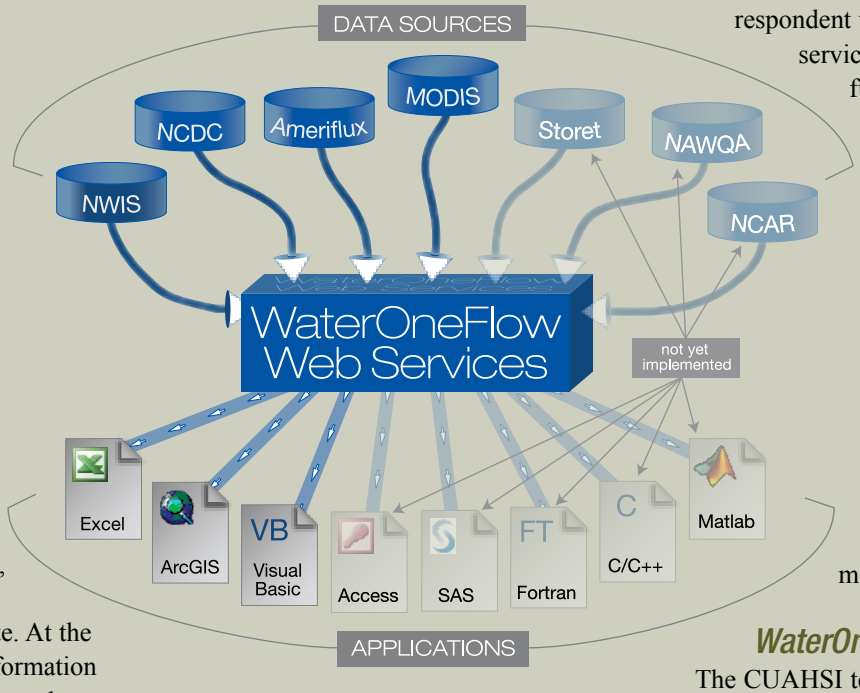
Hydrologic Data Access using Web Services

David R. Maidment, Ph.D. *– Center for Research in Water Resources, University of Texas at Austin;* ***Ilya Zaslavsky, Ph.D.*** *– San Diego Supercomputer Center, University of California, San Diego; and* ***Jeffery S. Horsburgh*** *– Utah Water Research Laboratory, Utah State University*

When a hydrologist wants to study a watershed, stream, or aquifer, one of the first tasks is to assemble pertinent information, including data on precipitation, streamflow, water quality, groundwater levels, and climate. At the national level, much of this information is available on public websites, such as the USGS National Water Information System (NWIS), the National Climatic Data Center's Climate Data Online, and the EPA Storet system for water quality. Additional information is available from the web pages of state and local water agencies. All these websites are designed uniquely, each with its own method of presenting data. The end result is that the hydrologist sets off on an Easter egg hunt searching through all these information sources trying to discover how each one operates and eventually assembling a set of output files in various formats that then have to be homogenized into the format needed to support a particular analysis. More extensive data sources, such as Nexrad precipitation or MODIS remote sensing, have large files in arcane formats that require significant effort before they can be accessed and used.

CUAHSI's (Consortium of Universities for Advancement of Hydrologic Science) Hydrologic Information System team developed the Hydrologic Data Access System (HDAS) with the aim of providing a common "data window," or web portal, to these disparate data sources. The idea is for a hydrologist to be able to quickly find out what information is available for a particular region, select the required data, then import that into an application such as Excel, ArcGIS, Matlab, or a programming language such as Visual Basic. The magic needed to make this happen is an innovation in computer science called "web services," which is based on a set of protocols developed by the World Wide Web Consortium and enables one computer to request services of another.

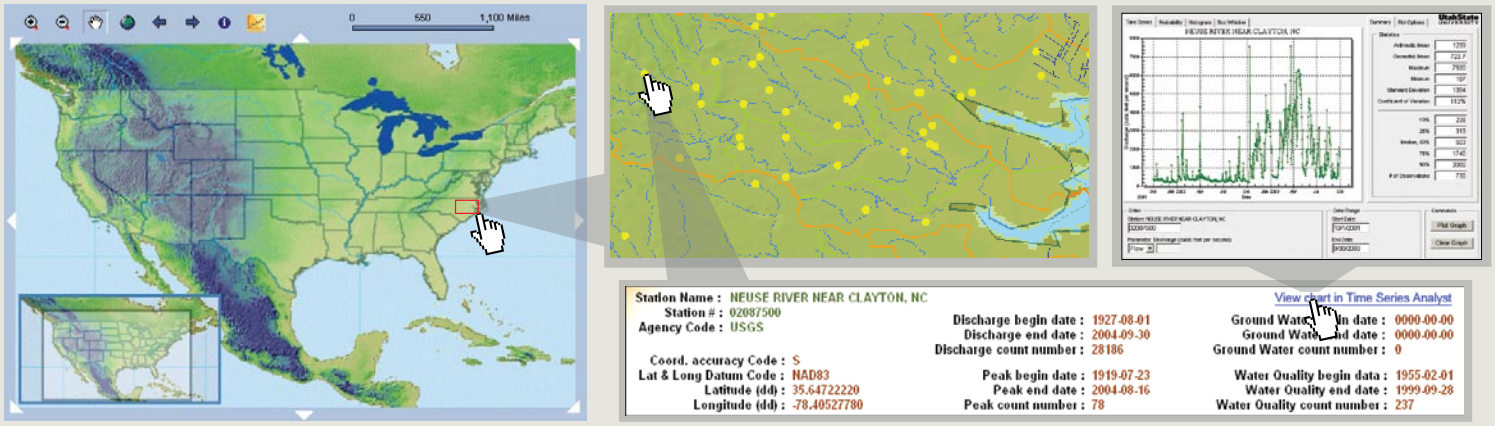


A rough analogy can be made with e-mail: when we write an e-mail message, we don't worry about what software or operating system the recipient will use to read the message, we only know their e-mail software will somehow properly interpret the sent message to allow the respondent to read it and reply. Web services perform this mediating function among computers in a more general way and enable particular functions to be executed, such as "get stations" to identify the locations where measurements have been made, "get parameters" to determine what parameters have been measured there, and "get data" to obtain the measured information itself.

WaterOneFlow

The CUAHSI team has also developed WaterOneFlow, a web services library that enables users to extract data from national water data archives, transform the data to an appropriate format, and load it into common working environments (see figure above). HDAS and NWIS Time Series Analyst (discussed below) are two applications that use WaterOneFlow web services to facilitate data analysis.

Users have quickly recognized the value of web services. The availability of web services for NWIS was announced during a CUAHSI cyberseminar on Oct. 28, 2005. On Nov. 2, Jason Love, from RESPEC, a private firm in Sioux Falls, South Dakota, posted an announcement on the EPA Basins list server about what a valuable tool web services for NWIS is, and included a tutorial he had developed on how to use the services with Matlab. Thus, technology transfer from academia to the private sector to the public sector occurred in less than one week!



The Hydrologic Data Access System: locating a site (left), querying a station for available data (center), and viewing time series data (right).

Hydrologic Data Access System

HDAS is a web portal that provides a map interface to WaterOneFlow web services (see figure above). Users can zoom in anywhere in the United States, query a station and find out what variables are measured there, the period of record, and the number of measurements made. The information needed to respond to these queries is extracted from the data sources using web service requests and is locally compiled to form a database unique to each hydrologic observation network. This work is complete for the streamflow gauging stations in the NWIS, and will be accomplished later for other data sources accessible through WaterOneFlow web services.

Having discovered which variables are measured at a station and when, a user can retrieve a time series for a selected variable directly from NWIS through HDAS. HDAS contacts the "get data" method of the NWIS web service of the WaterOneFlow library, supplying the parameter code and the "from" and "to" dates as inputs. The NWIS web service then queries the USGS NWIS archive and returns the requested time series data. The time series data are then displayed as a graph and a table of values within HDAS, and can be downloaded in Excel or CSV formats. We are working on expanding its capabilities to serve other national hydrologic data archives as well as various local observation data sets.

Links

Hydrologic Data Access System
river.sdsc.edu/HDAS/

Web Services for NWIS
river.sdsc.edu/NWISTS/nwis.asmx

Time Series Analyst
water.usu.edu/nwisanalyst/

USGS National Water Information System
waterdata.usgs.gov/nwis/

National Climatic Data Center's Climate Data Online
hurricane.ncdc.noaa.gov/CDO/cdo

U.S. EPA Storet System
www.epa.gov/storet/

NWIS Time Series Analyst

Another application empowered by WaterOneFlow web services is the NWIS Time Series Analyst developed at Utah State University. This Internet-based application graphs hydrologic data as time series, cumulative frequency curves, frequency histograms, and box and whisker plots, and provides summary statistics about the data. The Time Series Analyst was originally programmed to query data from a local database into which was loaded data from NWIS and other sources for the Bear River watershed of northern Utah. When WaterOneFlow web services became available, the Time Series Analyst was adapted to acquire data from NWIS using web services, and in so doing, the Time Series Analyst became applicable anywhere in the nation without any special database compilation. It now effectively uses the national data archive as its local database. And, the Time Series Analyst is now available online so that it can be accessed from and applied anywhere in the nation, vastly increasing its value. Furthermore, the Time Series Analyst is now integrated with HDAS: when users select a station, variable, and measurement interval in HDAS, they can request advanced time series charts to be generated by the Time Series Analyst application. The two components remain at different locations, in San Diego and Logan, Utah, but work in concert and rely on the common set of WaterOneFlow web services.

Putting It All Together

HDAS and WaterOneFlow are the first available components of CUAHSI's Hydrologic Information System, a geographically distributed system of functions and datasets that are connected through the Internet. These tools allow a variety of national data archives to be directly queried in a consistent manner, and user applications can be written that automatically operate on data measured anywhere in the nation. By exposing tools through web portals, applications such as HDAS and the Time Series Analyst are available to users anywhere. Combining local applications, web services, and web portals into a connected Hydrologic Information System offers great potential. However, building information systems using web services is a new innovation and a significant amount of work remains to make these systems robust, efficient, and comprehensive.

The authors acknowledge the contributions of their colleagues in the CUAHSI HIS team, and the support of this research by the National Science Foundation grant EAR – 0413265. *Contact David Maidment at maidment@mail.utexas.edu, Ilya Zaslavsky at zaslavsk@sdsc.edu, or Jeffery Horsburgh at jeffh@cc.usu.edu.*