

# STREAMSTATS: A U.S. Geological Survey web site for stream information

**Kernell G. Ries III\***

\*U.S. Geological Survey, Office of Surface Water, 12201 Sunrise Valley Drive, MS415, Reston, VA 20192, phone 703-648-5307, [kries@usgs.gov](mailto:kries@usgs.gov)

**Abstract** Streamflow statistics, such as the 100-year flood and the 7-day, 10-year low flow, are needed for water-resources planning and management; for design of bridges, culverts, and flood-control structures; and for many other purposes. This paper describes the National Streamstats (STREAMSTATS) computer application currently in development by the U.S. Geological Survey for providing streamflow statistics to the public. STREAMSTATS works through a Web browser to provide various types of streamflow statistics for data-collection stations and for ungauged sites. STREAMSTATS users can point and click on data-collection stations shown on a map in their Web browser window to obtain streamflow statistics and other information for the stations. Users also can point and click on any stream shown on the map to get estimates of streamflow statistics for ungauged sites. STREAMSTATS determines the watershed boundaries and measures physical and climatic characteristics of the watersheds for the ungauged sites by use of a Geographic Information System, and then it inserts the characteristics into previously determined regression equations to estimate the streamflow statistics. Compared to manual methods, STREAMSTATS reduces the average time needed to estimate streamflow statistics for ungauged sites from several hours to several minutes.

**Keywords** Surface water, flood frequency, low flow, statistics, United States.

## Introduction

Government agencies need streamflow statistics, such as the 100-year peak flow, the 7-day, 10-year low flow, and the mean flow, to plan and manage use of water resources and to regulate activities in and around streams. Engineers, environmental consultants, utilities, industries, and others use the statistics to design and operate water supplies, hydropower facilities, industrial facilities, wastewater-treatment facilities, roads, bridges, and other structures.

The U.S. Geological Survey (USGS) periodically publishes streamflow statistics for data-collection stations on a statewide or regional basis. Published streamflow statistics have often been difficult to obtain, however, because they are scattered among many reports and because many old reports are not readily available to the public. In addition, streamflow statistics are often needed for locations where no data are available.

To provide simple methods for estimating streamflow statistics for ungauged sites, the USGS has developed and published numerous regression equations. Estimates of streamflow statistics can be obtained by inserting measured watershed and climatic characteristics for sites of interest into the regression equations. Equations for estimating flood-frequency statistics are available for every State, the Commonwealth of Puerto Rico, American Samoa,

and several metropolitan areas in the United States (for example, see [http://webdiaiwc.cr.usgs.gov/pubs/reports/WRIR\\_00-4233.pdf](http://webdiaiwc.cr.usgs.gov/pubs/reports/WRIR_00-4233.pdf)). Equations for estimating other types of streamflow statistics are available in many areas, as well (for example, see <http://water.usgs.gov/pubs/wri/wri004135/pdf/report.pdf>).

This paper describes the National Streamstats (STREAMSTATS) program currently under development by the USGS. STREAMSTATS works through a Web browser to provide various types of streamflow statistics for data-collection stations (stream-gauging stations, low-flow partial-record stations, peak-flow partial-record stations, and miscellaneous-measurement stations) and for ungauged sites. STREAMSTATS users can point and click on data-collection stations shown on a map in their Web browser window to obtain streamflow statistics and other information for the stations. Users also can point and click on any stream shown on the map to get estimates of streamflow statistics for ungauged sites. STREAMSTATS automatically determines the watershed boundary and climatic characteristics for the ungauged site by use of a Geographic Information System (GIS), and then it inserts the characteristics into previously determined regression equations to estimate the streamflow statistics. Compared to manual methods, STREAMSTATS reduces the average time needed to estimate the streamflow statistics for ungauged sites from several hours to several minutes. Estimates provided by STREAMSTATS for ungauged sites assume rural (non-regulated) flow conditions only.

## **STREAMSTATS development**

An initial version of STREAMSTATS was released in 2000 for Massachusetts, in the northeastern United States. The Massachusetts version of STREAMSTATS can be accessed on the Web at <http://ma.water.usgs.gov/streamstats>. The Web site provides full documentation (Ries and others, 2000) and instructions for use of the application. The USGS and MassGIS, the Massachusetts Geographic Information Systems (GIS) Agency, jointly developed Massachusetts STREAMSTATS in cooperation with the Massachusetts Departments of Environmental Management and Environmental Protection. The cooperating agencies wanted to use STREAMSTATS primarily as a planning and management tool. The State of Massachusetts requires applicants for water-withdrawal permits to use STREAMSTATS to determine the natural flow of any stream from which 378.5 cubic meters (100,000 gallons) per day or more of water would be withdrawn.

Since its release, more than 20 States have expressed interest in implementing STREAMSTATS; however, Massachusetts STREAMSTATS is not easily modified for use in other areas because it was built to use GIS data sets that are unique to Massachusetts. To meet this need, the USGS is developing a new version of STREAMSTATS that will be able to provide streamflow statistics and other information for any State. It is designed to have all of the features of Massachusetts STREAMSTATS, but it will be easier to use and will have several new features. The new program will take advantage of advances in the technology for serving maps and data over the Web, and in other advances in computer technology, since Massachusetts STREAMSTATS was developed.

As in Massachusetts STREAMSTATS, the new program has four major components. These include (1) a user interface that displays maps and allows users to select stream locations where they want streamflow statistics; (2) a database of previously published streamflow statistics and descriptive information for 725 USGS data-collection stations; (3) a GIS database that contains all the map layers needed for measuring the watershed and climatic characteristics in addition to numerous map layers that can be used to identify locations of interest; and (4) an automated GIS procedure that measures the watershed and

climatic characteristics needed to solve the regression equations and inserts those characteristics into equations to estimate the streamflow statistics. Each of these components is described below.

### **User interface.**

Orkand<sup>1</sup>, Inc., of Vienna, Virginia, is developing the STREAMSTATS user interface and the automated GIS procedure for the USGS using ArcIMS (<http://www.esri.com/software/arcims/index.html>) and ArcGIS 8.1 (<http://www.esri.com/software/arcgis/index.html>). The new user interface will appear within a Web browser window rather than as a separate window, as in Massachusetts STREAMSTATS.

The appearance of STREAMSTATS is designed to be similar that of most other USGS Web mapping sites (for example, see [http://gisdata.usgs.net/website/National\\_Map/viewer.htm](http://gisdata.usgs.net/website/National_Map/viewer.htm)) that are deployed using ArcIMS. The main features in the user interface will be a map frame and a map legend frame.

The map frame for displaying default and user-selected map layers will dominate the user interface. Initially, the map frame will show a map of the United States and State boundaries. Clicking on a state will cause a map of the state to appear in the map frame. The state map will show county boundaries, major streams, and data-collection stations. Zooming in further will cause other default data layers to display, depending on the scale of the map.

The map legend frame and a set of map navigation and information query tools will fill most of the rest of the interface. The map legend frame will show the names and the map symbols for the map layers that can be displayed in the map frame. Check boxes next to the layer names in the legend frame will allow the user to turn on and off the display of the layers. A pull-down menu will allow the user to select an active map layer. Activating a layer will allow use of the tools to operate specifically on the activated layer.

STREAMSTATS will include the standard ArcIMS tools for navigation, selection, and other actions, in addition to some custom tools. The standard tools will allow zooming in, zooming out, and panning to selected areas of the map; to select areas, lines, or polygons of interest on the map and obtain information stored as attributes for the selected map features; and to print whatever is shown in the map frame. Custom navigation tools will allow zooming to a specific place on the map by specifying a street address, a geographic name from the Geographic Names Information System of the USGS (GNIS, <http://geonames.usgs.gov/>), the latitude and longitude of the site, or a stream-reach address from the National Hydrography Dataset (NHD, <http://nhd.usgs.gov/>). Other custom tools will allow querying the database of information for data-collection stations, and obtaining estimates of streamflow statistics for ungauged sites.

### **Streamflow statistics database.**

Microsoft Access was used to develop the streamflow statistics database for national STREAMSTATS, named StreamstatsDB. Aqua Terra Consultants, of Decatur, Georgia, USA, under contract to the USGS, has built a custom user interface that allows for easy entry and update of the database. The initial release of StreamstatsDB contains 144 types of physical and climatic basin characteristics and 492 types of streamflow statistics. Each

---

<sup>1</sup> The use of trade, firm, or product names in this paper is for identification purposes only and does not imply endorsement by the U.S. Geological Survey.

streamflow statistic has an associated standard error of estimate in the database. All USGS-developed regression equations for estimating peak-flow frequency statistics were reviewed to determine the basin characteristics needed initially in StreamstatsDB.

StreamstatsDB has been released internally to each of the 48 District offices within the Water Discipline of the USGS so local staff can populate the database. District staff can easily add any new basin characteristics or streamflow statistics that they need in the database. There is a District office in each State and in the Commonwealth of Puerto Rico, except there are single District offices for New Hampshire and Vermont, Massachusetts and Rhode Island, and Maryland and Delaware. When the Districts have completed populating StreamstatsDB, the information for their States can be made available over the Web through national STREAMSTATS.

### GIS database

STREAMSTATS will use nationally available GIS data layers as much as possible for measuring basin characteristics and for use as base map layers. In many areas of the United States, however, local data layers will be used instead of or in addition to the national data layers because they have better resolution than the national data layers or they provide information that is not available from a national data layer. Some of the national data layers that will be used in STREAMSTATS for measuring basin characteristics include:

- the National Elevation Dataset (NED) (<http://edcnts12.cr.usgs.gov/ned/>),
- the Elevation Derivatives for National Applications (EDNA) dataset (<http://edcnts12.cr.usgs.gov/ned-h/>),
- the National Hydrography Dataset (NHD) (<http://nhd.usgs.gov/>),
- the Watershed Boundary Dataset (WBD) ([http://www.ftw.nrcs.usda.gov/huc\\_data.html](http://www.ftw.nrcs.usda.gov/huc_data.html)),
- the National Land Cover Dataset (NLCD) (<http://landcover.usgs.gov/>),
- Parameter-elevation Regressions on Independent Slopes Model (PRISM) climate data (<http://www.ocs.orst.edu/prism/>), and
- State Soil Geographic Database (STATSGO) soil survey data ([http://www.ftw.nrcs.usda.gov/stat\\_data.html](http://www.ftw.nrcs.usda.gov/stat_data.html))

STREAMSTATS will use many of the boundary, water, and transportation layers available for viewing and downloading from the USGS National Atlas Web site, <http://nationalatlas.gov/natlas/natlasstart.asp>, and the National Hydrography Dataset (<http://nhd.usgs.gov/>), as base map layers. In addition, STREAMSTATS will use the Digital Raster Graphics (DRG) (<http://mcmweb.er.usgs.gov/drg/>) scanned USGS topographic maps as the primary base layer for precise site selection.

STREAMSTATS requires three data layers to determine watershed boundaries for ungauged sites: (1) a Digital Elevation Model (DEM) (Elassel and Caruso, 1983), (2) a networked data layer of streams with lines connecting through the centers of wetlands and water bodies, and (3) a watershed boundary data layer. In much of the United States, STREAMSTATS will use the EDNA data layers to determine watershed boundaries and to measure several watershed characteristics. In some areas, however, STREAMSTATS will use other data layers, such as the NHD, the WBD, the NED, or locally developed data to determine watershed boundaries and characteristics.

The EDNA data layers were developed through an interagency effort, and are derived from the NED, a DEM with grid cells of 30 meters on a side and a vertical accuracy of about 1.5 meters. The EDNA includes a hydrologically correct DEM, a networked synthetic streams layer, a catchment layer that includes catchments derived for each reach in the

synthetic stream network, and several other layers. The DEM was made hydrologically correct by filling spurious depressions so that surface runoff could not collect in the depressions on the DEM land surface. The synthetic streams were obtained by setting a minimum flow-accumulation value of 5,000 DEM grid cells (about 4.5 square kilometers) to generate a stream. The catchments define the drainage boundary that is unique to each synthetic reach.

All local and national GIS data used in STREAMSTATS will be formatted to conform to the Hydrology Data Model developed by the GIS Water Resources Consortium (<http://www.crrwr.utexas.edu/giswr/>). The data will be stored in geodatabases (<http://www.esri.com/news/arcuser/0701/migrating.html>) organized by major watersheds. Watershed characteristics needed to solve the regression equations in STREAMSTATS will be pre-calculated for each catchment and attached as attributes to the catchment data layer to increase the efficiency of the application.

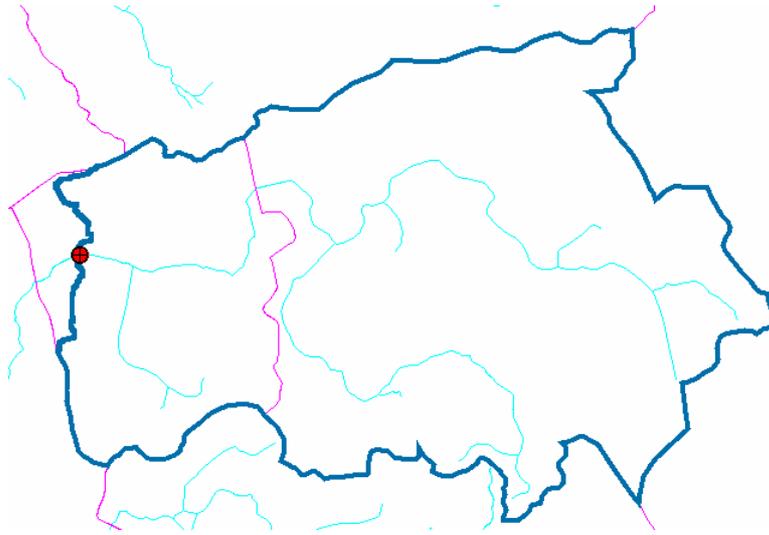
### **Automated GIS procedure.**

The STREAMSTATS automated procedure is being developed as a Dynamic Link Library (DLL), a bundle of Visual Basic subroutines that will operate within ArcGIS 8.1. The DLL will transfer information to and from ArcIMS to allow implementation over the Web.

When the user selects an ungauged site along the stream network, the coordinates for the site will be sent to a computer on which the STREAMSTATS automated GIS procedure is running. The procedure will then determine the watershed boundary for the location from the DEM up to the points at which the DEM-defined boundary coincides with existing catchment boundaries (fig. 1). STREAMSTATS then combines the new boundaries with the existing boundaries for any upstream catchments and dissolves the internal boundaries to obtain the complete drainage-basin boundary for the location. Use of the existing catchment boundaries minimizes errors associated with determining boundaries for ungauged sites from the raw DEM and minimizes processing time.

STREAMSTATS will display the watershed boundary for the user-selected site in the map frame, with the stream network and digital images of USGS topographic maps shown as the base data layers. This will allow users to verify that the boundary determined by STREAMSTATS is correct before proceeding with estimating streamflow statistics for the site.

After the drainage-basin boundary has been verified, STREAMSTATS determines the applicable watershed and climatic characteristics and solves the regression equations to obtain estimates of streamflow statistics for the ungauged site. These estimates will be printed in a pop-up Web browser window, along with the date, the coordinates of the site, the watershed and climatic characteristics, and the standard errors or prediction intervals. STREAMSTATS provides either standard errors or 90-percent prediction intervals as indicators of the reliability of the estimates. Approximately two-thirds of estimates for ungauged sites will have errors within the standard errors provided. There is a 90-percent probability that the actual streamflow statistics for a site are within the 90-percent prediction interval.



**Figure 1.** Example of a watershed boundary determined for an ungauged site, with the ungauged site shown as a red dot, the streams shown with a thin light blue line, the existing watershed boundaries shown with a pink line, and the watershed boundary for the ungauged site with a thick dark blue line.

### **STREAMSTATS implementation**

The new version of STREAMSTATS will work initially only for the States of Vermont, New Hampshire, Massachusetts, Kansas, and Idaho. These States were chosen as prototype areas because appropriate GIS data were available and the USGS District offices were able to help with the implementation.

Steps involved in implementing STREAMSTATS include (1) preparation of GIS data needed for the application, (2) population of StreamstatsDB with streamflow statistics and other information for data-collection stations, and (3) verification that the watershed and climatic characteristics are measured and the streamflow statistics are estimated without bias and with the same accuracy as that indicated in the USGS reports that describe the regression equations for each State. Support from the USGS District offices will be needed to complete these steps. The USGS will not implement STREAMSTATS for any State until the local District office has confirmed the reliability of the estimates. Several Districts already have started the implementation process; however, full national implementation of STREAMSTATS will likely take several years.

### **References**

- Elassel, A.A., and Caruso, V.M. (1983) USGS digital cartographic standards: digital elevation models. U.S. Geological Survey Circular 895-B.
- Ries, K.G., III, Steeves, P.A., Freeman, Aleda, and Singh, Raj. (2000) Obtaining streamflow statistics for Massachusetts streams on the World Wide Web. U.S. Geological Survey Fact Sheet 104-00.