INTegrating Water Resources Information Using GIS and the Web

Jack Dangermond and David Maidment

ABSTRACT: GIS is a vital tool for building a digital information base for water resources management and analysis. Thus far, the focus in application of GIS in water resources has been on processing and synthesis of geospatial data layers and building GIS preprocessors for water simulation models. Besides desktop and server-based GIS, a new field of web-based GIS is emerging, supported by broader bandwidth, faster computers and massive data storage. Organizations are creating geo-services on the web to open access to their information. A new challenge in water resources is to use GIS and the web to synthesize access to water observations data – the time series of flow, water levels and water quality about surface and groundwater resources that are presently held in a myriad of independently managed tabular databases maintained by water agencies. The Consortium of Universities for the Advancement of Hydrologic Science, Inc (CUAHSI) has invented a language, WaterML, for the transmission of water observations data through the internet, and has built a national water metadata catalog for water observations data in the United States. By geo-enabling and extending the CUAHSI model for water data services a new web-GIS for water resources can be created. This will enable the easy integration of large volumes of water data and complex models into simple to use applications that become pervasive.

INTRODUCTION

The use of GIS in water resources has a rich heritage. Some key contributions include: base maps of water features of the landscape, digital elevation models delineating watersheds and stream networks, geospatial data preprocessors supporting water resources simulation models, floodplain maps built using automated hydrology and hydraulics, and precisely measured land surface elevation using LIDAR (Tarboton, 1997; Maidment and Djokic, 2000; Wilson and Gallant, 2000; Maidment 2002). GIS serves as the key to integrate these layers of geospatial information, and to provide a framework for water resources management and enhanced environmental management. The AWRA biennial series of Specialty Conferences in GIS in Water Resources has served as an important forum for creating and sharing ideas, concepts and experience – for example, it was at the Nashville Conference in 2004 at which the ideas solidified that now form the basis of the NHDPlus dataset integrating the national coverages of Hydrography, Elevation, Land Cover and Watershed Boundaries for the United States. These geospatial datasets describe with national coverage the physical environment through which water flows.

Separate data archives exist to describe the properties of water itself – observational data about water flows, levels and quality that are measured at point locations using gages, or by taking water samples and analyzing them in a laboratory. Water observations data are collected by federal, state and local water agencies to support water management and by researchers to support advancements in water science. The US Geological Survey’s National Water Information System stores water observations data measured at about 1.5 million locations and is by far the nation’s largest repository of water resources data. The EPA’s STORET database contains water quality information collected at about 400,000 locations by state and local agencies to assess the health of the nation’s natural water systems. The National Climatic Data Center is the nation’s repository of precipitation, weather, and climate information. The USDA Natural Resources Conservation Service collects and publishes data about snow and soil water. Hundreds of state and local water agencies collect water observations data about the water resources in their jurisdiction. Many agencies share water observations data as part of their core mission through web sites. Water data web sites and clearinghouses have improved access but these efforts haven’t resulted in an integrated geospatial system for water information because each web site has its own interpretive structure and output data format, and there is little integration across organizations to thematically synthesize water data, as has occurred in the GIS field (Figure 1).

There is a strong demand for integrated water information. Comprehensive water management involves flood management and emergency response, long term water assessment to deal with droughts, water quality management to make the nation’s waters swimmable and fishable, and environmental modeling to ensure the ecological integrity of the nation’s rivers, lakes and estuaries. Looming over all these traditional water management challenges is the uncertainty of the impact of climate change on water resources – will sea level rise and changes in temperature and precipitation alter coastal and inland water environments significantly? How can advances in GIS help meet these challenges?

*Respectively: President, Environmental Systems Research Institute, Redlands, CA; Director, Center for Research in Water Resources, The University of Texas, Austin, TX.
GIS TRENDS

We are becoming a geodata rich society with more geospatial information, applications and access. Satellite imagery is expanding; GPS and locational data are more easily acquired; georeferenced information is becoming more available; real time change monitoring and locationally aware devices are spreading. Online GIS for mapping and visualization is improving consumer access to information. Our global society is increasingly spatially literate. This trend is benefiting the water resources community. More than $1 billion has been invested in FEMA floodplain mapping since 2003, perhaps the nation’s largest investment in mapping in recent years. LIDAR measurement of land surface terrain has emerged as a key new information source for supporting floodplain mapping, and for improving the National Elevation Dataset.

GIS is providing integrating technology, bringing together geospatial data and knowledge, providing methods for analysis and collaboration, and making them widely accessible. GIS applies the geographic approach, providing tools, methods and workflows that support collaboration and action. This leads to better decisions, greater efficiency, more effective communications, thus improving the way we do things. GIS technology is changing, becoming easier and more pervasive as new styles and techniques emerge. These include new media such as vivid web maps, real-time data, user-generated content, social networks, mobile GIS, web mashups – in short a set of distributed services that create many opportunities for leveraging water resources data (Figure 2).

Today, GIS is being implemented in multiple patterns that are evolving with advances in technology. Faster computers, broader internet bandwidth, massive storage, mobile resources and cloud computing are creating new environments that enhance using, sharing and authoring geographic information. Desktop, server, enterprise, and mobile GIS are increasingly integrated with the web. GIS server technology is fundamental to this evolution, and desktop-centric GIS now benefits from data provided by GIS servers within an enterprise and from servers distributed across the web.

Desktop GIS integrates water resources with widely deployed community tools such as the Arc Hydro data model (Maidment, 2002) for surface water hydrology. This is accompanied by the Arc Hydro toolset, now comprising more than a hundred tools for tasks such as terrain analysis, watershed and stream network delineation, and stream network tracing. A companion data model, Arc Hydro Groundwater, has been defined (Strassberg et al., 2010), and a new toolset for implementing that data model is now available, including a special extension for integrating ArcGIS and Modflow. GIS preprocessors, such as HEC-GeoHMS and HEC-GeoRAS, have been produced, and a discipline of automated hydrology and hydraulics has emerged to support floodplain mapping. Powerful, web-based applications such as Streamstats, have been created to allow probabilities of flow to be computed anywhere on a stream network. A large community of users of GIS in water resources exists that is constantly evolving and improving the application of GIS in this field.
While Desktop and Server GIS will continue to play a key role in the future, a new pattern of web-centered GIS is emerging. This involves GIS-ready services that supply data, maps, imagery and models, and community sharing of information through search and discovery portals. Online GIS now provides a rich array of base maps for hydrology, topography, roads, imagery, political boundaries and many other data layers. Desktop, Server and Online GIS form an integrated system that uses the internet as its platform. This new computing paradigm helps to integrate GIS through a services-based architecture that takes advantage of cloud computing and integrates diverse content through Web 2.0 based mash-ups. Organizations are increasingly creating geo-services, moving from data sharing to creating shared services (Figure 3). These include RESTful services that are easily discovered, support open standards, and can work with easy to use client applications and free application tools. These shared geospatial resources are opening access to many new users and applications. REST (REpresentational State Transfer) is an architectural style for the design of applications that make use of distributed resources through the internet (Fielding., 2000).

Organizations are Increasingly Creating Geo-Services

Opening Access to Many New Users & Applications

In contrast to the normal accessing of a web page, web services enable the user to access a network of distributed services that function as a collective whole. This is creating a new geospatial infrastructure supporting a large community of users and applications – linking government, business, citizens and education. This supports open access and transparency.
of information, and better collaboration. A national GIS or spatial data infrastructure is emerging as a distributed network of systems and services that integrates national, regional and local geospatial knowledge.

VISION FOR WATER RESOURCES

This vision of an interoperable linked geospatial infrastructure can also be applied for water resources data. A key challenge is to integrate the nation’s geospatial data describing the water landscape, such as watersheds, stream networks, aquifers, wells, gages and sampling sites, with the time series of observations data describing the flow, level and quality of water – in other words synthesizing the GIS description of the physical landscape with properties of the water that flows through that landscape (Figure 4). Only in this way can the interaction of the human and physical factors that influence water flow and quality be fully understood.

However, the current situation is complicated. Water observations data are stored in many distributed tabular databases, each having its own output data format. Commonly measured variables such as streamflow or dissolved oxygen are labeled differently from one organization to another. The tabular databases are independently managed, not spatially enabled, and have no over-arching community or sponsor. In large water agencies, it can occur that data for different geographic regions are managed independently from one district or field office to another. The implication of this vast heterogeneity of water data systems is that data access and integration is laborious, so much so that perhaps 80% of an analyst’s time is spent acquiring and processing data into useful forms before analysis can be carried out. As a result, water resources information is not leveraged as much as it could be, and big problems are not addressed effectively. For example, the operators of water supply reservoir systems maintain daily water balance records of the inflow, storage and release of water from their reservoirs, but this information is fragmented in many forms at many locations. As a consequence, the effect of climate change on the nation’s surface water supply reservoir network is difficult to assess.

Figure 4. Connecting GIS data with time series of water observations.

CUAHSI Water Web Services Standards

Many organizations are now collaborating to share information using the water web services data standards established by the Consortium of Universities for the Advancement of Hydrologic Science, Inc (CUAHSI). CUAHSI is an organization representing more than 120 US universities, which was established with the support of the National Science Foundation in 2001 for the purpose of creating infrastructure and services to advance hydrologic science. As part of this mission, the National Science Foundation has also supported, since 2004, a CUAHSI Hydrologic Information System project whose aim is to provide better access to and synthesis of the nation’s water data, http://his.cuahsi.org (Maidment, 2008, 2009). CUAHSI has invented a language called WaterML for communicating time series of water observations through the internet (Zaslavsky et al., 2007), and agencies such as the US Geological Survey and the National Climatic Data Center have adopted WaterML and now publish some of their data using this format.

CUAHSI is building an online water information community through registration of water data services at a single location, HIS Central, located at the San Diego Supercomputer Center. CUAHSI has designed a web-enabled observations data model, provides data and interoperability standards, maintains a national metadata catalog for water observations data, and is collaborating with ESRI to host an online community water map. This includes information about water flow and level in surface water bodies, soil water information, precipitation, water quality, groundwater levels, meteorology, and related forms of water data such as daily reservoir operations, and water diversions by pumping from surface and
groundwater systems. All the information from these diverse sources is communicated in a single language, WaterML (Figure 5).

Figure 5. USGS streamflow data published in the WaterML language.

The CUAHSI model for water web services leverages three essential elements of the internet: users, servers and catalogs (Figure 6). When the internet emerged as an information source, users employed a web browser and accessed web servers at known locations to acquire text and pictures in the HyperText Markup Language (HTML). Some years later, cataloging services, such as Google, Yahoo and Bing emerged, which greatly increased the utility of web information, and now it is commonplace to launch a search query using Firefox or Internet Explorer on a catalog such as Google, and then sort through the list of web sites that results from the search to get the information needed from a particular web site. CUAHSI has replicated these three functions—user, server, and catalog—and integrated them through its WaterML language (Horsburgh et al., 2010).

Since WaterML web services are public information sources, it does not matter what server or operating system produces them, nor does it matter what kind of user application receives them—just so long as both server and client conform to CUAHSI web services standards. These standards are being reviewed by a Domain Working Group in Hydrology, jointly sponsored by the Open Geospatial Consortium and the World Meteorological Organization, which has a goal of harmonizing them with existing OGC and WMO data standards to create a WaterML Version 2 that may become an international standard for water observations data transmission through the internet.

What this means is that water data services distributed over a state, a region, the nation, and even the world, can be integrated and accessed through different approaches. Traditionally, these include clearinghouses such as Geodata.gov and Data.gov, and search engines such as Google, Yahoo and Bing. CUAHSI has its own portal, http://hiscentral.cuahsi.org, that lists 47 water data services available nationally describing 15,000 variables, measured at 1.8 million locations, and comprising 9 million time series containing 4.3 billion data values, all accessible in the WaterML language (Figure 7). Other organizations, such as the Texas Water Development Board, are setting up similar CUAHSI water data services for state level water data sources (Whiteaker et al., 2010). Online GIS services can leverage these services by providing maps of observations data sites for water data in their region, and links to the observations information accessible at each site.
CUAHSI has also developed a HydroServer, based on Microsoft SQL Server and ArcGIS Server to allow a university or workgroup to host, organize and publish their own data (Horsburgh et al., 2009). The HydroServer stores point observations data in a relational data model called the CUAHSI Observations Data Model (Horsburgh et al., 2008), and publishes this information in WaterML, together with map-based information published as geospatial data services. Users may acquire information from HydroServers using the open source CUAHSI HydroDesktop application, http://www.hydrodesktop.org.

It is thus possible to create a web-GIS for water resources with a portal whose catalog supports data search, basemaps and content sharing, for clients such as desktop, browser and mobile applications, and provides distributed services for models, data and maps, for both commercial and open source applications. Web-GIS for water resources will extend the CUAHSI model by leveraging the rich capacity of GIS for integrating information by location across organizations that provide water data in a region. The CUAHSI data model supports description of site locations, time series measured there, and metadata about what variables have been measured over what time period and how many values are available (Figure 8).

The query functions include GetSites, which gives the (x,y) locations of the list of sites contained in an observations network; GetSiteInfo, which specifies the variables measured at each site, the period of record, and the number of data available; GetVariableInfo, which gives details of these variables such as the units of the data; and GetValues, which provides the time series of data itself. Because time series of water data are atomic in character – they can be queried one by one, and each one is a small data package that comes quickly through the internet, water observations data are ideally suited to a services-oriented architecture for information sharing. This simple data model has proven to be widely applicable across a very large class of water information sources.
As use of water web services expands, a number of application patterns will emerge. Desktop GIS and other desktop applications including hydrologic models will be able to access water data directly through the internet without the necessity of manual human intervention. This will support workflows, models, augment local data sources, and populate data models such as Arc Hydro. Web browser and mobile applications will emerge to leverage water data services. Situational awareness centers for emergency response will be better informed as to what is going on around them. In severe storm and flood events, locally measured data on rainfall and flood water levels can be quickly shared regionally to inform a wide range of emergency responders using public water information data services.

This environment will also open up the ability to use web services to integrate multiple dimensions of water data and modeling for use in planning and place-based approaches for projects and programs (Figure 9). These include water data and data of other types, such as land use, climate, vegetation and environmental conditions, as well as data about budgetary expenditures originating from multiple agencies focused around a place (watersheds, neighborhoods, regions, aquifers, etc). Water data incorporated within GIS services allows for the easy integration of water within this place-based context. This will lead to improved water resources planning and management through collaboration across multiple levels of government. Alternative plans for flood mitigation in an urban stream could be assessed and compared. When a potential new water user applies for a pumping permit from an aquifer, a web-based application could run a groundwater model with the potential new pumpage in the correct geographic location so as to assess its impact on aquifer water levels. Various approaches for meeting Total Maximum Daily Loads could be explored for improving the health of water bodies impaired by water quality contamination. Possible time patterns of environmental flows to preserve the ecological integrity of rivers could be defined and compared.

**GIS Supporting Place-Based Policies**

*Increasing the Impact of Water Resources Programs and Policies*

- Targeting Priority Places
- Coordinating Projects & Programs
- Developing Coordinated Plans
- Supporting Integrated Decision Making

**Using geography & GIS as a Framework for Integrated Problem Solving**

Figure 9. Using GIS for place-based problem solving in water resources.

**IMPLEMENTATION**

Web-centered GIS is being implemented in a new cloud computing architecture called ArcGIS Online, which is an online geographic information system that has two types of content: Maps and Applications, or “Maps and Apps”. Maps are ready to use high quality geoinformation that have cartography built in, they are multiscale so that more detail becomes apparent as you zoom in and less as you zoom out, and they can be used in a standard viewer such as ArcGIS Desktop or other desktop applications, a web browser, or a mobile device. The “Map” concept is more than simply a visual image – it can also include groups of interrelated data layers, tables and datasets whose content is displayed in the map or connected with it. In this sense, the “Map” can be thought of as a visual interface to the information set as well as analytical tools that lie behind it, rather like the title page of a book.

A Map is retrieved from ArcGIS Online using a “layer package”, which can automatically be opened in ArcMap to display the map image and may be associated with a zip file containing the supporting data. The layer package may also contain feature attachments of any information type, such as photos and explanatory supporting documents. The second kind of content in ArcGIS Online is Applications that provide access to information and functionality, rather than simply data. Applications use maps and services and leverage them to create new information.

**Data Carts and Themes**

A CUAHSI water data service delivers a collection of time series describing a set of variables measured on a network of observation sites over a particular time horizon by a single organization or individual. Each time series in this collection is represented by a metadata record that contains all the information needed to describe the site, the variable, the period of record, and the web service addresses from which the data can be obtained. A “Data Cart” is a tabular catalog, with one record for the metadata of each time series. A simple ArcGIS Online “layer package” consists of a Data Cart and an accompanying cartographic symbolization so that the observation sites are rendered appropriately as points and displayed over the top of a “hydro basemap” that shows the streams, waterbodies, aquifers, watersheds and other locational information.
that provide geospatial context for the observational data. A user downloads this layer package into ArcGIS Desktop, and uses Arc Hydro and ArcGIS tools to select particular sites and variables and download the resulting time series into an Arc Hydro time series database stored in the ArcGIS geodatabase. Using time-enabled feature layers in ArcGIS version 10, the time series information can be used to build maps describing the time variation of water properties over a region, and to support many other applications integrating spatial and temporal data.

The Data Cart just described is a catalog of information describing one water data service, but the concept is readily generalized – if the Data Carts describing several such services in a region, such as those from the principal water agencies, are merged, it follows that the resulting catalog provides an inventory of water information for the region. From the GIS perspective, it is very attractive to thematically classify this information – to classify the time series into groups describing streamflow, dissolved oxygen, nutrients, precipitation, salinity, evaporation, groundwater levels, and so on. In this way, time series “themes” can be constructed, which are collections of time series describing water observations of a particular type measured in a particular geographic region.

### Data Carts can be Built for Themes

*Each Organization Publishes its own Water Data Services*

<table>
<thead>
<tr>
<th>Themes</th>
<th>USGS</th>
<th>EPA</th>
<th>State Agency</th>
<th>Water District</th>
<th>University</th>
</tr>
</thead>
<tbody>
<tr>
<td>Streamflow</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Groundwater Level</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dissolved Oxygen</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nutrients</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Salinity</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Theme Data Carts Reference Data Services...
...across a Group of Organizations*

Figure 10. A Theme Data Cart references water data services across organizations.

If the observations data themselves are then harvested from the water data services referenced in the catalog, the resulting information can be inserted into a new kind of layer package, now containing not just the times series metadata but also the data itself. One way of presenting such a Theme is as an Arc Hydro time series dataset, which consists of a feature class of observation sites and three tables representing variable definitions, the series catalog and the time series data themselves. This Theme can be packaged as file geodatabase and set inside a layer package that has appropriate observation site symbolization and background basemap.

### Publishing Water Observations Data in ArcGIS Online

Information can not only be downloaded from ArcGIS Online but also uploaded. Suppose a water agency or scientist has a water observations dataset that has been loaded into a relational database, and made accessible as a WaterML web service. A water agency might be using the WISKI time series management system from the Kisters firm, which has an addon package for WaterML web services access, or a water scientist at a university might be using the CUAHSI Observations Data Model provided with the same WaterML web services to manage their data from an experimental watershed (Horsburgh et al., 2009). The water agency or scientist could decide to publish their observations in ArcGIS Online by using ArcGIS Desktop to “mashup” the data and map services to create the appropriate layer package for either just the Data Cart, or for a set of one or more fully curated data themes. These would then be uploaded to ArcGIS Online, and would be immediately available for others to download and access. In ArcGIS Online, information can be shared within a Group for their use only, or can be published for public access. ArcGIS Online supports social networks for sharing maps and geospatial information.

We intend to create a global map of water observations by federating and visually displaying on a single basemap in ArcGIS Online the observations data services and themes that have been posted to ArcGIS Online in the manner just described, and those presently registered at the CUAHSI HIS Central. The CUAHSI HIS Central already supports search services for more than 9 million time series collected mostly in the United States, which are tagged with concepts derived from a hydrologic ontology, a hierarchical organization of key terms that are used to describe the physical, chemical and biological properties of water. Providing a water data catalog within ArcGIS Online will facilitate access to water data for the GIS community and also simplify the synthesis of water observations with other geospatial layers of water information, such as watersheds, stream networks, aquifers, and the like.
QUALITY ASSURANCE

In building a federated water information system as we are here proposing, many questions arise about quality assurance of the information accessed through this system. It is not our intention to endorse or scientifically justify the information provided by water agencies and scientists through this system, but rather to simplify and make uniform the means of access of this information. It will still be up to the user to make judgments as to the degree of credibility of information from particular sources. What is carried along in the metadata for all the water observations time series in this system is the source from which the information came, and a description of the method by which the data were determined. Indexing water data services using a Data Cart requires assurance of the accuracy and currency of the metadata – are they complete and up to date? Publishing water observations Themes involves selection of observations data from various sources, and in doing that there is opportunity to filter the information for appropriateness, and to convert the units of the data to a common set of units for the variables in the Theme. Building water observations data Themes has many of the same characteristics as does building geospatial data layers – information is ingested from various sources, has to be checked, synthesized, and presented in a uniform format.

CONCLUSIONS

What we are proposing here is using GIS and the web to connect the geospatial description of the water landscape with the time-varying observations of the properties of water flowing through that landscape. This is an historic conjunction of two water data worlds that have up to this time been largely stored and processed separately. A series of observations of water level, flow, or quality, measured regularly or irregularly through time at a single point in space by a particular organization, is considered as a single entity whose data and metadata are transmitted through the internet as web services in a specially designed computer language, WaterML, developed by CUAHSI. Organizations publishing their information in WaterML can compile the metadata describing their water data services into a Data Cart, and publish that in ArcGIS Online as a layer package, containing a map of their observation sites overlaid on a standardized hydro basemap of the water landscape provided in ArcGIS Online. By federating these observations maps and Data Carts, a map-based enquiry system for water data services can be provided across water organizations. This approach also permits indexing the information by Themes, that is, collections of time series data services for particular kinds of information, such as streamflow, groundwater levels, dissolved oxygen, nutrients or salinity, and preparing Data Carts that index such Themes. Users of the water data services can compile downloaded time series data and Data Carts with other geospatial information, and upload that into ArcGIS Online as a new “mashup” of geospatial and temporal water information displayed over a customized base map. ArcGIS Online thus provides a social networking system for individuals, groups and agencies to share water information privately within a group or publicly to all users. Applications leveraging that information can also be included in ArcGIS Online so as to facilitate access to the information through web browsers and mobile applications.

Implementing this web-centered GIS for water resources requires a community approach: leadership and management support, technical architecture, creating and sharing water data services, adoption of standards, creation and support of community water data portals, and ongoing participation. This system will create a new approach transforming how water resources data is managed and used that fuses physical and socioeconomic data with water observations and informs GIS that integrates spatial and temporal information. The users of this system will include a cross-cutting, multidisciplinary
community – federal, state and local government, universities and research centers, water districts, engineering consulting firms, schools, communities, non-governmental organizations and citizen science. The key to the success of this system will be a diverse group of applications – map viewing and query, simple mashup’s, simple analysis and visualization, customized applications for particular purposes, access to server side modeling, data discovery and download for local analysis and developing a common operating dashboard for situational awareness. Water information integrated using GIS and the web will improve our ability to integrate, plan, manage, and more holistically design water resources systems.

ACKNOWLEDGEMENTS

The authors wish to acknowledge the many contributions to this work from the partners in the CUAHSI Hydrologic Information System project at the San Diego Supercomputer Center, University of Texas at Austin, Utah State University, University of South Carolina, Idaho State University, Drexel University, and the CUAHSI Program Office. They also wish to acknowledge the contributions of many individuals at ESRI to this work, in particular Clint Brown, Christine Dartiguenave, Dean Djokic, Charlie Frye, Caitlin Scopel, and Zichuan Ye. The contributions of reviewers of this paper, Jim Cameron, Sandra Fox, Jon Goodall, Jack Hampson, Richard Hooper, Jeff Horsburgh, Katherine Lins, David Tarboton and Tim Whiteaker, are appreciated.

REFERENCES


