PROJECT PROPOSAL

A STREAMFLOW STATISTICS WEB APPLICATION FOR KENTUCKY

Submitted to the

Kentucky Transportation Cabinet  
Department of Highways (KYTC)

and the

Kentucky Energy and Environment Cabinet, Department for Environmental Protection, Division of Water (KDOW)

Contact: Gary Martin

Prepared by the  
U.S. Geological Survey,

Kentucky Water Science Center

**July 2012**

**Objective**

The objective of this investigation is to modify and extend the functionality of a USGS internet-based hydrologic-design computer application, known as StreamStats, to quickly and easily provide estimates of design streamflow statistics for gaged and ungaged stream sites in Kentucky. Currently, StreamStats provides estimates of rural peak-flow frequency statistics only for five of seven peak-flow hydrologic regions in Kentucky. StreamStats will be modified to provide estimates for ungaged sites of rural peak-flow frequencies for the remaining two hydrologic regions, and estimates of low-flow and mean-annual flow statistics for the entire State.

**Background**

Regression equations often provide the most accurate estimates available of streamflow sta­tistics for ungaged stream sites where no hydrologic data are available. These regression equations are developed by statistically relating calculated streamflow statistics for a group of data-collection sites to measured physical and climatic characteristics of the drainage basins for the sites. Esti­mates of streamflow statistics for an ungaged site can be obtained by measuring the appropriate physical and climatic characteristics for the site and entering them into these regression equations. Recent advances in GIS technology, coupled with increased availability of map data in digital form, now permit automated techniques for measurement of drainage-basin characteristics and calculation of corresponding design streamflow statistics.

The U.S. Geological Survey (USGS), in cooperation with Kentucky Natural Resources and Environmental Protection Cabinet, Department for Environmental Protection, Division of Water (KDOW), recently completed updating the regional regression equations for estimating low-flow statistics (Martin and Arihood, 2010). Previously, the USGS developed regional regression equations for estimating the magnitude and frequencies of rural floods (Hodgkins and Mar­tin, 2003) and mean annual flows (Martin, 2002) in Kentucky in cooperation with the Kentucky Transportation Cabinet (KYTC). And the USGS developed regional regression equations for estimating the magnitude and frequencies of urban floods in Jefferson County, Kentucky in cooperation witht the Louisville and Jefferson County Metropolitan Sewer District (Martin and others, 1997). (Development of updated regression equations for estimating flood characteristics in rural and urban streams in Kentucky is proposed in cooperation with KYTC, and update of the regression equations for estimating harmonic-mean streamflow is proposed in cooperation with KDOW.)

**Problem and Need**

Planning, permitting, design, and operation of hydraulic structures including bridges, culverts, and other drainage facilities and water- and wastewater-treatment facilities depends on reliable, accurate estimates of streamflow statistics, such as the peak-flow frequencies, mean annual streamflow, and the 30-day mean low flows for recurrence intervals of 2 and 5 years (30Q2 and 30Q5) and the 7-day mean low flows for recurrence intervals of 5, 10, and 20 years (7Q2, 7Q10, and 7Q20) and the probability of zero low flow. Streamflow statistics are needed for locations where streamflow data are collected as well as for ungaged locations where no data are available. The U. S. Geological Survey (USGS) periodically publishes various stream­flow statistics for data-collection sites. Streamflow statistics have been published for more than 400 stream sites in Kentucky. However, users of streamflow statistics may find it difficult to obtain needed statistics for sites of interest because they may be required to search through several reports to locate the statistics, and in some cases, the needed reports may be difficult to obtain, if they are out of print.

Historically, estimating streamflow statistics by use of regression equations, which may be somewhat complex, has been problematic because measuring the drainage-basin characteristics needed to solve the equations can be difficult and time-consuming. Most physical drainage-basin characteristics were formerly measured by hand from various maps, a process that could take from several hours to days to complete for a single site. Use of manual methods to measure basin characteristics from maps may not provide accurate, reproducible results. In addition, maps needed to measure the physical characteristics are often not easy for users to obtain.

Recent equations for estimating selected streamflow statistics have been developed for streams in Kentucky using digital Geographic Information System (GIS) technology to derive explanatory variables (Martin and Arihood, 2010). The use of a GIS can greatly improve the efficiency of estimating streamflow statistics and may result in estimates with higher precision than users can obtain by use of manual methods; however, proper application of the equations with GIS-based explanatory variables depends on use of GIS databases and techniques that are equivalent to those used to measure the drainage-basin physical and climatic characteristics that were used in developing the regression equations.

An automated, internet-based application for obtaining drainage-basin characteristics and calculating design streamflow statistics at user-selected sites in Kentucky is needed to provide the required hydrologic-design information quickly and reliably, but without requiring large investments in computer hardware and software, and without requiring advanced knowledge of computer science, geographic analysis, or hydrol­ogy. Such an application would ensure that basin characteristics used as explanatory variables in the regression equations are computed correctly. In addition, the application would make previously published streamflow statistics for streamgages readily available to the public.

The USGS has developed and continues enhancement and extension of a national Web-based application called StreamStats (Ries and others, 2008) that provides published streamflow statistics to the public and facilitates the estimation of streamflow statistics for ungaged stream sites. StreamStats is an integrated web-based GIS application that uses ArcGIS Server, and the ArcHydro Tools (Maidment, 2002) to make the process of computing streamflow statistics much faster, more accurate, and more consistent than previous manual methods. In short, StreamStats allows a user to select any point on a stream via a Web-based GIS interface and obtain a display the area contributing drainage to that point. StreamStats then determines the drainage area and the other basin characteristics required to solve the regression equations that are available at that location, computes the streamflow statistics, and presents them in a report. The program provides an option to download a shapefile of the drainage boundary, and any computed basin characteristics and streamflow statistics, which can be imported into a local GIS and used for further analyses.

StreamStats incorporates (1) a map-based user interface for site selection, (2) a database that provides previously published streamflow statistics and other information for gaged sites, (3) a GIS program that determines boundaries of the drainage basins for ungaged sites and measures other physical characteristics from GIS data layers, (4) the National Streamflow Statistics Program (Ries, 2006), which solves equations for estimating streamflow statistics at ungaged sites, even for drainage basins in more than one hydrologic region, and (5) ancillary GIS data layers to assist users in selected sites of interest, such as Digital Raster Graphic (DRG) topographic maps, and coverages of stream gages, roadways, and political boundaries. More information about the StreamStats application can be found at: <http://water.usgs.gov/osw/programs/streamstats.html>.

The USGS has been developing StreamStats since 2001. Although considerable amounts of effort and money have gone into development of the StreamStats application, it is being left to each State’s Water Science Center to obtain funding for local implementation through cooperative agreements with State and (or) local agencies. To date, 24 states have fully implemented applications, 3 states are partly implemented, and 11 states are in various stages of implementation. Kentucky StreamStats is partly implemented—developed for five of the seven rural peak-flow regions that use drainage-area-only regression equations for estimating peak flow statistics (See <http://water.usgs.gov/osw/streamstats/kentucky.html>). ). StreamStats in Kentucky was built using the detailed 1:24,000-scale National Hydrography Dataset (NHD) in combination with a digital elevation model (DEM) with grid points spaced 10 meters apart. Further information about the NHD can be found at <http://nhd.usgs.gov>. Rural peak-flow regression equations for regions 1 and 4 were not implemented because those equations included the main channel slope as an explanatory variable. Computation of this basin characteristic requires the preparation of specialized digital datasets for which funding was not available when Kentucky StreamStats was initially implemented.

In addition to the capabilities described above, StreamStats has network-navigation functionality that allows searching upstream or downstream along the stream from a selected site to identify connected stream reaches and to locate and provide information on natural or manmade features, such as dams and wastewater discharges, that may affect the quantity or quality of the flow in the stream. This functionality is not fully implemented in the existing StreamStats application for Kentucky. In particular, full implementation of network navigation functionality will allow the estimation of streamflow statistics for ungaged sites to be based on the flows per unit area for the statistics at nearby upstream and downstream streamgages. This estimation method can be more accurate than estimates obtained from regression equations when the drainage area for the streamgage is between 0.5 and 1.5 times the drainage area for the ungaged site. Also, a StreamStats application with the ability to summarize upstream of water discharges and withdrawals has been developed for Maryland ([Ries](http://pubs.usgs.gov/sir/2010/5111/) and others, 2010).

The presence of karst-terrain drainage features in certain regions of Kentucky creates uncertainty regarding actual drainage area at some stream sites. More than 50 percent of land area in Kentucky is underlain by rocks with potential for karst development and 25 percent of the land area is mature, well developed karst terrain (Crawford and Webster, 1986). Hydrologic discontinuities in karstic terrain, such as large springs and sinks, can cause large variations in streamflow over short reaches. Accurate determination of actual drainage area in karst terrain on the basis of topographic divides alone is difficult, if not impossible, in some areas of Kentucky. Supplemental information concerning significant, known subsurface flow paths, or at minimum, delineation of areas of extensive, mature karst development where stream drainage-basin bound­aries are not known with certainty, is needed to avoid making unreliable estimates of the design streamflow statistics. Currently, there is a karst layer in StreamStats and warnings are given to alert users when they have selected a basin that has karst.

**Approach**

Major components to be developed for the proposed StreamStats hydrologic-design toolset and database for Kentucky include (1) a data base of previously published streamflow statistics (low flows, peak flows, and mean annual flows) for data-collection sites in the State, (2) a GIS database that contains the digital map data needed for measuring the basin characteristics, and (3) GIS-derived measurements of main-channel slope for 45 basins in flood regions 1 and 4 and a validated, or revised as appropriate, set of regression equations for estimating peak-flow frequencies in flood regions 1 and 4.

Streamflow statistics for gages reported in Martin (2002), Hodgkins and Martin (2003), and Martin and Arihood (2010) will be compiled and entered into the NSS database used by StreamStats. This will allow these statistics to be readily available to the public and also will allow the transfer of the statistics for the gages to the locations of nearby user-selected sites based on the flows per unit area at the gages. Once implemented, the tool that will allow this functionality is labeled “Estimate Flows Based on Similar Streamgaging Stations” in the StreamStats user interface.

The Kentucky rural peak-flow estimating equations used two independent explanatory variables (total drainage area and main-channel slope) in two of the seven flood regions (Hodgkins and Martin, 2003). Main-channel slope was originally measured primarily by use of paper topographic maps. Adoption of a new GIS-based methodology for determining main-channel slope requires that the flood-frequency equations be re-evaluated to determine new coefficients, if required, in the estimating equations. To do this, new GIS slope measurements will be made for each of 45 basins in Kentucky flood regions 1 and 4 used in the previous regression analyses. The values of main-channel slope measured by use of the GIS will be com­pared to values previously measured manually from paper maps. The values will be analyzed to define the magnitude of any differences and bias in the measurements. The regression-equation estimates of peak-flow statistics derived by use of the values of basin characteristics measured manually from paper maps and values measured by use of GIS methods will be compared simi­larly. New regression analyses, if required as appropriate, will be performed for these flood regions. The new regression equations will be evaluated to ensure that all assumptions required for a valid regression are met and to determine prediction errors. The revised estimating equations and prediction errors, as appropriate, will be published in the final report that will also describe the use of StreamStats.

USGS will coordinate the implementation of StreamStats for the KYTC for the peak- and mean-flow estimating equations with the implementation of StreamStats for the KDOW for the low-flow estimating equations. USGS will develop and test the StreamStats program for the estimating peak-flow statistics (Hodgkins and Martin, 2003), annual mean flows (Martin, 2002), and the newly developed logistic and WLS regression low-flow estimating equations (Martin and Arihood, 2010) to calculate streamflow estimates for any unregulated, rural stream site in the State and for any size of drainage area within the limits of the data used to develop the regression equations. USGS will review StreamStats measurement of basin-characteristic values, computation of streamflow statistics, confidence intervals, and estimation errors calculated for ungaged sites and other information displayed by StreamStats for ungaged sites; and will review flow statistics and other descriptive information for continuous-record, unregulated and regulated, gages displayed by the StreamStats program. USGS will simultaneously provide KYTC and KDOW with a statewide, internal implementation of the complete StreamStats program for review and testing. River reaches within the State that are regulated will be coded as such, so that when a user clicks on a regulated river reach, StreamStats will deliver a message saying that the river reach is regulated and the regional regression equations are not applicable. StreamStats will be released for public use following publication of the final fact sheet report.

The StreamStats application provides functionality to aid the user in verifying the accuracy of the basin boundaries and to help identify potential problems, such as basin characteristics being outside the range of values used in developing the regression relations, or stream reaches being downstream from major reservoirs*. Although the application will identify potential problems, in all cases, the user will be responsible for verifying the accuracy of the delineated basin boundaries and the validity of assumptions (e.g. unregulated, natural flow) necessary for computing streamflow estimates using the regional regression equations*.

# Benefits

# Implementation of an automated program to measure basin characteristics and automatically calculate streamflow estimates for Kentucky will provide critical information and improved predictive accuracy for calculations of design streamflow statistics. A user-friendly, web-based interactive program will reduce time and effort to calculate the estimates and will avoid possible user confusion and error in the selection of appropriate regional equations and the weighting of regression estimates for basins within more than one hydrologic region.

# The USGS has developed the StreamStats program to further the agency mission of providing hydrologic information and understanding needed for the best use and management of the Nation’s water resources. StreamStats for Kentucky will facilitate the rapid, accurate, and reproducible estimation of selected streamflow statistics for gaged and ungaged stream sites in the State. Also, the program will provide an updateable and extensible platform that will provide basin characteristics values (such as drainage area, main-channel slope, etc.) and create a downloadable shapefile of the basin boundary that will be useful for further analyses in custom applications. The stream-network navigation capabilities will allow users to identify upstream or downstream streamgages, water-quality stations, dams, and wastewater discharges and obtain information about these point events, with the possibility for adding additional point events in the future, such as bridges if a state bridge layer is available.

This investigation meets the requirements of the Federal Cooperative Water Program by advancing understanding of hydrologic processes, furnishing hydrologic data or information that contribute to protection of life and property, and providing standardized, quality-assured data to national databases available to the public that will be used to advance the understanding of regional and temporal variations in hydrologic conditions that are useful to multiple parties (W. Werkheiser, U.S. Geological Survey, written commun., January 2012). Natural Hazards such as floods also have been identified as one of six strategic science directions listed in the USGS Circular 1309 “Facing tomorrows challenges—U. S. Geological Survey science in the decade 2007-2017” (U. S. Geological Survey, 2007).

**Deliverables**

The updated Kentucky StreamStats web site will be the primary product of this project. Additional products will be vertically-integrated hydrologic GIS datasets for the study area, including (elevation, hydrography, basin boundaries, and gaging station locations). These datasets will be produced as independent products as the project advances geographically. These datasets will exist within the ArcHydro framework (Maidment, 2002), which will allow for integration and use in a multitude of water-resources projects where basin characteristics as predictive variables need to be linked to a hydrologic flow-network for modeling basin and stream-network processes. A brief USGS Scientific Investigations Report will be prepared, which discusses the use and features of the Kentucky StreamStats web site, and the modifications to the rural peak-flow regression equations for hydrologic regions 1 and 4.

**Work plan and budget**

A general time line of proposed activities is shown below with reference to the State fiscal year.

The total cost, approximately $320,000, will be shared by KYTC and KDOW.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| TASK | 2013 | | | | 2014 | |
| Q1 | Q2 | Q3 | Q4 | Q1 | Q2 |
| Compile flow statistics at gaged sites and prepare database |  |  |  |  |  |  |
| Preparation of GIS coverages required for delineations and explanatory variables in regression equations |  |  |  |  |  |  |
| Implement test site for computing basin characteristics and evaluate accuracy |  |  |  |  |  |  |
| Compute main-channel slope for 45 gaged sites using GIS |  |  |  |  |  |  |
| Test/revise peak-flow frequency equations to incorporate the new measure of main-channel slope |  |  |  |  |  |  |
| Entry of coefficients in the estimating equations into NSS Program as engine for computing estimates |  |  |  |  |  |  |
| Test application for computation of streamflow statistics |  |  |  |  |  |  |
| Report writing, reviewing, and release |  |  |  |  |  |  |

**Personnel**

Project staffing will come from the Kentucky Water Science Center. Project coordination, data analysis, modeling, and GIS will be provided by Gary Martin, Randy Ulery, Jeremy Newson, and Jeremiah Lant, Hydrologists, and Hugh Nelson, Geographer.

**REFERENCES CITED**

Crawford, N., and Webster, J., 1986, Karst hazard assessment of Kentucky--Sinkhole flooding and collapse: Boling Green, Ky., Western Kentucky University, Center for Cave and Karst Studies, prepared for the U.S. Environmental Protection Agency, Region IV, Atlanta, Ga., scale 1:1,000,000, 1 sheet.

Hodgkins, G.A. and Martin, G.R., 2003, Estimating the magnitude of peak flows for streams in Kentucky for selected recurrence intervals: U. S. Geological Survey Water Resources Investigations Report 03-4180, 68 p. Accessed 10/20/2011 at <http://pubs.usgs.gov/wri/wri034180/>.

Martin, G.R., Ruhl, K.J., Moore, B.L., and Rose, M.F., 1997, Estimation of peak-discharge frequency of urban streams in Jefferson County, Kentucky: U.S. Geological Survey Water- Resources Investigations Report 97-4219, 40 p. Accessed 07/10/2012 at http://pubs.er.usgs.gov/publication/wri974219/.

Maidment, David, ed., Arc Hydro: GIS for Water Resources: ESRI Press, Redlands, California, 203 p.

Martin, G. R., 2002, Estimating mean annual streamflow of rural streams in Kentucky, U. S. Geological Survey Water Resources Investigations Report 02—4206, 35 p.

Martin, G.R., and Arihood, L. D., 2010, Methods for estimating selected low-flow frequency statistics for unregulated streams in Kentucky: U.S. Geological Survey Scientific Investigations Report 2010-5217, 83 p.,

Ries III, K.G., 2006, The National Streamflow Statistics Program: A computer program for estimating streamflow statistics for ungaged sites: [U.S. Geological Survey Techniques and Methods Report TM Book 4, Chapter A6](http://pubs.usgs.gov/tm/2006/tm4a6/), 45 p.

Ries III, K.G.; Guthrie, J. G.; Rea, A. H.; Steeves, P. A.; and Stewart, D. W., 2008. StreamStats: A water resources Web application, U. S. Geological Survey Fact Sheet 2008-3067, 6p.

[Ries, K.G., III, Horn, M.A., Nardi, M.R., and Tessler, S., 2010, Incorporation of water-use summaries into the StreamStats web application for Maryland: U.S. Geological Survey Scientific Investigations Report 2010–5111, 18 p.](http://pubs.usgs.gov/sir/2010/5111/)