International Environmental Modelling and Software Society (iEMSs)
2012 International Congress on Environmental Modelling and Software
Managing Resources of a Limited Planet, Sixth Biennial Meeting, Leipzig, Germany
R. Seppelt, A.A. Voinov, S. Lange, D. Bankamp (Eds.)
http://www.iemss.org/society/index.php/iemss-2012-proceedings

Global Multiscale Hydrologic Web Services

Steve Kopp, Dean Djokic, Nawajish Noman
Environmental Systems Research Institute, Redlands, California, U.S.A.
skopp@esri.com ddjokic@esri.com nnoman@esri.com

Abstract: There are two problems slowing widespread adoption of spatially enabled hydrologic analysis, accessibility of data, and ease of use. This paper will describe and illustrate a system for sharing hydrologic analyses (geoprocessing) services and workflows that empowers GIS professionals, scientists, and citizens by making water data and its analysis accessible to a wide audience. The system leverages recent evolutions and standards in the information technology and GIS industries, particularly web services accessed through simple geospatial browser applications, cloud computing, and the increasing amount of elevation and hydrologic observations data. The GIS foundation of the system includes a high cartographic quality base map, an attributed vector hydrologic network, a hydrologically conditioned DEM, and linkage to temporal observation data. The system will provide a set of analytic services bound to best-available, authoritative data, and include tools and workflow documentation to enable others to more easily build similar systems. A key feature of the system is the development and serving of hydrologically conditioned elevation data at multiple scales. These datasets are the basis of geoprocessing services for interactive watershed delineation and flow tracing. Geoprocessing overlay services combine these areas with descriptive landscape information to explain or predict water quantity and quality. Extensive national and global datasets exist for land cover, soils, precipitation, and water observations, and are combined to provide hydrologic information in a highly accessible and interactive way. This system of web services is useful for GIS professional who may lack the time or skills needed to assemble and process all the needed data layers. But more importantly the system is critical to a large community of scientists, professionals, and citizens who want or need more detailed and useful water information in an easily accessible way.

***Keywords**:* GIS; Geoprocessing; Web services; Hydrology.

1 INTRODUCTION

This project builds upon a strong history of integrating GIS and water resources. There have been numerous excellent conference series emphasizing this topic, all starting in the early 1990s, which aided its fast and continued evolution; the American Water Resources Association specialty conference on GIS and Water Resources, NCGIA GIS and Environmental Modeling, HydroGIS in Vienna, Austria, and the Esri International Users Conference GISHydro preconference seminar series. A recurring theme was to find commonality among applications with the goal of creating a framework for integrating environmental models.

In 1999 Esri initiated a community effort to design a common data model and tools for GIS applications in water resources. That project, now known as Arc Hydro, created a template geodatabase model, and collection of data processing and analysis tools for the water resources community. Maidment [2002]. This project was quite successful in providing a common framework and tools for the GIS community working in water resources. Over the years it has been extended and is used extensively for support of various water resources applications. It is also a foundation for hydrologic and hydraulic model integration allowing efficient integration of engineering numerical models with GIS. Ackerman et al. [2010], Dunne et al., [2000].

The earliest web-based integration of GIS and hydrologic analysis to grow beyond a research project was StreamStats in 1999. Ries et al., [2008]. StreamStats was unique in providing custom hydrologic analysis for a single purpose, estimating flow using approved regression techniques. It has a high level of use and professional status, and has become recommended or mandated by law for use in some states. Pennsylvania, [2009], Massachusetts, [2008]. It is now in its 3rd version release as it evolves with changes in GIS and information technology. More recent projects and web applications developed combining GIS analysis and hydrology include CUAHSI, GeOnAS, and Maryland GIS Hydro, to name a few.

To accelerate the availability of spatially enabled hydrologic analysis this project will create easy to use analytic services bound to authoritative, curated data. It uses and teaches best practices for performance and integration with other applications. It provides a comprehensive software framework for authoring, sharing, and serving hydrologic data, maps, and analysis.

2 PROJECT COMPONENTS

2.1 Data

To achieve a vision of truly distributed web GIS, where users consume services provided by one another, it is imperative that the service being consumed is authoritative and reliable. Authoritative services may come from governmental or non-governmental agencies or their contractors, and are recognized within their geography and application domain as a respected provider. There are many places on the internet to find elevation data and measurement data, but finding the data you need and understanding its quality is difficult. To empower authoritative sources to process and serve their data this project provides a streamlined approach to prepare their map and elevation data and host their own web services. Workflows and services are designed to work at global, regional, and local scales and provide a seamless user experience across scale ranges.

The project data consists of an attributed vector hydrologic network, a hydrologic base map service, hydrologically conditioned elevation data and derivatives, and water resources observation data.

The stream network and base map are based upon NHDPlus (USEPA and USGS [2006]), and DEM derived streams with additional attribution. The elevation data is based upon the 90 meter resolution HydroSHEDS dataset (Lehner et al. [2008]) covering much of the world, 30 meter resolution NHDPlus elevation data in the US, 20 meter resolution CEM 2.0 elevation data for Mexico from INEGI, and 10 foot resolution LiDAR derived elevation data for Williamson County, Texas.

2.2 Services

Web services, such as map services and geoprocessing services solve two problems for the GIS community. First, they make it possible to provide GIS professionals with ready-to-use data and solutions. Second they provide an easier way for non-GIS users to access and leverage GIS capabilities.

Map and Feature Services

Map services provide a way to deliver high cartographic quality maps very quickly. They are constructed from a collection of prerendered map tiles stored on a server as images which are tiled together into the display when requested. Quinn and

Gehegan [2010]. The tiled map cache has become an industry standard approach used by Google, Bing, ArcGIS, and others.

One reason map services are popular and seem easy to use, is that they work at many map scales. As the map scale changes, the amount of information in the map changes. This is achieved by creating many different maps, each for a specific viewing scale. For this project design template a tile cache was created with 14 maps from 1:147 million to 1:18,000. These maps include streams symbolized by mean annual flow and stream names.

Figure 1. Examples from the multiscale global hydrologic base map.

Geoprocessing Services

Geoprocessing services easy-to-consume approach to providing powerful analytic capabilities. They are authored with the ArcGIS Desktop and can be consumed in ArcGIS Desktop, a wide variety of Esri web client APIs, or in other products through use of industry standards.

Input features such as gage locations can be entered by drawing features on the web map, by querying a feature service, or by getting the results from another geoprocessing service. Input raster may be accessed by referencing a URL to a raster file, by querying an image service, or by getting results from another geoprocessing task. This pattern is true for all the data types, and all OGC service protocols are also supported.

Figure 2. Geoprocessing service data flow

Geoprocessing services can be chained in the client or server as illustrated in Figure 2. When chaining in the client the results of one task can be used as input to another task. For example the query service to find a stream gage, can chain to a watershed delineation service which creates a polygon, which can be used as input to an overlay service to calculate mean annual precipitation over the area. Services can also be chained on the server using the geoprocessing authoring environments (ModelBuilder and Python). This enables combining geoprocessing services with other services and local tools. This new tool can then be published as a geoprocessing service, simplifying web application development.

2.2.1 Foundation Services

There are two types of analytic services developed in this project, foundation services and solution services. Foundation services are building blocks that encapsulate a generic GIS task into an easily reusable, configurable component. Foundation services are used by GIS professionals and others within their desktop application or in conjunction with other web services to build solutions. These services shorten the learning curve while ensuring best practices are applied, increasing overall productivity. This project created a collection of 3 types of foundation services; watershed delineation, flow path tracing, and overlay.

To foster the use of best practices and reuse, a foundation service consists of two parts, the geoprocessing tasks workflow to construct the base data, and the geoprocessing tasks that are run as a web service against this data. Having well defined base data tailored to specific services provides great opportunity to improve the performance and scalability of the web service. These tools and workflow are key to others being able to apply this methodology to other geographies. The tasks and workflows are shared as geoprocessing packages as described in section 3.1.

The single largest obstacle to anyone who wants to model the flow of water across the landscape is to modify the elevation data to mimic how water flows. Elevation data contain errors and inaccuracies that need to be addressed for proper modeling to be possible. This "hydroconditioning" of elevation data requires the user to understand a bit of hydrology, geomorphology, how elevation data is constructed, how the analytic tools operate against it, and also often some local knowledge of drainage patterns. This is too much for most people and prevents further progress. The Arc Hydro tools made it easier to do, but the user still needed deep understanding. This project seeks to leverage the skills of those who have the expertise by providing a platform through which others can use their work. We believe that this will result in significant cost savings for governments and significantly increase access to hydrologic flow analysis by a much broader audience.

The project has defined three resolution workflows for preprocessing, which are distinguished by differences in the elevation hydroconditioning process. Generally these are greater than 90 meter, 10-30 meter, and less than 5 meter. For projects that fall in between they can choose an appropriate workflow based upon documented criteria, influenced by the source of the data and the geomorphology of the landscape being modeled. The results of preprocessing are stored in the same data model regardless of resolution, therefore the services scripts work seamlessly against them.

Hydrologic modeling is based upon knowing what is upstream, therefore watershed delineation is the key foundation service. Overlay services can then use the watershed to intersect with precipitation, land use, or other variables and summarize this information for input to a runoff or water quality model.

Performance and scalability of watershed delineation is an important and interesting subject both with respect to core operation and delivery of results to the client. As the area of interest increases and the resolution of DEM increases, the watersheds not only become larger in extent, but their geometry (number of vertices) increases in complexity.

We have developed a workflow and collection of scripts that are openly available as geoprocessing packages (see Section 3.1) to help users of the system build their own high performance watershed delineation web service. This includes a collection of data preprocessing scripts as well as the scripts to enable delineation and other analysis on the server against this preprocessed data. In short, the preprocessing divides up the landscape by defining major drainage basins and using those as processing units to minimize the quantity of data to be accessed when delineating large basins.

This methodology allows watershed delineation for any point in the preprocessed extent within a few seconds. Computation is consistently under 5 seconds when the result watershed contains less than 100,000 vertices. As a reference, this is a watershed of approximately 500,000 km2 derived from a 30 meter resolution elevation model. For areas with a larger number of vertices the performance non-linearly decreases. We continue working on expanding this technique to further improve scalability with increasing of watershed boundary complexity.

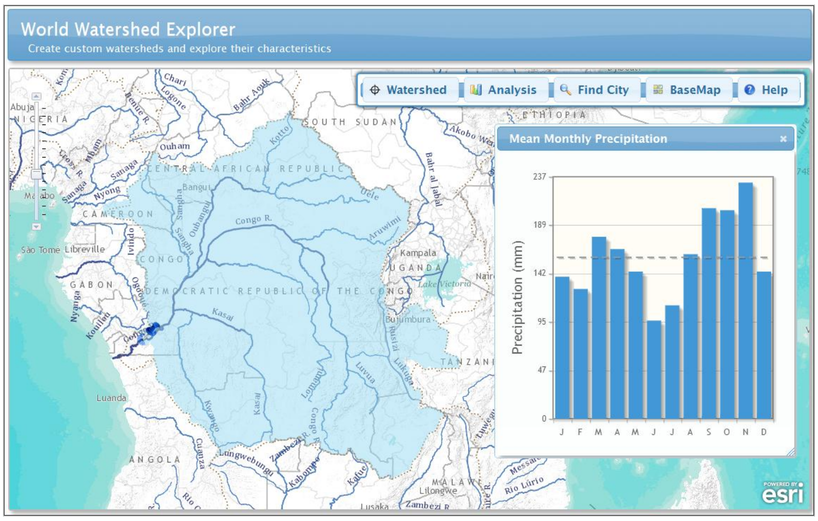

Figure 3. Congo basin with mean monthly precipitation

High performance of the interactive web application during watershed delineation is achieved in two ways. The application is designed to support delineation at multiple spatial resolutions, and since we are working in the context of a map, the default behavior is to use the most appropriate resolution based on the map scale of the current display. For example, if the map extent is currently half of the United States, it is not necessary to use the 30 meter resolution DEM, the 90 meter DEM provides sufficient detail. Similarly when the map is zoomed into a small area of a few hundred square kilometers, the application will use the best available data, such as the 10 meter DEM of Williamson County, Texas. The second technique for maintaining performance of the web application is to initially return a map service (a picture of the result) instead of an actual feature containing lines and vertices, which is more data and therefore more time transmit. If the user wants the features they are available.

Additional services are also provided for tracing of flow paths, both overland flow and in-stream flow. These services provide upstream and downstream tracing from a user-specified location. The lines returned from these services can be overlaid with the land use, soils, and other data and summary information calculated.

To drive parameterization of hydrologic models the system provides an initial set of overlay services that will continue to expand in geography and scale. Currently available data for the US includes the USGS NLCD land cover data, the USDA SSURGO soils data, and the PRISM mean annual rainfall data.

2.2.2 Solution Services

Solution services are web services designed with functionality and graphical interface tailored to the needs and expertise of a particular community of people. They provide a simple user experience by limiting the number of choices, with only the capabilities they need, in language they understand. These services make hydrologic inquiry and basic analysis more approachable for a broad community of scientists and citizens.

These services like others discussed above are provided with source code for the geoprocessing service as well as the user interface to make them easily configurable for other applications. Similar initiatives to build solution services and provide them as configurable components are happening across many industries and are designed to be tied together to create new custom applications.
For example, common operating picture dashboard for flood response may use a geocoding service, a reverse 911 notification service, an evacuation routing service, and flood modeling service.

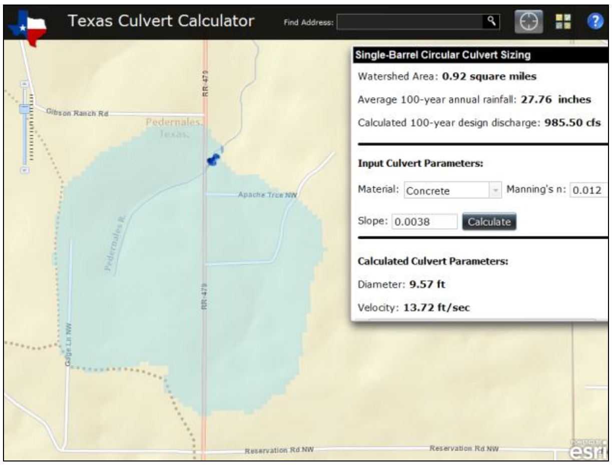

Figure 4. Interactive runoff calculation and culvert sizing application

A specific and relatively simple example is an application developed for estimating flow at ungaged road crossings to determine the appropriate size of culvert to use. This uses a methodology similar to the StreamStats application mentioned earlier, tailored to a specific task.

The only tools in the application are a geocoder for zooming the map to an address, and a button for the user to click where they want to make the flow estimation. The user does not need to know where to find the data, how to condition an elevation model, or how to use a GIS. The application returns to him the information he needs in a format that is relevant to his work.

3 Software Framework for sharing and serving spatial data and analysis

The purpose of this project is to make it easy for people to package, serve, and consume spatial (especially hydrologic) data and analysis. This requires a software system for authoring maps and analysis, serving maps and analysis, and a collection of supporting source code and workflows which users can configure into custom applications tailored for specific geographies and applications.

There are two patterns of sharing in this system, sharing between GIS professionals, and sharing from the GIS professional to a broader community. Sharing between GIS professionals is done with *packages*, and web services. Sharing with the broader community is done through web services with a simple application specific user interface.

3.1 Framework for Sharing

Sharing between GIS professionals of data and analysis tradecraft often involves acquiring someone else's work and modifying it for your use, such as using someone else's watershed boundaries to overlay with your data, or taking someone's runoff estimation code and modifying it to run against your data in your location. To provide a consistent pattern for transfer of data and analysis between

professionals, we developed a simple wizard driven packaging technique. This packaging ensures that data and analysis being prepared for delivery is complete and will work on the other end when consumed. Packaging makes it easy to consolidate data and scripts from multiple locations into a single file, avoiding common problems of missing data and incorrect pathnames to files, and ensuring that a minimum amount of descriptive information about the package is provided.
Creating packages can be done by simply right clicking and stepping through a wizard, or by using one of the packaging geoprocessing tools. The tools provide finer control over parameters and also enable packaging to be scripted into an automated workflow.

There are several kinds of packages, the ones of interest for this project are layer packages, map packages, and geoprocessing packages. A layer package is used to share a single dataset and its symbology and labeling. A map package is used to share multiple datasets that are usually part of a project. A geoprocessing package is used to share a geoprocessing workflow including all of its tools, data, and optionally presentations or other related materials.

Consuming a geoprocessing package can be done by double clicking the package file or by dragging and dropping it into ArcMap. Packages can be shared through the free hosting in ArcGIS Online, as email attachments or like any other file.

3.2 Framework for Serving

ArcGIS Server provides the service framework for this project. It provides a number of GIS service types (map, geocode, geoprocessing, etc.) with properties and methods that communicate through simple object access protocol (SOAP) or representational state transfer (REST). This project uses the GeoServices REST Specification which Esri [2010] released in an open document, and which all developers are invited to use in their own GIS server implementations. Server administrators can also choose to expose their services through standards as defined by the Open Geospatial Consortium (OGC). Maps can be served as web map services (WMS), and analysis tasks as web processing services (WPS). Vector data can be served as web feature services (WFS) and raster data as web coverage services (WCS).

Creating services is similar to creating packages, by right clicking a context menu or using a geoprocessing tool and filling in parameters as illustrated in Figure5. Map and geoprocessing services are commonly used in web applications and can also be used directly inside ArcGIS Desktop just like any other data on disk or tool in the toolbox.

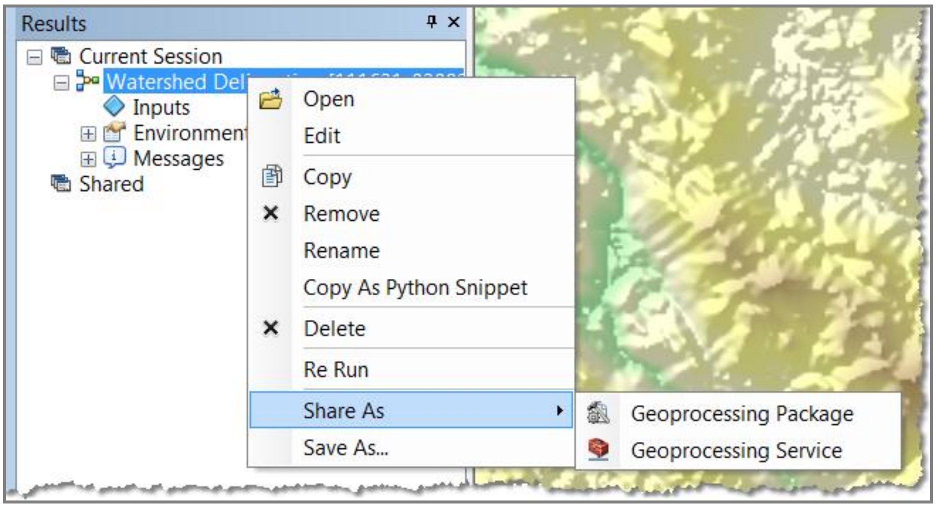

Figure 5. Creating a Geoprocessing package or Geoprocessing service

4 CONCLUSIONS AND FUTURE WORK

This project has created a framework and methodology for authoring, serving, and consuming hydrologic geoprocessing services and additional hydrologic analysis. It provides ready-to-use services on a collection of freely available data, as well as templates and workflows for authoritative curators of data to build and host their own services.

The true value of this project will be seen in the future collaboration with the GIS and water resources communities in standing up services for their geography, and developing additional analytic capabilities which leverage this foundation.

ACKNOWLEDGEMENTS
The authors would like to acknowledge the assistance of Kris Verdin of the USGS for her guidance in use of the HydroSHEDS data and further refinement of this data she provided.

REFERENCES

Ackerman, C.T., Jensen, M.R., and Brunner, G.W., Geospatial capabilities of HEC-RAS for model development and mapping. In: Proceedings 2nd Joint Interagency Conference, Las Vegas, NV. 2010.

Dunn, C.N., Evans, T.A., Ackerman, C.T., Doan, J., Hydrologic modeling for the Sacramento and San Joaquin basins using GeoHMS. In: Proceedings 2000 Esri International User Conference, San Diego, CA. 2000.

ESRI., GeoServices REST specification version 1.0, Esri, Redlands, California, USA , 220 pp. http://www.esri.com/library/whitepapers/pdfs/geoservices-rest-spec.pdf (accessed 12 March, 2011). 2010 .

Lehner, B., Verdin, K., Jarvis, A., New global hydrography derived from spaceborne elevation data. Eos, Transactions, AGU, 89(10): 93-94. 2008.

Massachusetts., Massachusetts wetlands protection act regulations 310 CMR 10.00, Preface Appendices, 53pp. 2008.

Maidment, D.R., *Arc Hydro: GIS for Water Resources*, Esri Press, pp. 203. 2002.

Pennsylvania., Water management plan example format instructions for Marcellus shale gas well development. Pennsylvania Department of Environmental Protection, 5500-PM-OG0087. 7pp. 2009

Quinn, S., and Gahegan, M., A predictive model for frequently viewed tiles in a web map, *Transactions in GIS*, 14(2), 193-216, 2010.

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

USEPA and USGS (US Environmental Protection Agency and US Geological Survey). NHD Plus. US Environmental Protection Agency, Washington, DC. http://www.horizon-systems.com/nhdplus/index.php (accessed 8 May 2012). 2006.