q_{10} , 10-year flood quantile; Q_{25} , 25-year flood quantile; Q_{50} , 50-year flood quantile; Q_{100} , 100-year flood quantile; Q_{500} , 500-year flood quantile; %, percent]

Q_T	Ratio of tested Q_T to published Q_T									
	40%	60%	80%	90%	95%	105%	110%	120%	140%	160%
(a) Regions 1, 3, and 5										
Q_2	0.50	0.68	0.85	0.92	0.96	1.04	1.07	1.15	1.29	1.42
Q_5	0.51	0.68	0.85	0.92	0.96	1.04	1.07	1.15	1.28	1.42
Q_{10}	0.51	0.69	0.85	0.92	0.96	1.04	1.07	1.14	1.28	1.42
Q_{25}	0.51	0.69	0.85	0.93	0.96	1.04	1.07	1.14	1.28	1.41
Q_{50}	0.51	0.69	0.85	0.93	0.96	1.04	1.07	1.14	1.28	1.41
Q_{100}	0.51	0.69	0.85	0.93	0.96	1.04	1.07	1.14	1.28	1.41
Q_{500}	0.51	0.69	0.85	0.93	0.96	1.04	1.07	1.14	1.28	1.41
(b) Regions 2, 6, and 7										
Q_2	0.51	0.69	0.85	0.93	0.96	1.04	1.07	1.14	1.28	1.41
Q_5	0.52	0.69	0.85	0.93	0.96	1.04	1.07	1.14	1.27	1.40
Q_{10}	0.52	0.69	0.85	0.93	0.96	1.04	1.07	1.14	1.27	1.40
Q_{25}	0.52	0.69	0.85	0.93	0.96	1.04	1.07	1.14	1.27	1.40
Q_{50}	0.52	0.69	0.85	0.93	0.96	1.04	1.07	1.14	1.27	1.40
Q_{100}	0.52	0.69	0.85	0.93	0.96	1.04	1.07	1.14	1.27	1.40
Q_{500}	0.52	0.69	0.85	0.93	0.96	1.04	1.07	1.14	1.27	1.40
(c) Region 4										
Q_2	0.51	0.69	0.85	0.93	0.96	1.04	1.07	1.14	1.28	1.41
Q_5	0.49	0.67	0.84	0.92	0.96	1.04	1.08	1.15	1.30	1.44
Q_{10}	0.48	0.67	0.84	0.92	0.96	1.04	1.08	1.16	1.31	1.45
Q_{25}	0.48	0.66	0.83	0.92	0.96	1.04	1.08	1.16	1.31	1.46
Q_{50}	0.47	0.66	0.83	0.92	0.96	1.04	1.08	1.16	1.32	1.47
Q_{100}	0.47	0.65	0.83	0.92	0.96	1.04	1.08	1.16	1.32	1.48
Q_{500}	0.46	0.65	0.83	0.91	0.96	1.04	1.08	1.17	1.33	1.49

The sensitivity of the estimated flood quantiles, Q_T , to the computation of average permeability is shown in table 12. The range of average permeability differences tested was from 60 to 140 percent of the published (BC_{2004}) values, with a corresponding range in the flood quantile changes, expressed as the ratio Q_T^{TEST}/Q_T^{2004} , of 1.12 to 0.93 for Q_2 through Q_{500} . The average permeability is used in the regional regression equations for only regions 1, 3, and 5. The sensitivity of the flood quantiles to average permeability is relatively low, with very large percent changes in the basin characteristic resulting in only small changes in the computed flood quantile. Differences in average permeability are expected to be the effect of differences in the watershed delineation only, as the permeability layer that was implemented in ILSS is identical to the one used in Soong and others (2004).

The sensitivity of the estimated flood quantiles, Q_T , to the computation of %Water area is shown in table 13. The range of %Water differences tested was from 60 to 140 percent of the average base values for all 100 stations, including individual stations that may have a value of zero percent, making a base-value increase or decrease impossible. The corresponding range in the flood quantile changes, expressed as the ratio Q_T^{TEST}/Q_T^{2004} , ranged from 1.37 to 0.75 to as little as 1.27 to 0.80 for Q_2 through Q_{500} . The %Water basin characteristic is used in the regional regression equations for only regions 2, 6, and 7. The sensitivity of the flood quantiles to %Water area is relatively low, with very large percent change in the basin characteristic, resulting in only small changes in the computed flood quantile. Differences in %Water are expected to be the effect of differences in the watershed delineation only,

Table 11. Sensitivity of flood quantiles based on the published (Soong and others, 2004) regression equations for specified return-intervals, Q_T , to differences in stream slope for (a) regions 1, 3, and 5; (b) regions 2, 6, and 7; and (c) region 4, expressed as the ratio of the tested Q_T to the published Q_T , corresponding to percent of published stream slope.

[Q_T , flood quantile; Q_2 , 2-year flood quantile; Q_5 , 5-year flood quantile; Q_{10} , 10-year flood quantile; Q_{25} , 25-year flood quantile; Q_{50} , 50-year flood quantile; Q_{100} , 100-year flood quantile; Q_{500} , 500-year flood quantile; %, percent]

Q_T	Ratio of tested Q_T to published Q_T							
	60%	80%	90%	95%	105%	110%	120%	140%
(a) Regions 1, 3, and 5								
Q_2	0.81	0.91	0.96	0.98	1.02	1.04	1.08	1.14
Q_5	0.80	0.91	0.96	0.98	1.02	1.04	1.08	1.16
Q_{10}	0.79	0.90	0.95	0.98	1.02	1.04	1.09	1.17
Q_{25}	0.78	0.90	0.95	0.98	1.02	1.05	1.09	1.17
Q_{50}	0.78	0.90	0.95	0.98	1.02	1.05	1.09	1.18
Q_{100}	0.77	0.89	0.95	0.97	1.02	1.05	1.10	1.18
Q_{500}	0.76	0.89	0.95	0.97	1.03	1.05	1.10	1.19
(b) Regions 2, 6, and 7								
Q_2	0.84	0.93	0.96	0.98	1.02	1.03	1.06	1.12
Q_5	0.83	0.92	0.96	0.98	1.02	1.04	1.07	1.13
Q_{10}	0.82	0.92	0.96	0.98	1.02	1.04	1.07	1.14
Q_{25}	0.81	0.91	0.96	0.98	1.02	1.04	1.08	1.15
Q_{50}	0.80	0.91	0.96	0.98	1.02	1.04	1.08	1.15
Q_{100}	0.80	0.91	0.96	0.98	1.02	1.04	1.08	1.16
Q_{500}	0.79	0.90	0.95	0.98	1.02	1.04	1.09	1.17
(c) Region 4								
Q_2	0.83	0.92	0.96	0.98	1.02	1.04	1.07	1.13
Q_5	0.81	0.91	0.96	0.98	1.02	1.04	1.08	1.15
Q_{10}	0.80	0.91	0.96	0.98	1.02	1.04	1.08	1.15
Q_{25}	0.80	0.91	0.95	0.98	1.02	1.04	1.08	1.16
Q_{50}	0.79	0.90	0.95	0.98	1.02	1.04	1.09	1.17
Q_{100}	0.79	0.90	0.95	0.98	1.02	1.05	1.09	1.17
Q_{500}	0.78	0.90	0.95	0.97	1.02	1.05	1.09	1.18

as the water bodies and herbaceous wetland layer that was implemented in ILSS is identical to the one used in Soong and others (2004).

The sensitivity of the estimated flood quantiles, Q_T , to the basin length is shown in table 14. The range of basin length differences tested was from 40 to 130 percent of the published (BC_{2004}) values, with a corresponding range in the flood quantile changes, expressed as the ratio Q_T^{TEST}/Q_T^{2004} , from 1.28 to 0.93 for Q_{500} to as little as 1.01 to 1.00 for Q_2 . Basin length is used in the regional regression equation only for region 4.

The sensitivity analysis indicates that for the common range of 60 to 120 percent of the published basin characteristics that was tested, the estimated flood quantiles were most sensitive to drainage-area differences, with the range in quantile changes, expressed as the ratios Q_T^{TEST}/Q_T^{2004} , varying

from 0.45 (1.14–0.69) to 0.52 (1.17–0.65). The flood quantiles were less sensitive to %Water, with the range of quantile changes expressed as the ratios Q_T^{TEST}/Q_T^{2004} , varying from 0.35 (1.27–0.92) to 0.48 (1.37–0.89). The range of quantile changes, expressed as the ratios Q_T^{TEST}/Q_T^{2004} , for stream slope was 0.22 (1.06–0.84) to 0.34 (1.10–0.76), and for basin length the range of quantile changes expressed as the ratios Q_T^{TEST}/Q_T^{2004} , was from 0 (1.0–1.0) to 0.20 (1.15–0.95). For average permeability the minimum and maximum range of quantile changes expressed as the ratios Q_T^{TEST}/Q_T^{2004} , was 0.16 (1.12–0.96). The relative sensitivity does not indicate the likelihood of computing any particular basin characteristic difference, but rather the influence of the basin characteristics in the equation.

Table 12. Sensitivity of flood quantiles based on the published (Soong and others, 2004) regression equations for specified return-intervals, Q_T to differences in average permeability for regions 1, 3, and 5, expressed as the ratio of the tested Q_T to the published Q_T , corresponding to percent of published average permeability.

[Q_T , flood quantile; Q_2 , 2-year flood quantile; Q_5 , 5-year flood quantile; Q_{10} , 10-year flood quantile; Q_{25} , 25-year flood quantile; Q_{50} , 50-year flood quantile; Q_{100} , 100-year flood quantile; Q_{500} , 500-year flood quantile; %, percent]

Q_T	Ratio of tested Q_T to published Q_T							
	60%	80%	90%	95%	105%	110%	120%	140%
Regions 1, 3, and 5								
Q_2	1.12	1.05	1.02	1.01	0.99	0.98	0.96	0.93
Q_5	1.12	1.05	1.02	1.01	0.99	0.98	0.96	0.93
Q_{10}	1.12	1.05	1.02	1.01	0.99	0.98	0.96	0.93
Q_{25}	1.12	1.05	1.02	1.01	0.99	0.98	0.96	0.93
Q_{50}	1.12	1.05	1.02	1.01	0.99	0.98	0.96	0.93
Q_{100}	1.12	1.05	1.02	1.01	0.99	0.98	0.96	0.93
Q_{500}	1.12	1.05	1.02	1.01	0.99	0.98	0.96	0.93

Table 13. Sensitivity of flood quantiles based on the published (Soong and others, 2004) regression equations for specified return-intervals, Q_T to differences in selected percentage water bodies and herbaceous wetland area (%Water) for regions 2, 6, and 7, expressed as the ratio of the tested Q_T to the published Q_T , corresponding to percent of published %Water.

[Q_T , flood quantile; Q_2 , 2-year flood quantile; Q_5 , 5-year flood quantile; Q_{10} , 10-year flood quantile; Q_{25} , 25-year flood quantile; Q_{50} , 50-year flood quantile; Q_{100} , 100-year flood quantile; Q_{500} , 500-year flood quantile; %, percent]

Q_T	Ratio of tested Q_T to published Q_T								
	60%	80%	90%	95%	105%	110%	120%	140%	160%
Regions 2, 6, and 7									
Q_2	1.27	1.11	1.05	1.02	0.98	0.96	0.92	0.85	0.80
Q_5	1.31	1.12	1.06	1.03	0.97	0.95	0.91	0.84	0.78
Q_{10}	1.32	1.13	1.06	1.03	0.97	0.95	0.90	0.83	0.77
Q_{25}	1.34	1.14	1.06	1.03	0.97	0.95	0.90	0.82	0.76
Q_{50}	1.35	1.14	1.06	1.03	0.97	0.95	0.90	0.82	0.76
Q_{100}	1.36	1.14	1.07	1.03	0.97	0.94	0.90	0.82	0.75
Q_{500}	1.37	1.15	1.07	1.03	0.97	0.94	0.89	0.81	0.75

Table 14. Sensitivity of flood quantiles based on the published (Soong and others, 2004) regression equations for specified return-intervals, Q_T , to differences in basin length for region 4, expressed as the ratio of the tested Q_T to the published Q_T , corresponding to percent of published basin length.

[Q_T , flood quantile; Q_2 , 2-year flood quantile; Q_5 , 5-year flood quantile; Q_{10} , 10-year flood quantile; Q_{25} , 25-year flood quantile; Q_{50} , 50-year flood quantile; Q_{100} , 100-year flood quantile; Q_{500} , 500-year flood quantile; %, percent]

Q_T	Ratio of tested Q_T to published Q_T								
	40%	60%	80%	90%	95%	105%	110%	120%	130%
Region 4									
Q_2	1.01	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
Q_5	1.09	1.05	1.02	1.01	1.00	1.00	0.99	0.98	0.98
Q_{10}	1.14	1.07	1.03	1.01	1.01	0.99	0.99	0.97	0.96
Q_{25}	1.18	1.10	1.04	1.02	1.01	0.99	0.98	0.97	0.95
Q_{50}	1.21	1.11	1.05	1.02	1.01	0.99	0.98	0.96	0.95
Q_{100}	1.23	1.12	1.05	1.02	1.01	0.99	0.98	0.96	0.94
Q_{500}	1.28	1.15	1.06	1.03	1.01	0.99	0.97	0.95	0.93

Conclusions and Limitations

The ILSS implementation of the regional regression equations for estimated flood quantiles at ungaged, unregulated rural sites for Illinois was found to be an adequate method for applying the most current (2010) published equations (Soong and others, 2004). The basin characteristics computed by ILSS were compared to the basin characteristics published in Soong and others (2004) and used in the rural flood-frequency equations at the applicable regional subsets of 283 rural streamflow-gaging station locations. There were no significant differences in drainage area and percentage of open water and herbaceous wetland, although relative differences were larger for smaller drainage areas. StreamStats enables the user to manually edit the drainage basin if errors in the drainage-area delineation are found. Average permeability had a small significant difference (less than 0.25 percent). Significant differences were found in BL and slope; BL was adjusted using a linear best-fit regression line. The adjusted BL did not differ significantly from the published values at the 95-percent confidence level. The slope was adjusted by a linear best-fit regression line on log-transformed slope values. This removed the statistical significance of the differences determined by the t-test, but the adjusted values were significantly different from the published values under the Wilcoxon signed-rank test, with a mean difference of -4.76 percent. The adjusted distribution was selected as preferable.

A sensitivity analysis was done to determine the sensitivity of the estimated flood-peak quantiles to the basin-characteristic differences. For the common range of 60 to 120 percent of published (Soong and others, 2004) basin

characteristics that were tested, the greatest average range of sensitivity of the resulting flood-peak quantiles was (in order from greatest to least) drainage area, %Water, slope, average permeability, and basin length. The relative sensitivity does not indicate the likelihood of computing any particular basin characteristic difference, but rather the influence of the basin characteristics in the equation.

The flood-peak quantiles produced by ILSS were compared to the published values at an approximately random sample of 170 streamflow-gaging stations. There were no significant difference at the 95-percent confidence level between the log-transformed flood-peak quantile estimates published in Soong and others (2004) and those computed by ILSS, either taken as a whole or sorted by the hydrographic region identified in Soong and others (2004), except for region 1. Region 1 had a small statistically significant difference ranging from 3.8 percent for the 2-year flood-quantile estimate to 4.3 percent for the 500-year flood-quantile estimate at the 95-percent confidence level. All 21 stations were considered in the analysis, because of the few stations available in region 1. The total number of stations in region 1 was small (21) and the mean difference is less than one-tenth of the average prediction errors for the 2- to 500-year regression-equation estimates, which range from 39.5 to 54.9 percent respectively.

A test of usage reliability was conducted by having at least 7 new users compute ILSS estimates at 27 locations. The average maximum deviation from the mode value of the 100-year flood quantile estimate at each site was 1.31 percent after four mislocated sites were removed. A comparison of manual 100-year flood-quantile computations with ILSS estimates at 34 sites indicated no statistically significant difference. The

estimates of flood quantiles computed by ILSS are based on the assumption that streamflow at the site is not appreciably regulated. ILSS does not return a warning of the presence of regulation or urbanization in a delineated drainage basin; consequently, the user should consider this possibility when using the tool. All other limitations described in Soong and others (2004) apply to ILSS.

The basin characteristics computed by ILSS should be compared to the range of basin characteristics used in developing the flood-frequency regional regression equations. This comparison is facilitated by the information provided in the ILSS output, showing the range. Extrapolations outside the range of values should be avoided. Although, in general, the user must determine whether a desired site is outside the range of applicability of the regional regression equations (for reasons such as regulation or urbanization), the application is not implemented for the Illinois, Mississippi, Ohio, and Wabash Rivers. The basin characteristics are based on the geospatial data sets and the computer algorithms as described in this report and are subject to the differences that are described in the sections on the evaluation and adjustment of basin characteristics. ILSS appears to be an accurate, reliable, and effective tool for flood-quantile estimates and the determination of a consistent set of basin characteristics.

Summary

The rural flood-frequency equations for Illinois were implemented in a U.S. Geological Survey (USGS) StreamStats Web-based application, the Illinois StreamStats (ILSS), in cooperation with the Illinois Center for Transportation, the Illinois Department of Transportation, and the Illinois Department of Natural Resources—Office of Water Resources. ILSS computes selected basin characteristics and flood-peak quantiles based on the most recently (2010) published (Soong and others, 2004) regional flood-frequency equations at any rural stream location in Illinois. Limited streamflow statistics including general statistics, flow durations, and base flows also are available for USGS streamflow-gaging stations. ILSS can be accessed on the Web at <http://streamstats.usgs.gov/> by selecting the State Applications hyperlink and choosing Illinois from the pull-down menu.

The basin characteristics produced by ILSS were compared to the basin characteristics published in Soong and others (2004) and used in the rural flood-frequency equations at 283 rural streamflow-gaging station locations by testing for significant differences at the 95-percent confidence level using the paired t-test. There were no significant differences in drainage area and percentage of open water and herbaceous wetland, although relative differences were larger for smaller drainage areas, where local hydraulic-control features are relatively important. StreamStats enables the user to manually edit the drainage basin. Average permeability had a small significant difference (less than 0.25 percent), and no correction was

considered necessary because of the very low sensitivity of the flood quantile to this characteristic. Significant differences were found in basin length and slope. Basin length was adjusted using a linear best-fit regression line. The adjusted basin length did not differ significantly from the published values. The slope was adjusted by a linear best-fit regression line on log-transformed slope values. The adjusted values were not significantly different from the published values according to the paired t-test, but did have a significant difference according to the Wilcoxon signed-rank test, with a mean difference of 4.22 percent.

A sensitivity analysis was used to determine the sensitivity of a large sample (271) of the estimated flood-peak quantiles to the basin-characteristic differences. For the common range of 60 to 120 percent of published (Soong and others, 2004) basin characteristics that were tested, the greatest average range of sensitivity of the resulting flood-peak quantiles was (in order from greatest to least) drainage area, %Water, slope, average permeability, and basin length. The relative range in sensitivity does not indicate the likelihood of computing any particular basin characteristic difference, but rather the influence of the basin characteristics in the regional regression equations.

The flood-peak quantiles produced by ILSS were compared to the published values at an approximately random sample of 170 streamflow-gaging stations. There were no significant differences between the log-transformed flood-peak quantile estimates published in Soong and others (2004) and those computed by ILSS, either taken as a whole or sorted by the hydrographic regions identified in Soong and others (2004), except for region 1. Region 1 had a small statistically significant difference ranging from 3.76 percent for the 2-year flood-quantile estimate to 4.27 percent for the 500-year flood-quantile estimate at the 95-percent confidence level. All 21 stations were considered in the analysis, because of the few stations available in region 1. The total number of stations in region 1 was small, and the mean difference is less than one-tenth of the average prediction errors for the 2- to 500-year regression-equation estimates, which range from 39.5 to 54.9 percent, respectively.

A test of the ILSS usage reliability was conducted by having at least 7 new users compute flood-quantile estimates at 27 locations. The average maximum deviation of the 100-year flood quantile estimate from the mode value at each site was 1.31 percent after four mislocated sites were removed. A comparison of manual 100-year flood-quantile computations with ILSS at 34 sites indicated no statistically significant difference.

ILSS appears to be an accurate, reliable, and effective tool for flood-quantile estimates and the determination of a consistent set of basin characteristics.

Acknowledgments

The authors thank the many Illinois Department of Transportation (IDOT) and Illinois Department of Natural Resources—Office of Water Resources (IDNR—OWR) hydraulic engineers who helped with the testing and evaluation of the functionality of Illinois StreamStats. The authors thank Matthew R. O'Connor, IDOT, the chair of the Technical Review Panel (TRP) and the other TRP members including Arlan Juhl, IDNR-OWR, Neil Vanbebber, IDOT, and Mark Gadwedzinski, IDOT, for providing guidance and oversight for the project.

The authors thank Ann Holtrop, formerly of the Watershed Protection Section, Illinois Natural History Survey, Illinois Department of Natural Resources, who processed the majority of the dendritic network by June 2006 and graciously provided the data for use in this study.

The authors thank the U.S. Geological Survey (USGS) StreamStats team, especially Alan Rea, Kernell Ries, and David Stewart who assisted with technical implementation and testing issues. The authors thank Christine Dartiguenave and Sreeresh Sreedhar of ESRI, Inc., who completed the XML programming for the basin-length characteristics. Finally, the authors thank Thomas Over and Timothy Straub of the USGS Illinois Water Science Center, and David Eash of the USGS Iowa Water Science Center who offered suggestions on the statistical analyses and watershed-geomorphology issues.

References Cited

- Berk, K.N., and Carey, P., 2004, Data analysis with Microsoft Excel: Brooks/Cole—Thomson Learning, 579 p.
- ESRI, Inc., 2005, Arc Hydro tools overview—Version 1.1 final, July 2005, accessed April 8, 2009, at <http://www.crwr.utexas.edu/gis/gishydro06/ArcHydro/ArcHydroTools/Doc/Arc%20Hydro%20Tools%20-%20Overview.pdf>.
- Harvey, C.A., and Eash, D.A., 1996, Description, instructions, and verification for BasinSoft, a computer program to quantify drainage-basin characteristics: U.S. Geological Survey Water-Resources Investigations Report 95-4287, 25 p.
- Helsel, D.R., and Hirsch, R.M., 2002, Statistical methods in water resources: U.S. Geological Survey Techniques of Water-Resources Investigations, book 4 (Hydrologic Analysis and Interpretation), chap. A3, 510 p.
- Jennings, M.E., Thomas, W.O., Jr., and Riggs, H.C., 1994, Nationwide summary of U.S. Geological Survey regional regression equations for estimating magnitude and frequency of floods for ungaged sites, 1993: U.S. Geological Survey Water-Resources Investigations Report 94-4002, 194 p.
- Natural Resources Conservation Service, 1993, State soil geographic base (STATSGO): U.S. Department of Agriculture, Miscellaneous Publication No. 1492, 110 p., and computer data.
- Ries III, K.G., 2007, The national streamflow statistics program: A computer program for estimating streamflow statistics for ungaged sites: U.S. Geological Survey Techniques and Methods 4-A6, 37 p.
- Ries, K.G., III, Guthrie, J.G., Rea, A.H., Steeves, P.A., Coles, J.D., and Stewart, D.W., 2008, StreamStats—A water resources Web application: U.S. Geological Survey Fact Sheet FS 2008-3067, 6 p.
- Schlotzhauer, S.D., and Littell, R.C., 1987, SAS system for elementary statistical analysis: Cary, N.C., SAS Institute, 410 p.
- Soong, D.T., Ishii, A.L., Sharpe, J.B., and Avery, C.F., 2004, Estimating flood-peak discharge magnitudes and frequencies for rural streams in Illinois: U.S. Geological Survey Scientific Investigations Report 2004-5103, 147 p.
- Wolock, D.M., 2003a, Flow characteristics at U.S. Geological Survey streamgages in the conterminous United States: U.S. Geological Survey Open-File Report 03-146, digital data set, accessed May 11, 2009, at <http://water.usgs.gov/lookup/getspatial?qsitesdd>.
- Wolock, D.M., 2003b, Base-flow index grid for the conterminous United States: U.S. Geological Survey Open-File Report 03-263, digital data set, accessed May 11, 2009, at <http://water.usgs.gov/lookup/getspatial?bfi48grd>.

