Comprehensive Flow Analysis Using Cloud-Based Cyberinfrastructure

Tyler Wible and Mazdak Arabi, Department of Civil and Environmental Engineering, Colorado State University

In order to facilitate a better understanding of streamflow analysis and take advantage of recent improvements in Web technology, a cloud-based Web tool was developed to analyze the multiple aspects of streamflow and water quality data.

Introduction

Streamflow data are widely used by water managers, engineers, scientists, and policy makers to support a variety of environmental and water resources decisions. Water data at varying spatial and temporal scales provide information for flood control, drought management, water allocation, water quality control, and ecological assessment. Despite the development of complex models of the environment, analyses of water data from stream monitoring locations remain essential for sustainable management of water resources and environmental protection.

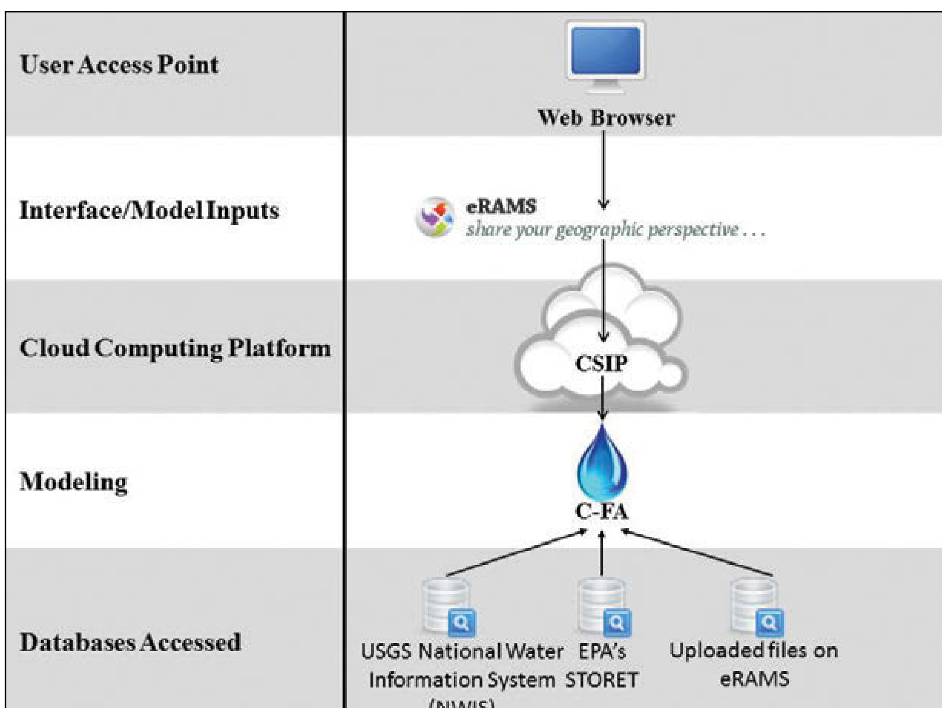
In the Unites States, extensive historic and real-time water quantity and quality data for many rivers and streams may be obtained from the U.S. Geological Surveys (USGS) National Water Information System (NWIS), U.S. EPA STORET, and other similar agencies. However, lack of accessibility along with limited scalability of water analytics impede the broader use of these data by interested stakeholders. Existing technologies have extensive hardware requirements, require installation of platform-dependent software, and often require substantial training.

However, these Web-based tools offer few if any ways to analyze the streamflow data. To work around this deficiency a person would have to find, download, and install software to analyze the streamflow data. Additionally, there are numerous aspects to streamflow analysis, and currently, no comprehensive software package to analyze all of them. This then requires someone to become proficient in each individual streamflow analysis software package, its installation procedure, computer requirements, and the creation and format of its input files.

In light of recent advances in Web development and modeling services, the next step in streamflow analysis is a Web-based platform-independent software tool. This tool could also auto-extracts data from the aforementioned databases. One option for a flexible, scalable, Web-deployment of environmental models is the use of Infrastructure-as-a-Service (IaaS) clouds. Cloud-based modeling inside a Web tool offers a platform-independent package requiring no software from the user beyond their Web browser.

Infrastructure Outline

The purpose for this study was to develop a Web-based, platform-independent, stream water analysis tool. The result of this work is the Comprehensive Flow Analysis (CFA) tool designed and implemented on the environmental Risk Assessment and Management System (eRAMS) platform (www.erams.com/flowanalysis). The eRAMS technology provides a platform for seamless integration data with complex modeling systems using a cloud computing infrastructure. Major components



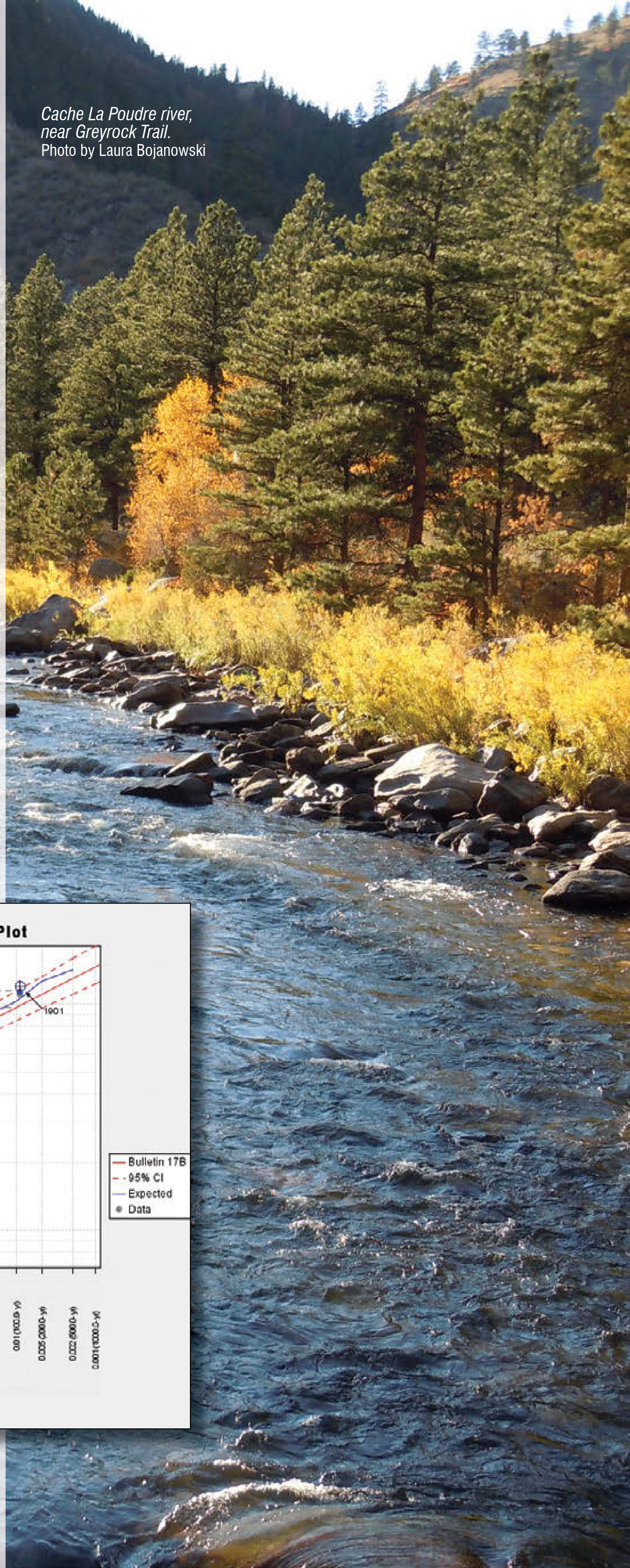
The Comprehensive Flow Analysis tool infrastructure of eRAMS include a content management tool, a collaboration platform, a mapping and geospatial analysis tool, and platform for model integration. eRAMS uses the Cloud Service Integration Platform (CSIP) to access the flow analysis models of CFA. The combination of cloud based computing from CSIP with the website services of eRAMS allows Web-based access to the streamflow analysis models of CFA from any desktop, laptop, smart phone, or tablet devices.

Analysis Capabilities

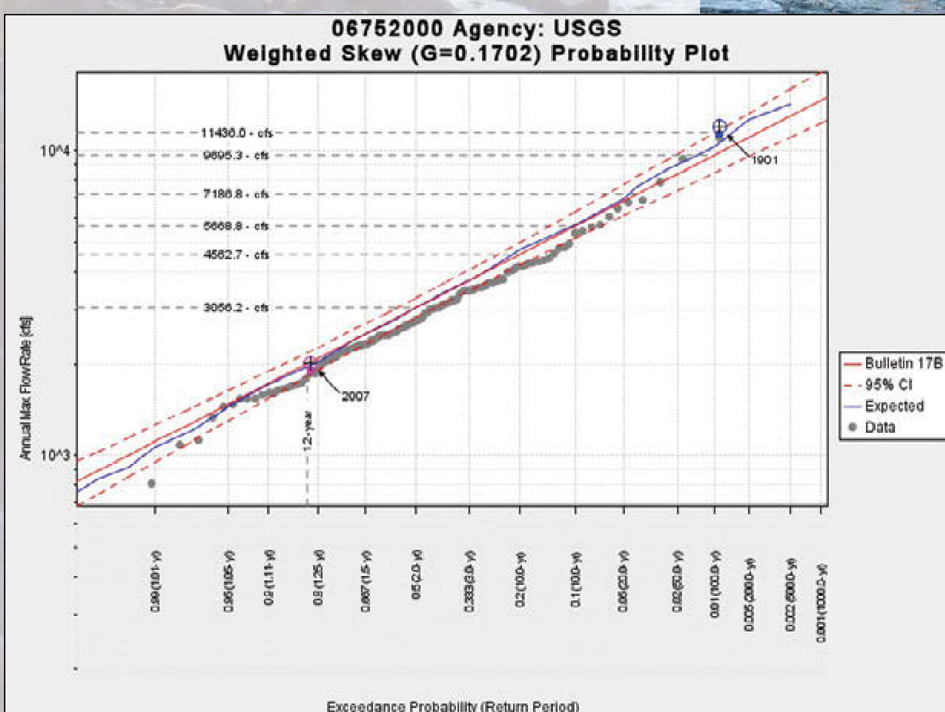
The CFA's streamflow analysis models allow a person to examine the various aspects of streamflow, including flood and drought frequency, duration curve analysis of streamflows, as well as base flow separation tools which isolate the influence of groundwater on stream discharge.

The first model in CFA is a simple time series analysis which graphs and summarizes basic statistics on streamflow or water quality concentrations. The next model in CFA is a flood analysis method based on the procedures outlined by the USGS for flood frequency analysis. The flood model calculates the recurrence interval of historic floods as well as predicting confidence limits for future flood magnitudes based on the historic data.

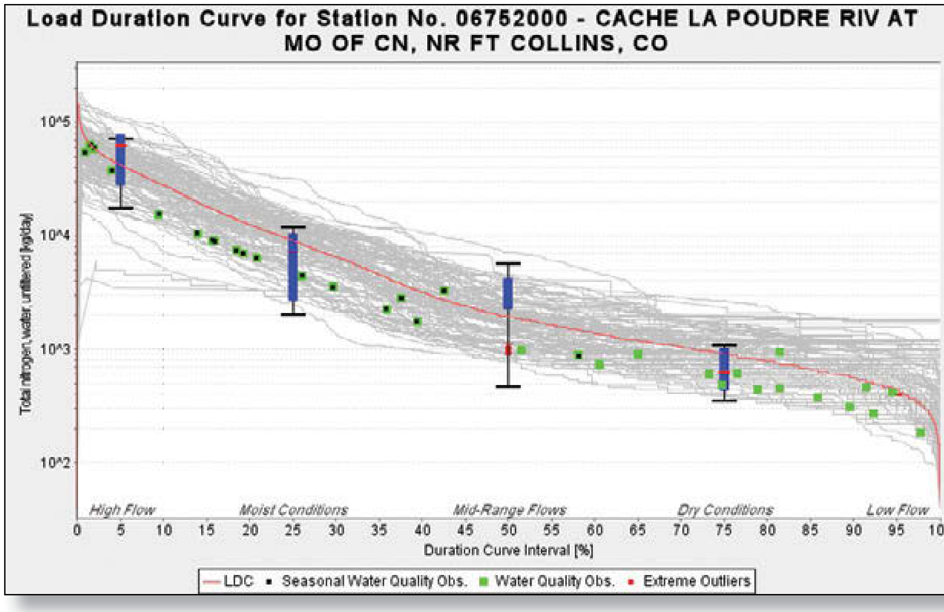
An opposite, but equally important, aspect of river flood flows is the consideration and analysis of droughts from streamflow records. For this reason a generalized



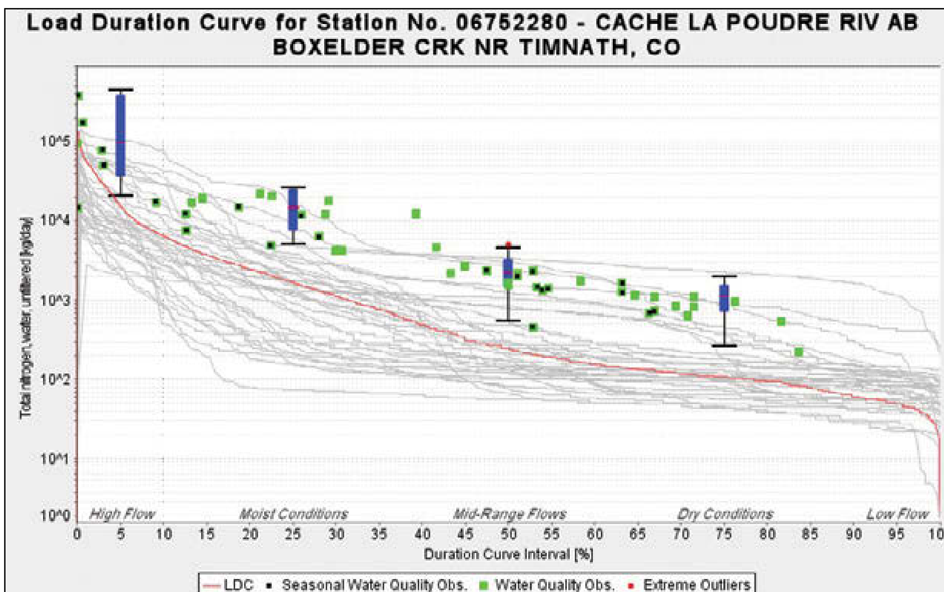
Cache La Poudre river, near Greyrock Trail.
Photo by Laura Bojanowski



An example of flood modeling, a tool based on USGS procedures for flood frequency analysis drought analysis tool was included in CFA. The drought analysis method included in CFA fits Auto-Regressive (AR) or Auto-Regressive-Moving-Average (ARMA) model to annual streamflow data. The purpose of fitting the regression model to the data is to increase the size of the dataset while maintaining its statistical properties, mean, and standard deviation. After fitting the regression model, a 100,000 year forecasting is performed using the fitted model to create a dataset sufficiently large to 'observe' high recurrence interval droughts. The forecasted dataset is then analyzed and summarized based on the drought deficit (deficit = supply – demand), its length (years), and its average recurrence interval.



Models can show comparisons, such as the graphs of nutrient loading before (above) and after passing through Fort Collins. Red lines represent the total average historical limit.



Another useful aspect of streamflow analysis is river base flow, the groundwater contribution to streamflow. CFA includes the base flow separation tool BFLOW, developed by Jeff Arnold, which is an automated digital filter base flow separation tool that performs a multi-pass separation of base flow from total streamflow. A final model included in CFA is a duration curve analysis tool. This tool uses the frequency with which a given flow value is exceeded to determine a percent exceedence and create a flow duration curve—how long (duration) a flow is exceeded. This flow duration curve can also be combined with water quality observations to create a load duration curve, useful in determining total maximum daily loads of pollutants into the river.

These CFA models were developed so that a person could take a river, like the Cache La Poudre River in Fort Collins, and examine something like the impacts of nutrient loading before and after the river passes through the city. Let us do just that using CFA's Load Duration Curve tool. First, look at the river monitoring station at the mouth of the Poudre River canyon upstream of Fort Collins, USGS 06752000. If we compare the results for total nitrogen loading on the river to the EPA's legal standard of 10 mg/L, notice that almost all of the samples are below the legal limit, which is desirable, before the river enters the city. The red line is the total average limit over the history of streamflow discharges; the grey lines are per year limits. However, if we examine the same information for the river monitoring station just downstream of Fort Collins, USGS 06752280, notice that almost all of the nitrogen samples are above the EPA's limit, indicating that the city increases the concentration of nitrogen in the river (see graphs).

Closing Remark

The purpose of this work was to develop a standardized, publicly available, Web tool for flow analysis. Beyond this, the tool was designed so that public utility companies, watershed modelers, water resource managers, and average citizens can use the tool to learn more about the watershed they live in.