StreamStats: A U.S. Geological Survey Web Site for Stream Information

Kernell G. Ries III, John R. Gray

Abstract

The U.S. Geological Survey has developed a Web application, named StreamStats, for providing streamflow statistics, such as the 100-year flood and the 7-day, 10-year low flow, to the public. Statistics can be obtained for data-collection stations and for ungaged sites. Streamflow statistics are needed for water-resources planning and management; for design of bridges, culverts, and flood-control structures; and for many other purposes. StreamStats users can point and click on data-collection stations shown on a map in their Web browser window to obtain previously determined 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 ungaged sites. StreamStats determines the watershed boundaries and measures physical and climatic characteristics of the watersheds for the ungaged sites 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 streamflow statistics for ungaged sites from several hours to several minutes.

Keywords: surface water, flood frequency, low flow, statistics, United States

Introduction

Estimates of the magnitude and frequency of floods (flood-frequency estimates), such as the 100-year flood (1 percent recurrence probability), are needed

Ries is a Hydrologist for the U.S. Geological Survey, Maryland-Delaware-District of Columbia District, Baltimore, MD 21237. Gray is a Hydrologist, U.S. Geological Survey, Office of Surface Water, Reston, VA 20192.

for floodplain mapping and for the design of bridges, culverts, and flood-control structures. Estimates of other streamflow statistics, such as the mean annual flow, the 7-day, 10-year low flow, and flow-duration statistics, are also needed for design of infrastructure and for water-resources planning and management. These estimates are commonly needed at gaged stream sites, where data are available to determine the statistics, and more often at ungaged sites, where no observed data are available.

The U.S. Geological Survey (USGS) has developed and published numerous equations to provide simple methods for estimating flood frequencies and other streamflow statistics at ungaged sites. The equations are developed from regression analyses on data for gaged sites, using streamflow statistics as the dependent variables and measured physical and (or) climatic characteristics of the drainage basins for the sites as explanatory variables. Flood-frequency equations are available for every State, Puerto Rico, American Samoa, and a number of metropolitan areas in the United States (for example, see http://me.water.usgs.gov/99-4008.pdf). Equations for estimating other streamflow statistics are also available in many areas (for example, see http://water.usgs.gov/pubs/wri/wri004135/).

The USGS has developed a prototype Web application named StreamStats that provides estimates of streamflow statistics for gaged and ungaged 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 ungaged sites. StreamStats uses a GIS to automatically determine the watershed boundary and physical and climatic characteristics for the ungaged site, and then it inserts the characteristics into previously determined regression equations to estimate the streamflow statistics for the site. Compared to

manual methods, StreamStats reduces the average time needed to estimate the streamflow statistics for ungaged sites from several hours to several minutes. Estimates provided by StreamStats for ungaged sites assume rural (non-regulated) flow conditions only.

StreamStats Development

An initial StreamStats prototype that was released in 2000 for Massachusetts can be accessed on the Web at http://ststdmamrl.er.usgs.gov/streamstats/. The Massachusetts StreamStats Web site provides full documentation (Ries et al. 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.

Representatives from many 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 and because the technology on which it was based is now obsolete. Because of this, the USGS has developed a new StreamStats prototype that will be able to provide streamflow statistics and other information for any State. The national StreamStats prototype was designed to have all of the features of Massachusetts StreamStats, but it is easier to use and has several new features. The new program takes 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.

The new prototype currently has been implemented only for Idaho (Figure 1). It is still being tested, and it is not yet available to the public. Plans have been made to make it available to the public and to add several more states within a year.

StreamStats has four major components, including (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

The StreamStats user interface and the automated GIS procedure were developed 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 appears within a Web browser window rather than as a separate window, as in Massachusetts StreamStats.

The map frame for displaying default and user-selected map layers is the largest feature in the user interface. For the Idaho prototype, the map frame initially shows a shaded-relief map of Idaho and the surrounding area, with locations of data-collection stations. When users zoom in to locate sites of interest, different scales of digital topographic maps are shown depending on the selected scale. When more states are added to StreamStats, a map of the United States and State boundaries will be shown initially. Clicking on a state will cause a map of the state to appear in the map frame.

Above the map frame is a set of buttons to activate tools that allow users to (1) zoom in, zoom out, and pan to areas of the map; (2) get information about features displayed on the map; (3) zoom to the full or last extent shown in the map; (4) delineate a drainage basin for an ungaged site; (5) get streamflow statistics for the delineated basin; (6) download GIS data displayed in the map; and (7) print the map.

The map legend frame is at the right of the map frame, and shows 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 allow users to turn on and off the display of the layers. A locator map can be shown instead of the map layers by clicking on the Locator Map tab at the top of the frame.

Above the map layers frame is the Zoom To tool, which allows users to zoom to an address, a water feature fromthe National Hydrography Dataset (NHD) (http://nhd.usgs.gov/), or a geographic name from the Geographic Names Information System (GNIS) of the USGS (http://geonames.usgs.gov/). A scale bar above the Zoom To tool allows users to zoom to a specific scale.

Streamflow statistics database

The streamflow statistics database, named StreamstatsDB, was developed using Microsoft Access. The database has a custom user interface that allows for easy entry and update of data. The initial release of StreamstatsDB contains 144 types of physical and climatic basin characteristics and 492 types of streamflow statistics. Each 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 was released internally to each of the 48 District offices within the Water Discipline of the USGS so local staff can populate the database with whatever previously published data are available. District staff can easily add any new basin characteristics or streamflow statistics that they need in the database.

StreamStatsDB has not yet been linked to the StreamStats prototype we application. When the linking is completed, users will be able to click on data-collection sites shown on the map and get information for the sites from the database.

GIS database

The GIS database contains map layers needed to assist users in locating sites of interest and map layers needed to measure basin and climatic characteristics used as explanatory variables in the regression equations used to estimate the streamflow statistics. 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 used in StreamStats include:

- Digital Raster Graphics (DRG) (http://mcmcweb.er.usgs.gov/drg/) scanned USGS topographic maps are used as the primary base layer for precise site selection
- National Elevation Dataset (NED) (http://edcnts12.cr.usgs.gov/ned/),
- Elevation Derivatives for National Applications (EDNA) dataset (http://edcnts12.cr.usgs.gov/ned-h/),
- National Hydrography Dataset (NHD) (http://nhd.usgs.gov/),
- Watershed Boundary Dataset (WBD) (http://www.ftw.nrcs.usda.gov/huc_data.html),
- 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)

StreamStats requires three data layers to determine watershed boundaries for ungaged 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 ENDA data are used in the Idaho prototype.

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.crwr.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 was developed as a Dynamic Link Library (DLL), a bundle of Visual Basic subroutines that operates within ArcGIS 8.1. The DLL transfers information to and from ArcIMS to allow implementation over the Web.

When users select an ungaged site along the stream network, the coordinates for the site are sent to a computer on which the StreamStats automated GIS procedure is running. The procedure then determines the watershed boundary for the location from the DEM up to the points at which the DEM-defined boundary coincides with existing catchment boundaries. 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 ungaged sites from the raw DEM and minimizes processing time.

StreamStats displays 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. Users should verify that the boundary determined by StreamStats is correct before proceeding with estimating streamflow statistics for the site.

After the user verifies the drainage-basin boundary, StreamStats determines the applicable watershed and climatic characteristics and solves the regression equations to obtain estimates of streamflow statistics for the ungaged site. The estimates are computed by the National Flood Frequency (NFF) computer program (Ries and Crouse 2002), which is linked as a subroutine to StreamStats. Information on NFF can be obtained and the program can be downloaded at http://water.usgs.gov/software/nff.html.

Output from StreamStats is provided in a pop-up Web browser window. The output consists of the date, the coordinates of the site, the measured watershed and climatic characteristics, the estimated streamflow statistics, and either standard errors or 90-percent prediction intervals as indicators of the reliability of the estimates. Approximately two-thirds of estimates for ungaged 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. The estimates assume natural flow conditions at the ungaged site.
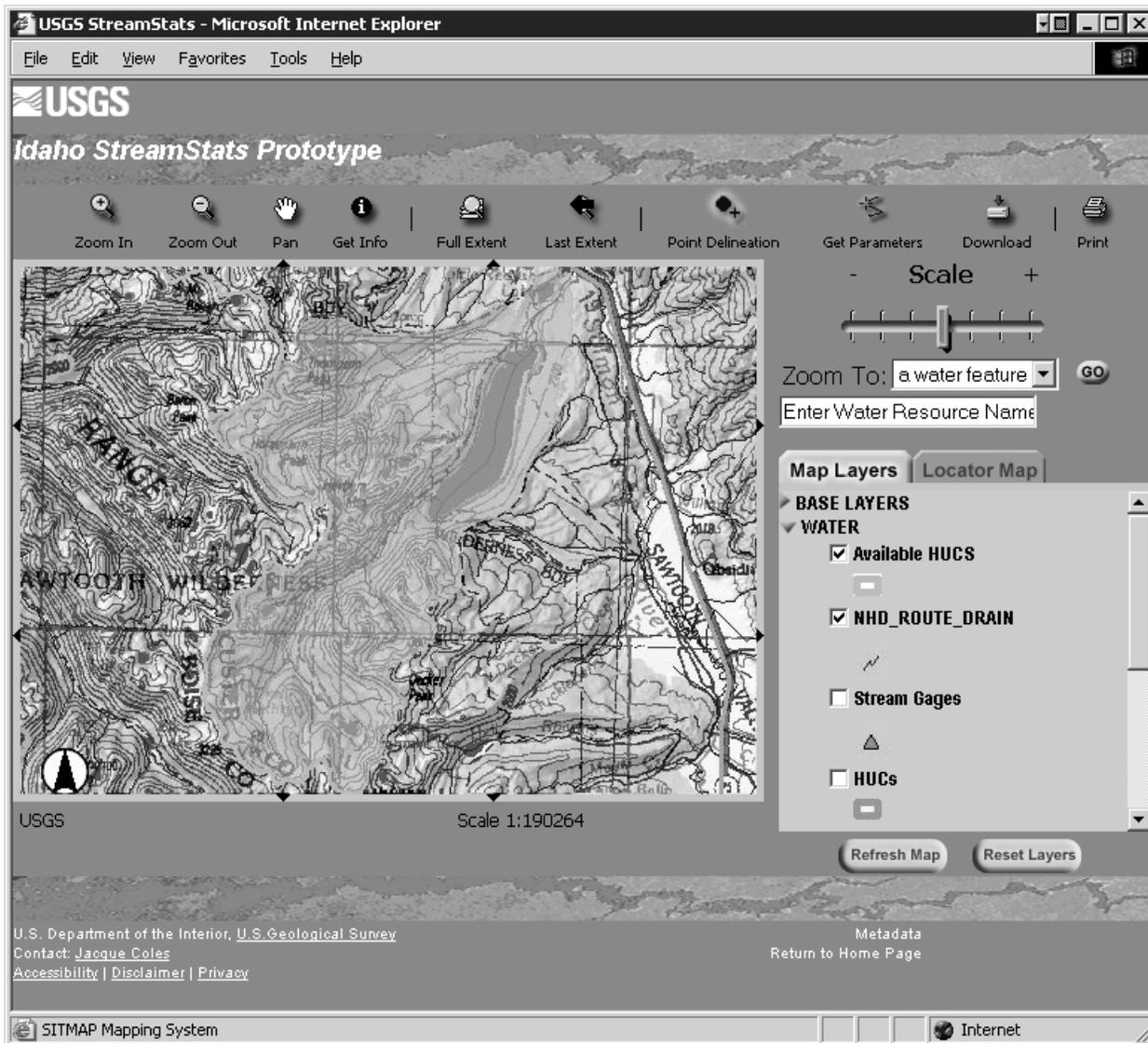
Figure 1. View of national StreamStats prototype user interface showing a delineated drainage basin for a selected site imposed over a digital topographic map.

StreamStats Implementation

Work by the USGS District offices will be needed to implement StreamStats. Usually, cooperative agreements for cost sharing with state and local agencies will be needed to complete the work.

Steps involved in implementing StreamStats include at least (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. Because regression equations for many states were developed using basin characteristics that were not measured using a GIS, it is likely that new equations will need to be developed before the ungaged site process can be implemented for these states.

After StreamStats is made available to the public for Idaho, work will begin on implementing it for other

313

states. Massachusetts, New Hampshire, and Vermont are the next states scheduled for implementation. Though several states have started the implementation process, full national implementation is expected to take several years to complete. Additional information on the development approach and implementation plans for StreamStats can be obtained at http://water.usgs.gov/osw/programs/streamstats.html

Note

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.

Acknowledgments

The authors appreciate the reviews of Peter Steeves and David Morganwalp.

References

Elassel, A.A., and V.M. Caruso. 1983. USGS digital cartographic standards: Digital elevation models. U.S. Geological Survey Circular 895-B.

Ries, K.G., III, and M.Y. Crouse. 2002. The National Flood Frequency Program, Version 3: A Computer program for estimating magnitude and frequency of floods for ungaged sites, 2002. U.S. Geological Survey Water-Resources Investigations Report 02-4168.

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