INTEGRATED WEB-BASED DATA MANAGEMENT, ANALYSIS AND VISUALIZATION:

TOWARDS AN ADAPTIVE WATERSHED INFORMATION SYSTEM

by

IBRAHIM DEMIR

(Under the Direction of M. Bruce Beck)

ABSTRACT

Information Systems (IS) have become an important issue for watershed management. Large amounts of data have been collected, and several models and tools developed to visualize and analyze the data. Much of the data and models, collected by individual researchers or research groups, are not integrated since the data are stored in local databases, and models are not easily accessible or applicable to other environmental systems. It takes an enormous amount of work to transform and understand the data and models, so that they can be explored and utilized within other information systems. Developments in internet technologies such as web services provide interfaces for easy integration of data, digital resources, and tools from various sources. Improvements in client and server-side scripting languages (e.g. AJAX, Javascript) make interactive content (e.g. visualization tools, online mapping), user interaction, rich user interfaces and desktop-like applications for data analysis and visualization possible on web browsers. In principle, water quality can be managed through an integrative perspective on watersheds, facilitated by a web-based information system. This dissertation presents a prototype of a web-based Information System for Georgia watersheds called 'GWIS: Georgia Watershed Information

System'. The primary purpose of GWIS is to record and present the work of the Environmental Process Control Laboratory (EPCL) of the University of Georgia, whose purpose in turn was to develop the monitoring and modeling of the water environment required to move towards a "smart" urban water infrastructure. As a prototype, GWIS has been developed initially to demonstrate capabilities and boundaries of an online information system from basic to advanced tools for the research community. The system is developed in a way to make adaptation of its components and structure easier in the light of new technology. Once it is populated with the necessary data and tools, GWIS can serve a management-oriented function. Several data management, modeling, visualization, mapping and resource management tools for watersheds, as well as interfaces for integration across diverse and dispersed data sources, are included in the system. As its initial point of departure — to provide substantial and specific content — GWIS has been populated with the high-volume high-quality (HVHQ; near continuous) water quality data acquired during field monitoring campaigns between 1998 and 2008 with the Environmental Process Control Laboratory (EPCL) of the University of Georgia.

INDEX WORDS: Environmental Information Systems, Web-based Systems, Watershed Management, Watershed Information Systems, Scientific Visualization, Internet Technologies, Web Services, Georgia

INTEGRATED WEB-BASED DATA MANAGEMENT, ANALYSIS AND VISUALIZATION: TOWARDS AN ADAPTIVE WATERSHED INFORMATION SYSTEM

by

IBRAHIM DEMIR

B.S., Bogazici (Bosphorus) University, TURKEY, 2000

M.S., Gebze Institute of Technology, TURKEY, 2004

A Dissertation Submitted to the Graduate Faculty of The University of Georgia in Partial Fulfillment of the Requirements for the Degree

DOCTOR OF PHILOSOPHY

ATHENS, GEORGIA

2010

INTEGRATED WEB-BASED DATA MANAGEMENT, ANALYSIS AND VISUALIZATION:

TOWARDS AN ADAPTIVE WATERSHED INFORMATION SYSTEM

by

IBRAHIM DEMIR

Major Professor: M. Bruce Beck

Committee: C. Rhett Jackson
Todd C. Rasmussen
I. Budak Arpinar

Electronic Version Approved:

Maureen Grasso
Dean of Graduate School
The University of Georgia
May 2010

DEDICATION

I dedicate this dissertation to my family. A special feeling of gratitude to my wife who has been proud and supportive of my work and has shared the many uncertainties, challenges and sacrifices for completing this dissertation, and my daughter who has grown into a wonderful girl in spite of her father spending so much time away from her working on this dissertation. I also dedicate this dissertation to my loving parents whose words of encouragement supported me throughout the process.

ACKNOWLEDGEMENTS

I would like to express the deepest appreciation to my major advisor, Professor M. Bruce Beck, who continually and convincingly conveyed a spirit of adventure in regard to research and scholarship, and an excitement in regard to teaching. Without his guidance and persistent help this dissertation would not have been possible. I would like to thank my committee members, Professors C. Rhett Jackson, Todd C. Rasmussen and I. Budak Arpinar, for their valuable discussions and suggestions for my study and research.

In addition, I offer a thank you to my friends and colleagues at the Environmental Informatics and Control Program of UGA, Mr. Rodrigo Villarroel Walker, Dr. Shi Feng, and Dr. Feng Jiang. I would like to acknowledge Drs. Zhulu Lin, Wei Zeng, Femi Osidele and Amanda K. Parker for their valuable help and suggestions. I also appreciate our program coordinator, Jenny Yearwood, who has helped me so much on my school and life matters.

Finally, I would like to thank my wife, daughter and parents for supporting me and encouraging me with their best wishes. They were always there cheering me up and stood by me through the good times and bad.

TABLE OF CONTENTS

LIST OF TABLES

Page

LIST OF FIGURES

Page

CHAPTER 1

INTRODUCTION

1.1. Background

1.1.1. Environmental Information Systems

Computer systems have been in use for more than three decades for managing environmental information. A wide range of applications in environmental science is covered by these systems; including data storage, management, preprocessing, modeling, analysis, mapping and sharing of environmental information to help solve watershed management problems. The management of environmental information requires a broader approach to information technology, such as an information framework to provide interoperability, integration and sharing of environmental data, models and digital resources. Interoperability of information systems refers to the degree of ability to share information between distributed systems. Interoperability is a critical issue facing researchers who need to access information from multiple sources in environmental science.

Information Systems (IS) can be subdivided into three categories (Pillmann et al., 2006) as; (a) systems for ongoing measured data, (b) systems for administrative data available in documents, databases, structured data and maps, and (c) systems for enhancing environmental awareness in the public. Large-scale information systems are generally established onsite or on local infrastructure for offline use with features of systems (a) and (b). Offline infrastructures are preferred in order to keep data and information private, and because of security issues, and the cost of public access in terms of data storage, bandwidth and computing power.

There are many approaches for designing Environmental Information Systems (EIS), such as; geographical analysis, modeling with artificial intelligence (Expert Systems), statistics and many more (Denzer, 2005). Jeyaraj et al. (2006) discussed a methodology to identify predictors, linkages and biases in information systems design. Díez and McIntosh (2009) followed this methodology to evaluate information system design processes and suggested increased user participation during environmental IS development.

As management needs and the potential costs of EIS development increase, at a time of growing budget restrictions, developing economic solutions to complex environmental problems and the maintenance of EIS becomes critical (Pillmann et al. 2006). The importance of interoperability is growing for the management of spatially related environmental data and information. Effective interoperability provides an environment in which information from heterogeneous sources may be easily combined for comprehensive exploration, analysis and modeling (Frehner and Brandli, 2006).

There are two main approaches to Information System design (Rajabifard et al, 2003). The first and traditional approach is data-oriented, and focused on preparing and providing harmonized data. Today there is a growing interest in a user and application-oriented approach. This evolution has been accelerated by new Internet technologies, such as web services. There is a general decline in the use of the earlier systems (Crompvoets et al, 2004). Recent developments in web-based mapping and analysis tools show a promising potential and better user satisfaction for the later systems (Frehner and Brandli, 2006). Data integration tools are also implemented to a certain degree in the spatial information systems (Groot and McLaughlin, 2000; Williamson et al., 2003).

Web-based mapping and Geographical Information Systems (GIS; Peng and Tsou, 2003) provide additional tools for spatial data analysis. Web-GIS with analytical capabilities, and spatial analysis methods, have been demonstrated in various studies (Tsou, 2004, Anderson and Moreno-Sanchez, 2002). A web-based GIS, called River IS, has been developed to provide data management for the Croatian part of the Danube River (Pecar-Ilic and Ruzic, 2006). A study that links ArcIMS, a web-based GIS software, to a national and regional-scale river model is presented to create a WWW-GIS-based river simulation model called GIS-ROUT (Wang et al, 2005). Although there are examples for advanced Web-GIS, none of them makes use of accessing distributed data repositories for analytical purposes and most of them come with limited methods of web systems (Yang et al., 2005).

The trend of integrating distributed data sources clearly points towards the use of Web services (Frehner and Brandli, 2006). The concept and function of a web-based information system is described in the context of supporting decision-making within the UK planning framework (Culshaw et al, 2006). Kingston et al. (2000) discuss traditional and web-based systems for GIS and public participation and argue that new web-based technologies have the potential to widen participation in planning systems. Dupmeier and Geiger (2006) present concepts for the structure, content, and implementation of an EIS (Theme Park Environment) for the general public. They also point out the potential of web-based environmental information systems with web-services. Comprehensive environmental information can be disseminated rapidly over the web easily and in real-time (Zhu and Dale, 2000; Schimak, 2003; Sharma et al., 2003). For example, a web-based dynamic ambient air quality data management system has been presented (Triantafyllou et al, 2006) and with a conclusion that the system supplies the local community with a better awareness of air quality.

In addition, web-based information systems have advantages of utilizing available data and information by all interested parties, collaboration and participation opportunities in using models and data resources, and using the community for gathering metadata and digital resources. Challenges such as computing power, data storage and network bandwidth can be minimized through incorporating latest web technologies (e.g. AJAX, Web Services). Web services allow different applications from various sources to communicate with each other to exchange any type of information (text, data, maps, figures, and etc…) and distribute the computing load amongst multiple systems.

Developments in internet technologies, such as web services, provide interfaces for easy integration of data, digital resources, and tools from various sources. Improvements in client and server-side scripting languages (e.g. AJAX, Javascript) make available interactive content (e.g. visualization tools, online mapping), user interaction, rich user interfaces and desktop-like applications for data analysis and visualization possible on web browsers.

1.1.2. Watershed Information Systems

In principle, water quality can be managed through an integrative perspective on watersheds, facilitated by a web-based information system, and that will promote the most beneficial solutions to water quality problems, despite the geographic constraints of the collaborators and stakeholders. A fully functioning IS on this scale is a critical missing component for effective watershed research in Georgia watersheds and in many watersheds nationwide.

Information Systems (IS) have become an important issue for watershed research. Large amounts of data have been collected, and several models and tools have been developed to visualize and analyze the data. Much of the data and models, collected by individual researchers or research groups, are not integrated, since the data are stored in local databases and models are not easily accessible or applicable to other environmental systems. It takes an enormous amount of work to transform and understand the data and models so that they can be explored and utilized within other information systems.

Web-based systems have not been explicitly incorporated into watershed research. An information system with powerful tools for data/information management, mapping and collaboration is needed to be successful in watershed water quality management. The critical role of Information Systems in enabling all the wishes/ambitions for the future of watershed management is to provide an online system, so that information can be managed easily between participants and stakeholders can easily participate in gathering data and digital resources, analyzing data, running models, and accessing digital resources of a watershed.

Some of the early examples of watershed information systems have been presented by Thompson (1980), Band (1989), Lam et al. (1994), and Busse (1998). Alves and Loucks (2007) described the development, implementation and evaluation of a WIS for water quality control in river basins. The WIS presented in these studies feature data management and modeling tools and are implemented on local infrastructures.

Currently the web-based Arkansas WIS (2008) provides a statewide repository, including maps and practical reports for watershed units across the state. HydroSeek (Piasecki, 2008) features an integrated search capability across multiple data sources. The CUAHSI Hydrologic Information System (2008) provides web services, tools, standards and procedures to access data for hydrologic analysis. The Watershed Information System (WISE, 2008) developed by HSMM, Inc., and WebH2O (2008) developed by WebH2O, Inc, are commercial systems (the former being a desktop application). The WIS for the Bear River Basin (Horsburgh, 2004) provides a central repository for data and information.

To summarize, web technologies are starting to replace desktop applications and opening up possibilities such as online modeling, mapping, visualization of data and modeling results, and interfaces for stakeholder involvement, including members of the general public.

1.2. Objectives of Research

The primary purpose of GWIS is to record and present the work of the Environmental Process Control Laboratory (EPCL) of the University of Georgia, whose purpose in turn was to develop the monitoring and modeling of the water environment required to move towards a "smart" urban water infrastructure. As a prototype, GWIS has been developed to demonstrate capabilities and boundaries of an online information system from basic to advanced tools for the research community. The system is developed in a way to make adaptation of its components and structure easier in the light of new technology. Once it is populated with the necessary data and tools, GWIS can serve a management-oriented function.

The second objective is to develop a prototype of a comprehensive IS with data management, visualization, analysis, modeling, mapping and digital resources management capabilities. A long-term objective will be to provide a platform to help integration of state-wide efforts for an effective watershed management in terms of environmental information collection, collation, storage, analysis, retrieval, and dissemination to all potential stakeholders. Current information systems either focus on only one component of IS such as data management or GIS mapping, and lack features for interactive analysis and visualization of data.

The immediate contributions of the dissertation research will be as follows: (a) provide an Information System for Georgia Watersheds, (b) easy access and management of user provided data and information for observation sites on Georgia watersheds, (c) access ready-to-run and user-friendly visualization, analysis and modeling tools for data from users or various sources, (d) integration with USGS and EPA data and information sources, (e) provide interactive modeling tools powered with Expert Systems applications like 'Artificial Neural Networks' and 'Self Organizing Maps' for statistical and environmental analysis, (f) provide static and interactive visualization tools for comparison and model preparation, and (g) help in building up an inventory of watershed related digital resources.

This study is able to contribute in the long term as follows: (a) build up repositories and dissemination of information across diverse and dispersed sources for watersheds of Georgia, (b) adopt modern information technologies for watershed research and management, (c) support research and development on Georgia watersheds, and (d) promote state-wide collaboration and liaison for exchange of watershed related data and information.

This study describes a prototype of a web-based Information System for Georgia watersheds called 'GWIS: Georgia Watershed Information System'. Several data management, modeling, visualization, mapping and resource management tools for watershed research are included in the system. As its initial point of departure — to provide substantial and specific content — GWIS has been populated with the high-volume high-quality (HVHQ; near continuous) water quality data acquired during field monitoring campaigns from 1998 to 2008 with the Environmental Process Control Laboratory (EPCL) of the University of Georgia. It was a challenging, yet a required, test for an information system to have to deal with high-volume and high-quality data sets from the EPCL, since such data sets are not available for most information systems.

The Georgia Watershed Information System will have the following features in the general terms of Information Technology: (a) user-friendly user interface (UI) requiring minimal training, (b) fully customizable UI to implement the system for watersheds of other states or countries, or for different environmental research areas, (c) using open source and free web technologies to minimize implementation costs, (d) using latest web technologies, such as 'Web Services' for mapping applications and data sources; to distribute the load of network usage, data storage and computing power to different sources; such as USGS, Google, Yahoo and Microsoft, and (e) platform-independent access through any web browser on any device (personal computers and portable devices).

GWIS, with its advanced and novel features will provide distinct advantages when compared to traditional approaches to the following steps in environmental research: (a) information management, (b) assembling data, (c) analysis of data, (d) collecting digital resources, (e) modeling, (f) scientific visualization, and (g) accessing spatial data and maps. One of the most distinctive contributions of GWIS to environmental research is combining data collection (download and transfer), pre-processing (change format and structure), and examination (visualization and analysis) in one user-friendly system, which otherwise requires various information system tools, expertise, and time. GWIS can make watershed research and management easy, user-friendly, and more efficient with its novel and distinctive features powered by the latest web technologies.

1.2.1. Delimitations of GWIS

Since GWIS will be a prototype of a Watershed Information System, its features and a number of tools provided in its components (data management, modeling, visualization, mapping and digital resources) are limited. While providing high quality data, integration with different sources, and Expert Systems models, GWIS will provide only basic functionality on pre-processing raw data and mapping. The purpose of this delimitation is to keep the focus of the dissertation on building a comprehensive information system, rather than on the details of a specific component. GWIS does not provide complex environmental models, which require more computing power, data storage, training and expertise on the domain, so that the system will be available to wide range of users.

1.3. Dissertation Structure

This dissertation has 7 chapters. This first chapter has introduced Environmental Information Systems with a focus on watershed research, the objectives and goals of the research, and the dissertation structure.

Chapter 2 reviews the literature on Environmental Information Systems, web-based systems with main information system components, policy needs for an EIS in the light of the European Union Water Framework Programme and the US TMDL programs, and concludes with a discussion of Watershed Information Systems.

Chapter 3 describes how the Georgia Watershed Information System has been developed with special reference to six main system components and procedures incorporating distributed data sources, web services, mapping systems, and an integrated modeling environment. User interfaces, novel visualization tools and meta-data generation features are detailed extensively.

Chapter 4 basically illustrates through "mini-examples" how to use GWIS to the fullest extent for each tool and feature in six main components. User and content management, advanced tools and design principles of GWIS are discussed accordingly.

Chapter 5 introduces cases studies as demonstrations of how research and management can be carried out in a way that is much better using GWIS, as an Information System in general, and specifically for environmental research and watershed management, with a step-by-step comparison of a watershed study with and without the presence of GWIS.

Chapter 6 defines procedures for allowing GWIS to be systematically adapted (over the years) in order to accommodate and benefit from the furious pace of change in the technologies supporting and defining a GWIS for the following aspects: presentation, scalability, reusability and maintainability. Special emphasis is given to Web Services (WS) and Semantic Web (SW) technologies.

Chapter 7 presents conclusions of the research and recommendations on future research in the area.

CHAPTER 2

LITERATURE REVIEW

2.1. Introduction

In the 21st century, we face a set of important environmental problems, but we also have many opportunities provided by scientific and technological advancements. These challenges require all of our skills and imagination to respond in an integrated way to maintain the sustainability of environmental systems. Sustainable systems refer to the technologies that may help to decrease the human footprint on the environment while providing a desired living standard.

Sustainability is a complex field that cannot be effectively understood within the confines of a single discipline. On the other hand interdisciplinary communication and collaboration give rise to several operational difficulties. It is clear we need a platform for interdisciplinary collaboration on an increasing scale, widening our understanding of the physical world, aided by an increasing array of tools for exploration of that world. This platform should support flexible and dynamic integration of independent computational tools from a wide range of fields; such as data management and integration, scientific visualization, analysis, environmental modeling, online mapping, digital resources management and information delivery. Many of these tools can be considered under the Environmental Information Systems (EIS), and Information Technology (IT) in general since, by definition, IT is concerned with the collection, recording and sharing of information.

The literature review here starts with the key issues of environmental information systems and the main EIS components are discussed with respect to their challenges and trends. The review continues with the policy need for environmental information systems with respect to two comprehensive environmental programs: the European Union (EU) Water Framework Directive (WFD) and the US Total Maximum Daily Load program (TMDL). A special emphasis is placed on web-based information systems in environmental studies and watershed information systems.

2.2. Environmental Information Systems

Several definitions can be found in the literature for Environmental Informatics (EI). In general terms it can be considered to be the science of turning environmental data into information and understanding (Lancaster University, 2009). A more structural definition is given by the Natural Environment Research Council (NERC, 2008) as 'research and system development focusing on the environmental sciences relating to the creation, collection, storage, processing, modeling, interpretation, display and dissemination of data and information'.

Environmental informatics can be understood as an interdisciplinary field linking environmental information with social, economic, ecological, and environmental objectives. This new field brings various information technology tools and environmental monitoring networks together, in association with interdisciplinary modeling skills, to provide cost-effective solutions (Huang and Chang, 2003). This new field is centered on the development of technical and institutional standards and protocols for sharing and integrating environmental data and information. This requires interaction between fields such as environmental observation networks, environmental databases, environmental information systems, geographical information systems, internet technologies, scientific visualization, and public participation and understanding.

An Environmental Information System (EIS), on the other hand, is a computer system that integrates various technologies to improve the understanding of environmental information. Easy access and integration of environmental information is an important part of this process. The most essential part is providing environmental information in a comprehensive, accurate, cost-effective and timely manner.

The development of a comprehensive environmental information system should follow a systematic approach that consists of collection, storage, management, distribution, analysis of information, and decision support. The development process can be divided into four steps. The first step is data collection for the given environmental variables through various agencies. The core of the system should be developed to store and share data sources to ensure the flow and reliability of the data. The second level is management of the environmental information in a central database accessible by all stakeholders. The third step is the interaction and understanding of the information by using analysis and visualization tools. The fourth step is the decision support system (Al-Ghadban, 1997).

The structure, content and implementation of Environmental Information Systems has been discussed in many studies (Dupmeier and Geiger, 2006; Lorenz et al, 2006). The principle challenges in developing environmental information systems are the complexity of systems, interdisciplinary integration, volume of data (long-term data collection), modeling, security and privacy issues related with data sharing, and decision support systems that combine environmental and economic information with other data.

Various components in the design of environmental information systems are possible, for instance; artificial intelligence modeling (Expert Systems), geographical analysis, statistical analysis and many more (Denzer, 2005). As management needs and the potential costs of EIS development increase, at a time of growing budget restrictions, developing economic solutions to both complex environmental problems and the maintenance of EIS becomes critical (Pillmann et al. 2006). The importance of interoperability is growing for the management of spatially related environmental data and information. Effective interoperability provides an environment in which information from heterogeneous sources may be easily combined for comprehensive exploration, analysis and modeling (Frehner and Brandli, 2006).

2.3. Web-based Information Systems

Advanced data interoperability, integration and sharing services are required for the management and analysis of environmental data. Environmental information systems can be used for advanced web-based data and information retrieval, analysis, and visualization with the integration of distributed data sources (Frehner and Brandli, 2006). Web based systems can be used for managing and sharing environmental data to stakeholders with widespread use and flexibility (Smiatek, 2005). A web portal can function for web-based data input and upload, review of environmental data (Ahlers et al., 2005), management and editorial tasks, and data mining (Stadler et al., 2005).

Web based systems can be used for various purposes in the literature such as image-processing techniques for extracting solar radiation information from earth observation satellite images (Gschwind et al., 2006), and data mining analysis tools for researchers in ecology (Stadler et al., 2006). An advanced web-based infrastructure, Distributed Object-based Modeling Environment, DOME, has been presented (Kraines and Wallace, 2003) to allow domain experts to use the development tools and data to define interfaces for modeling, and provide software analysis tools to evaluate system behavior, decision trade-offs, or performance optimization.

A comprehensive web-based environmental information system can increase efficiency in environmental research, scientific productivity, educational effectiveness, accelerating discovery, foster new interdisciplinary science, and enhance the connection between science and society that could lead to significant social and economic benefits. The future environmental information system developments requires effective support of data acquisition, storage, exploration, visualization, spatial analyses, modeling, reporting, and decision support.

2.3.1. Web Services

Web Services enable a standardized communication framework (Figure 2.1) based on the internet to exchange information between organizations and information systems using open standards (Alonso et al., 2004), with an emphasis on openness, interoperability, extensibility, reusability, ability to compose, autonomy, and discoverability of these services. The term 'Open Standards' is defined by Free Software Foundation Europe as, "... a standard that allows sharing data freely and with perfect fidelity, preventing barriers to interoperability, and promotes choice between vendors and technology solutions".

Implementation of web services can make a better distribution and integration of tools and data, as presented in the Theme Park Environment IS (Dupmeier and Geiger, 2006). The potentials of web-based environmental information systems with web-services, mapping and data integration features for the public are presented in this system. Web services applied in many studies in the literature for facilitating the provision of input data to a particular multi-scale chemistry climate model (Smiatek, 2005), and a knowledge management system for waste management based (Eimer and Krause, 2005). The Grid on-demand project of the European Space Agency (ESA) shows how distributed data and computing resources can be managed using Web Services technology (van Bemmelen et al., 2005).

The need for distribution of spatial data clearly points towards the use of Web services (Frehner and Brandli, 2006). The concept and function of a web-based information system has been presented (Culshaw et al., 2006) to support decision making by making data and information easily accessible for environmental issues. Web-based technologies can improve public participation in planning systems (Kingston et al., 2000) compared to traditional systems. Implementing web services for distributed data repositories can greatly increase the interoperability of data and provide scientists a platform to easily utilize the data in scientific studies (Goodall et al., 2008), as shown in an online repository, National Water Information System (NWIS), for historical and real-time stream-flow, water-quality, and ground water level observations maintained by the United States Geological Survey (USGS).

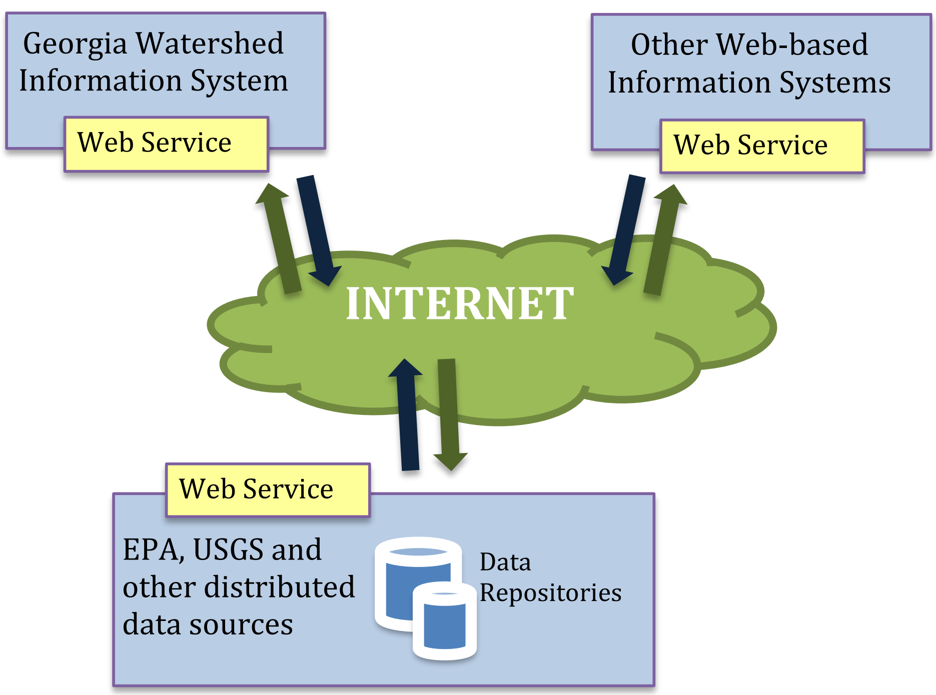


Figure 2.1. Web services and communication between distributed sources

Environmental information can be managed and shared with the public by a web based system easily and rapidly in real-time (Zhu and Dale, 2001; Schimak, 2003; Sharma et al., 2003). An example web-based dynamic ambient air quality data management system for data acquisition and monitoring showed (Triantafyllou et al., 2006) that the system provides a better understanding of air quality around their environment to the local community.

2.4. Information System Components

2.4.1. Design

Environmental Information Systems make information available to the users in many forms, from raw data to decision support systems (DSSs) with integrated data and dynamic models (Denzer et al., 1997; Power et al., 1995). The amount and quality of information available to users by EISs depend upon the level of pre-processing applied to the raw data during the collection and storage process, the relevance of the data to the needs of users, and the ability of users to use the EIS (Argent and Grayson, 2001).

A well-designed EIS can provide not only relevant information, but also an extra capacity to analyze the data (Argent and Grayson, 2001). In the process of EIS design, user interface and interaction with the information presented, data types, and formats should be considered carefully (Sprague and Carlson, 1982; Guariso and Werthner, 1989; Loucks and da Costa, 1991). There are many challenges for user interface design for environmental information systems, especially for geo-referenced data. A user interface combining both maps and catalog structures provides a better solution for integrated information systems with geo-referenced data (Guttler et al., 2001).

The user-interface of EIS should be designed in a way to facilitate the integration and presentation of information, and reflect the needs of users so that users can interact with the system and utilize it better to carry out analyses and visualizations (Figure 2.2).

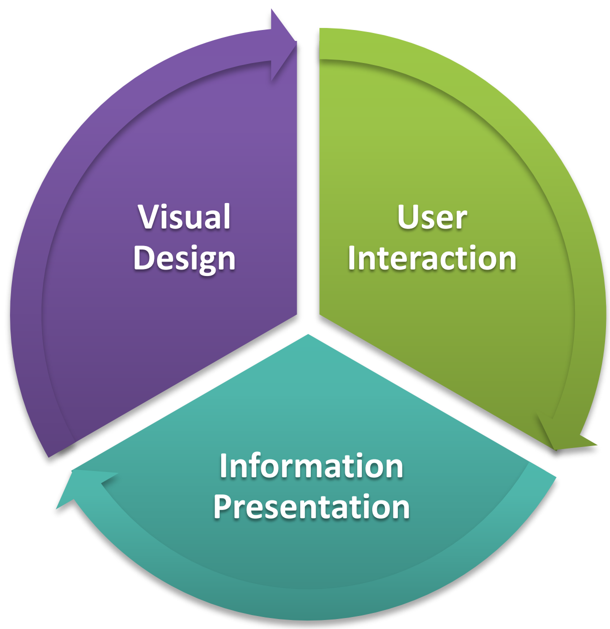


Figure 2.2. Three aspects of user-interface design for EIS

2.4.2. Data and Information Management

Data management in the environmental sciences includes data acquisition, quality control, quality assurance, data manipulation, analysis, interpretation, archiving and publication (Brunt, 2000). These tasks were often applied by individuals and small groups of researchers (Michener, 2000). Today environmental databases are mostly based on traditional file storage or commercial database management systems (DBMS). The needs for environmental data management have grown significantly with the high volume of data sets and their analysis in real-time applications. Both research and practical applications show that the traditional database architectures hardly satisfy current trends and new solutions are needed (Pokorny, 2006).

The volume of data sets containing environmental data has increased substantially in recent years, and they should not be stored within single data repositories any longer, but instead be held in distributed database systems. The number of studies on the integration of environmental information over a distributed network (Purvis et al., 2003), and environmental data management and monitoring system (Schimak, 2003) has increased over the last five years (2005-2010) with the developments in internet technologies.

Data from environmental observations provides a valuable resource that can be used to study environmental issues and build decision support systems. However, available data can be used effectively only if all the relevant details (meta-data) are also recorded in an accessible way, and this is best achieved using an information system. Details to be stored include the details of the campaign, physical properties of the area, monitoring methods, and all other details that might affect the variables and system. Researchers in many institutions are storing various details of their experiments into information systems on their computers.

Future researchers may wish to use this information to reanalyze the data using new statistical techniques. They may also want to produce analyses combining data from several different sites to study interactions. Data may be of different statistical and ecological qualities, and incomplete in these distributed sources (Stadler et al., 2004).

These issues should be handled during the data aggregation process for further analysis. The increasing complexity of data collection activities, using all kinds of environmental information in model development, and decreased funding for these activities, have led local community groups to cooperate with other agencies to gather such information. Information simply loses its value without easy, fast and affordable storage, distribution and analysis systems.

While local environmental information is important in order to understand a system in its local context, global information provides a larger perspective about the environmental systems. Local communities are storing their data on local infrastructures with a specific format and procedures. Many issues have emerged from these methods such as validity, accuracy, and precision of measurements, ownership, and copyright issues. Even if all these issues are handled carefully, further analysis raises more issues, such as formatting, storage, backup, distribution and use of information in other applications (Mayfield et al., 2001). There is great potential for information systems to help researchers in dealing with data complexities by providing computational tools to study and compare the data.

The primary problems of data management for users can be listed as follows (Tomasic and Simon, 1997): (a) data do not exist or are insufficient; (b) data are not referenced to the suppliers that makes it hard to locate, or referenced under domain-specific classification criteria; (c) data are hard to access because of being private, expensive, or costly to pre-process (e.g., paper documentation); (d) data sets are inconsistent or non-compatible; and (e) data quality is unknown due to lack of documentation.

One of the main issues of sharing data is to ensure an efficient data exchange platform between stakeholders. For that purpose, all related data would be stored in a distributed database and made available to all parties. Another important issue is the development of a common data format (e.g. XML, Extensible Markup Language) to provide consistency and improve data interchange. All available data sets can be converted into the same format using converters.

Extensible Markup Language (XML) is a well-known format for structured data on the web. It has been used in many studies to establish a link between data and applications. XML is a simple, very flexible text format derived from SGML (ISO 8879, 1986). XML plays an important role in exchanging a wide variety of data on the Web (W3C). XML makes it easy to manage data by providing a platform for sharing information, and supports the understanding of data. XML also facilitates proper documentation of great volumes of data. XML improves information sharing among all interested parties, including government, industry, community organizations, and public.

The management of spatially related environmental data increases the importance of data interoperability. Effective and efficient data interoperability is required to combine data from distributed sources to further explore and analyze (Frehner and Brandli, 2006). Environmental decision support systems (Guariso and Werthner, 1989; Poch et al., 2004) and participatory environmental decision making (Mustajoki et al., 2004) also require a database which dynamically interoperates (Bianconi et al., 2004).

Data integration and sharing tools have already been implemented to a certain degree in geographic data systems (SDI, Groot and McLaughlin, 2000; Williamson et al., 2003). There are two generations of these systems (Rajabifard et al., 2003): the data-oriented traditional approach; and the user- and application-oriented approach. New approaches are promoted by new internet technologies such as Web Services (WS). A general decline is observed in the use and management of the traditional approach, according to a comprehensive web survey (Crompvoets et al., 2004). Recent developments in Web-based mapping and analysis tools show a promising potential for the user- and application-oriented approach, and better user satisfaction (Frehner and Brandli, 2006).

The issues of environmental information collection and access deserve more attention by researchers, especially where the data are held by non-governmental entities and environmental groups (Haklay, 2003). The traditional approach to data management must be revised as new research challenges emerge along with technological developments and new user requirements (Frehner and Brandli, 2006).

2.4.3. Scientific Visualization

The National Science Foundation (NSF) initiated the use of scientific visualization in the 1980s to support researchers in visualizing their simulation results (Wang, 2005). However, this has not had an effect on many areas of environmental science. The importance of scientific visualization for environmental research comes from its potential for improving the quality of management and decision-making. Scientific visualization can help to present large volumes of data, examine trends and problems with data, and improve understanding of data (Ware, 2000).

In environmental management, it is essential to support interaction among stakeholders, who have different levels of involvement, from collecting and analyzing data to making decisions. Shiffer (2001) discussed several projects that support collaborative planning with an information system. Most of the presentations to stakeholders are made using two-dimensional (2D) maps and static charts and images. Complex environmental data have to be visualized with powerful scientific visualization approaches for presenting large amounts of diverse technical data to both technical and non-technical audiences.

Environmental research and management can benefit extensively from the integration of mapping, analysis and scientific visualization. Proper use of scientific visualization can enhance a researcher's ability to explore and analyze data, and interaction with stakeholders. Scientific visualization can help to increase the use and understandability of analysis and modeling tools through a user friendly interface. Online mapping can help the integration of images and animations on an interactive spatial map, and analysis tools can provide data for visualization of decision scenarios. All these benefits can lead to an increased understanding of the outcomes of environmental research and different management scenarios. In addition, integration of visualization, mapping and analysis tools can improve environmental research, management and decision making through an interactive system for evaluating different future scenarios. Scientific visualization supported with mapping and internet technology can bring better and faster public access to information (Ward, 2007).

2.4.4. Analysis and Modeling

Traditional environmental models are useful tools for understanding environmental processes. The key challenge is how to make use of models with different features, scales, and complexity as well as data collected from distributed data sources for a full understanding of the environmental mechanisms. Recent advances in high performance computing have shown significant potential to improve modeling and prediction accuracy in studies of environmental systems.

Grid computing, as one of the most promising systems nowadays, is associated with computer clustering. A grid is a collection of distributed computing resources such as storage and processors, available over a network, that appear to an end user as one large computing system with high computational capacity (Zhang, 2003). The purpose of building a grid system is to create virtual and interactive platforms to share computing resources among individuals and institutions in a secure way (Le Vine et al., 2003). The users do not need to know about the actual type, state and features of these distributed resources (Kacsuk et al., 2002).

Development of web-based analysis tools are based on four types of functionality (Foote and Kirvan, 2000; Peng and Tsou, 2003; Plewe, 1997; and Tsou, 2004a), as illustrated in Figure 2.3. These four types are the client-side user interface, web server, application layer and mapping layer. Recent studies on web-based mapping projects (Ghaemi et al., 2006; Swift et al., 2004; and Zimmermann et al., 2003) indicate the possibility of providing analysis and visualization tools for geo-referenced data.

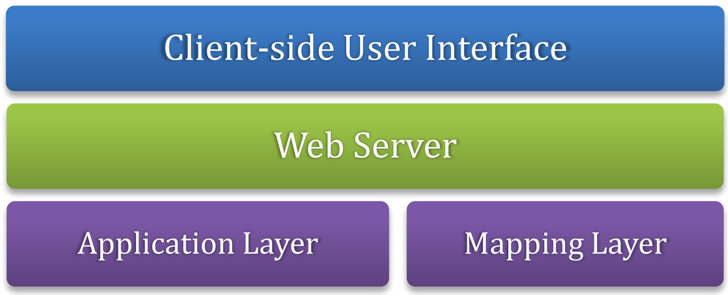


Figure 2.3. General architecture of web-based tools

The implementation of web-based analysis and modeling tools can provide many advantages. The internet extends the availability of analysis and modeling capabilities to the community and other stakeholder groups. Web-based tools provide an interactive service to non-expert users, especially those with otherwise limited resources, and professionals, which can be accessed concurrently from remote locations as a distributed resource.

2.4.5. Online Mapping

Since environmental events occur in time and space, data used in information systems are geo-referenced. Geographic Information Systems (GIS) are computerized mapping systems to capture, store, manage, analyze, and visualize spatial and descriptive data (Aronoff, 1989; Burrough and McDonnell, 1998). Due to advances in Internet technologies and increasing use of the Internet, local and closed structures of GIS are gradually changing towards distributed systems (Tsou and Buttenfield, 2002; Preston et al., 2003). Distribution includes both the storage of data and geo-processing capabilities (Peng and Tsou, 2003).

GIS is useful in handling complex spatial information which is essential for many environmental studies, as well as providing platforms for integrating various environmental models, systems, and interfaces (Lovejoy et al., 1997; Huang et al., 1999). In the past decade, many environmental modeling and decision-support systems based on GIS have been developed (Mailhot et al., 1997; Huang et al., 1999). The next step in distributed GIS is the web-based GIS platforms, which offers a system to discover and access geographic web services (Maguire and Longley, 2005). However, available functionality is limited to searching, mapping, publishing and limited querying of geo-data (Tait, 2005). Missing features in these systems are integrated analysis tools (Frehner and Brandli, 2006).

Web-based GIS (Peng and Tsou, 2003) provides additional tools for spatial data analysis. Web-GIS with analytical capabilities and spatial analysis methods have been demonstrated in many studies (Tsou, 2004b; Anderson and Moreno-Sanchez, 2002). Other applications of web-based GIS include waterway data management (Pecar-Ilic and Ruzic, 2006) and river simulation models (Wang et al., 2005). Although there are examples of advanced web-based GIS, most of them have limited technical capabilities and few of them make use of accessing distributed data sources for analytical purposes (Yang et al., 2005).

2.5. Policy Needs

2.5.1. European Union - Water Framework Directive

In this section, literature is reviewed on information systems from the perspective of the EU - Water Framework Directive (WFD). The Water Framework Directive (WFD) is recognized as one of the most comprehensive pieces of European environmental legislation to date (European Commission, 2000). It aims to establish a framework for the European Community in the field of water policy (Usländer, 2005). Objectives related with Information Technology (IT) mentioned in the Directive 2003/4/EC (article 1) are (i) the right of access to environmental information, and (ii) to ensure that environmental information is progressively made available and disseminated to the public.

Usländer (2005) has given a brief analysis of the existing IT recommendations for the WFD implementation strategy and pointed to the need to develop a common IT Framework Architecture integrating different views (organizational, process, data and functional) in one single concept. He concludes that IT developments are required to efficiently support the WFD implementation with data acquisition, storage, exploration, visualization, decision support, spatial analyses, reporting, analysis and modeling features.

A comprehensive environmental information system has been developed by different EU organizations and member states for water related issues, called the Water Information System for Europe (WISE, 2008). The purpose of WISE is to provide a platform for the reporting process, so that more information can be gathered with an efficient data exchange process (European Commission, 2005b). These data exchange processes are supported by projects like Orchestra and INSPIRE by promoting new European standards (Klingseisen et al. 2006).

In 1998, the European Union started the concept of River Information Services (RIS) to describe information services for terrestrial navigation. Several research and development projects, focused on RIS, have worked on the technological developments, system integrations, standardization, international cooperation and networking (European Commission, 2005a). For example, an FP5 Program, the Pan-European project Consortium Operational Management Platform River Information Services (COMPRIS, 2005), was started in 2002. The INTERREG project data warehouse for the Danube waterway (D4D) is another example of the exchange of waterway related data and information between the stakeholders (Via donau, 2002).

A series of workshops and joint activities has been launched to support the development and testing of WFD methodologies. As a part of the International Symposium on Environmental Software Systems (Usländer, 2003), a workshop with participants from environmental organizations, scientists, and information technology consultants took place (ISESS, 2004) and the following results have been agreed between the participants, that there is an urgent need for: (i) a better information exchange between water and IT sector at all levels; (ii) defining IT perspective for the implementation of the WFD; (iii) an information system with organizational, process, data and functional view perspectives; and (iv) validation of the current and future WFD-IT specifications in selected river basins.

On the one hand, WFD requires a comprehensive analysis of all impacts in a river basin, and on the other hand, the available data sets are not adequate for a complete analysis to cover all impacts from various sources. Even if the data will be available, they are not comparable due to the scale and time inconsistencies throughout the river basin. These issues are often common in international river basins (Dublin, 2004). The Directive introduces the intention to share spatial data by using Geographical Information Systems (GIS) between stakeholders.

The role of stakeholder participation and decentralization of responsibilities from authorities to community groups is increasing in the context of the WFD (Blomqvist, 2004). Participation may contribute to better understanding of the environmental issues, public awareness, knowledge and information transfer between local, regional and national groups, and reduction of conflicts between stakeholders (Lauber and Knuth, 2000; Pahl-Wostl, 2002). In the Final Summary of Proceedings of Pan-CEE NGO Conference on Public Participation and the Water Framework Directive (2001), it was agreed that public participation is required for sustainable environmental management and successful implementation of the Water Framework Directive.

Several studies highlighted the role of public participation (GWP, 2000; Belfiore, 2002; Olsson and Folke, 2001) and data and information transfer, with transparent and publicly accessible information systems (Langaas et al 2004, Denisov and Christoffersen, 2001, Timmerman et al 2002), in the context of sustainable integrated water management. In sum, the EU projects, WFD requires and suggests a broader approach to Information

Technology, such as a shared information framework (data, analysis tools, reports, and etc.) to secure interoperability, integration, standardization, international cooperation, public participation, and harmonization of a large amount of data/information in a trans-boundary, cross-disciplinary environment.

2.5.2. US - Total Maximum Daily Load Program

In this section, literature is reviewed on information systems in US - Total Maximum Daily Load (TMDL) program perspective. The TMDL is defined by the U.S. Environmental Protection Agency (EPA, 2007) as the calculated maximum amount of a pollutant that a water body can receive and still meet applicable state water quality standards, and an allocation of that amount to the pollutant's sources. The TMDL Program, mandated by the US Federal Clean Water Act, integrates watershed planning with water quality assessment and protection. TMDL differs from other pollution management efforts by requiring that loads from all pollution sources within an impaired watershed be allocated, where other efforts focus on loads from a few, identifiable sources. TMDLs generally require that a number of programs and groups work together to achieve the desired level of pollution control, while other efforts are often limited to a single program or entity (EPA, 2007).

Since the TMDL process depends on models to evaluate alternative scenarios for remediation, it is important to have adequate data for model calibration to achieve confidence in appropriate and efficient remedial plans (Richards, 2004). The development and implementation of many TMDLs is obstructed by limited data. This results in often ineffective and overly restricted TMDLs, and a lack progress on water quality improvement. Developing TMDLs in a watershed begins with the collection of large amounts of data within that specific watershed (Tate et al., 1999, Spooner and Line, 1993). This results in the need for data management techniques and tools to manage high volumes of new water quality data. These data must be managed in an integrated system to store, pre-process, modify, analyze, and share within the community of researchers and other stakeholders (Carleton et al., 2005).

The challenge is to find an approach and have appropriate tools for limited data and resources situations (Freedman et al. 2004). The US EPA, with the support of the West Virginia Department of Environmental Protection and Tetra Tech, Inc., has developed a simulation system (Mining Data Analysis System) to support TMDL development in areas affected by mining operations. The system can simulate large watersheds for assessment and allocation of point and nonpoint sources (Henry, et al., 2002). The US EPA now promotes a watershed-based approach that considers both point-source and nonpoint-source pollution. It requires stakeholder involvement in the development and implementation of the TMDL process. Stream flow and water quality data can be collected across the nation by researchers, state and federal agencies, and local community and organizations. Stakeholders should be involved in the early stages of the TMDL, and in model development in order to improve confidence in the model (Maguire 2003a).

Stakeholder interactions with water quality models and modelers in a TMDL process (Maguire, 2003b) are also important. Some aspects of this interaction have helped stakeholders to become involved in the model development and enhance their understanding of the topic. A decision support system has been developed (Chen et al. 2004) to help stakeholders through the calculation of the total pollutant amount that can be discharged against the limits of the water quality criteria for the designated use of water quality in the study area. A decision support system can be used also to guide the stakeholders through the development of a TMDL implementation plan. A web-based stakeholder process can provide a participation platform for stakeholders in management decisions by presenting scientific information to make their own decisions (Chen et al. 2004).

Having a large number of participants in an information system will increase the information and tools available to stakeholders; on the other hand it will increase the challenges in terms of management of such an information system. Two of the critical technical challenges of information systems for watershed management are information management and participation. Implementation of TMDLs requires collaborative efforts among federal, state and local governments, and public stakeholders within each watershed.

The TMDL requires an information systems approach that can handle challenges especially for data management and collaboration among various stakeholders. Web-based systems are best suited to TMDL for sites with the following considerations: (i) having limited data; (ii) with strong stakeholder cooperation; and (iii) where voluntary actions can be implemented.

2.6. Watershed Information Systems

There are a few early examples of Watershed Information Systems (WIS) for watershed management and planning (Thompson, 1980; Band, 1989; Lam et al., 1994; Busse, 1998). GIS can offer new ways to handle and use information about catchment characteristics with a high spatial resolution (Schumann et al., 2000). GIS and remote sensing can be used in the identification and categorization of watersheds on the basis of natural resources and their limitations (Khan et al., 2001). In recent years, GIS applications have been utilized for mapping in many watersheds studies, such as predictive modeling capabilities of a GIS-based model to assess the effects of implementing land use change within a watershed (Millward and Mersey, 2001), and integration of GIS and agricultural nonpoint-source pollution model to help planners and decision makers develop a watershed management plan (He, 2003).

Alves and Loucks (2007) described the development, implementation and evaluation of a WIS, a suitable computational tool, for water quality control and identifying source areas of pollution in river basins (Alves and Loucks, 2007). Most of these WIS presented in these studies feature data management and modeling tools, and are implemented on local infrastructures. A web-based information system can increase the capabilities for watershed management and planning with online mapping, data integration and visualization components.

The latest commercial and federal implementations of web-based WIS offer more features for watershed management. The Arkansas WIS (2008) provides a statewide repository including online maps and reports for watersheds across the state of Arkansas. HydroSeek (Piasecki, 2008) features an integrated search capability across distributed data sources through an interface by utilizing Web Services. The CUAHSI Hydrologic Information System (2008) provides web services, tools, standards and procedures to access data for hydrologic analysis. The Watershed Information System (WISE, 2008) has been developed by HSMM, Inc. as a commercial desktop application to manage, access, and analyze large amounts of water resources data. The Bear River WIS (Horsburgh, 2004) provides a central repository for data and information for the Bear River Basin, a website with informational content, data visualization and analysis capabilities, and a map server with GIS capabilities. WebH2O (2008) has been developed by WebH20, Inc. as a commercial web-based EIS that combines database features with GIS and technical analysis tools for environmental data.

The increasing number of commercial and state-wide WIS efforts with more features or specialized components shows that web technologies are starting to replace desktop applications. Promising advances in web technologies make online analysis, modeling, mapping and scientific visualization possible for high volume data and models. Combined with data integration and collaboration capabilities, web-based WIS can provide a flexible and comprehensive platform for watershed management.

2.6.1. End-users of a Watershed Information System

Decisions that affect our future environment should be based on an understanding of the needs, desires and preferences of all stakeholders affected by these decisions. Public participation is therefore an important aspect of sustainable environmental development.

However, public participation in the development of Information Systems (IS) is still a relatively new concept. A watershed information system can benefit from several potential end-users, such as national governments, environmental managers and management agencies, decision-making bodies, policy-makers, decision-makers, scientific research community, private sector, non-governmental organizations, and public users and local interest groups.

Potential end-users of a watershed information system can benefit in general from following the IS features summarized in Table 2.1 (Evans, 1999).

Table 2.1. Various features of Information Systems

Feature and description
Improved access to local or national environmental data and information Many environmental data and information are stored in local databases. An information system can improve access to these distributed sources with various methods.
Integrating distributed environmental repositories Distributed sources can be integrated in a central information system to minimize data management efforts.
Presenting environmental data in a user-friendly format Current access to environmental information is complicated. This is because of the difficulty of accessing the data, format problems for the accessible data, or not having the necessary software to visualize and understand the data.

Table 2.1 (continues)

Enabling the integration of various disciplines of environmental research Sharing together a variety of environmental information from a range of sources and presenting it through a web-based interface helps integration and encourages discussion of the environment as a whole.
Encouraging public participation regarding the requirements of the system Public users can be encouraged to visit the IS to share their views about the key issues and information that the IS should be able to address.
Developing web access and utilization of the EIS via a web interface The challenge of developing a fully functioning web-based IS has still not been met. It needs to be addressed to access, query and visualize the environmental data in an easy and accessible way.
Integrating other digital resources and multi-media into the IS Environment information is not only about raw observation data. The integration of other digital environmental resources also needs to be addressed, such as reports, articles, meta-data, multimedia information, and other media files. This would allow information to be viewed and understood in context.
Allowing multiple stakeholder levels of access to the IS While most of the users have access to certain common data sets, some users will have access to restricted ones. Also, special mechanisms are required to protect data from user mistakes and misuse.

A large number of public and private organizations collect environmental information. Current data collection efforts are unsystematic and inadequate to monitor the quality of environment and reach goals of each end-user. Either data collection efforts are not well coordinated, or there are difficulties in these efforts. The difficulties to coordinate data collection efforts can be summarized as follows (GAO, 2004):

- ✓ End-users often collect data to achieve specific goals, which makes them unwilling to change their data collection efforts to make the results more widely usable or they are reluctant to share their data with other end-users.
- ✓ End-users often use different data collection methods, resulting in inconsistent standards (e.g. definitions, detection limits (MacDougall and Crummett, 1980), quality assurance, meta-data).
- ✓ Many end-users are unaware of who is collecting which types of data because they do not have a centralized information system on data.
- ✓ Data coordination efforts are often assigned a low priority.

These challenges are not only caused the gaps and the duplication of efforts, but also complicated the integration of data from different sources. Difficulties in data management are commonly focused on two areas: (1) complexity of using current Information Systems (e.g. EPA's STORET, 2008); and (2) challenges of integrating data from various sources to provide a better understanding of environmental systems. Difficulties in integration often are caused by using different formats in different databases.

It is recommended that data collection efforts can be effectively improved with special mechanisms through a Watershed Information System with sufficient authority and resources. The main responsibilities of the Information System would be as follows:

- ✓ Support the development and continuation of data collection efforts
- ✓ Coordinate which end-users are collecting what types of data
- ✓ Allow users easy access to environmental data and information
- ✓ Develop metadata standards to integrate data from various sources

In particular, the following steps can help data incompatibility challenges:

- ✓ Delivering data in a common format to users
- ✓ Ensuring that the data are documented to describe their quality for determining the utility and comparability of the data
- ✓ Including metadata associated with each data collection effort
- ✓ Encouraging the use of metadata to enhance the usefulness of the information
- ✓ Developing a geospatial query tool for sharing data

2.7. Conclusion

The development of new information technologies has experienced extraordinarily rapid progress, and some new techniques and tools have yet to be applied to environmental research. There is a significant opportunity for information systems to contribute to sustainability and thereby minimize the risk of environmental problems. Therefore, further development of information systems associated with advanced data management, analysis and visualization techniques is desired to promote the long-term success of environmental research. Many challenges exist in the applications of analysis and modeling techniques in environmental research. Most environmental models can only handle limited spatial and temporal data and components because of the difficulties in available computational tools and data. On the other hand, environmental data collection has many challenges due to wide range of environmental systems, data sources, and stakeholder involvement (Briggs, 1995).

Traditional database technologies are limited in dealing with spatial and temporal queries. Geographical Information Systems are now integrated into environmental information systems to handle spatial environmental data. Advanced database technologies are needed to store, manage, pre-process high volumes of data from different sources in several formats and types. Almost all participants in environmental research are using computerized systems for environmental information processing. Typically, these systems consist of multiple components serving various purposes. Environmental information systems should be developed to allow integration of existing subsystems and components. The integration of different components (e.g. database systems, mapping, analysis, or visualization systems) in the development of environmental information systems would not be a challenge to information technology, if these systems are supported with high-end web technologies.

Research and applications in literature relating to Environmental and Watershed Information Systems context is far from complete, and the authors reviewed often emphasize the need for greater attention to this area. Information Systems reviewed here are either implemented on a local infrastructure, focused on a specific problem of watershed management, or not utilizing the latest web technologies. More broadly, the EU WFD and the US TMDL programs give us an idea about the missing points, needs and trends in Information Systems for environmental studies.

No research was found that provides a comprehensive framework for Watershed Information System with high quality data management, interactive visualization, statistical analysis, expert systems modeling, online mapping, and digital resources components. As a conclusion from this review, the architecture of the Georgia Watershed Information System will follow the trend of the latest web technologies, and take advantage of open standards, open source systems and web-services. It will particularly extend current Watershed Information Systems by providing online analysis, visualization, modeling and mapping capabilities that can be applied to distributed data sources.

One of the original contributions of GWIS will be combining 3 main processes of environmental research: data collection (download and transfer); pre-processing (change format and structure); and examining (visualization and analysis) data in one user-friendly interface, which is a time- and resource-consuming process with various information system tools and expertise. The next chapter describes the development process of, and components of, the Georgia Watershed Information System, with special reference to procedures incorporating distributed data sources, web services, mapping systems, and an integrated modeling environment. Design methodologies, novel visualization interfaces and meta-data generation tools are detailed extensively.

CHAPTER 3

STRUCTURE AND COMPONENTS OF GWIS

3.1. Introduction

This chapter describes how the Georgia Watershed Information System has been developed with special reference to procedures incorporating distributed data sources, web services, mapping systems, and an integrated modeling and visualization environment. An adaptive user-interface design, novel visualization interfaces and meta-data generation tools are discussed in detail. The chapter also includes the structure of GWIS, details of GWIS components, a comparison of GWIS with contemporary watershed and environmental information systems, and an initial setup for GWIS.

3.2. Initial Setup of GWIS

To begin with — to provide substantial and specific content — GWIS has been populated with the high-volume high-quality (HVHQ; near continuous) water quality data acquired during field monitoring campaigns (Table 3.1) between 1998 and 2008 with the Environmental Process Control Laboratory (EPCL) of the University of Georgia. These campaigns have covered monitoring of nutrient (and biological) dynamics in environments such as wastewater treatment plants, aquaculture ponds, and rural and urban watersheds subject to nonpoint-source runoff. GWIS also features several modeling and visualization tools integrated with distributed sources from USGS and continually growing digital library and online mapping capabilities. GWIS has several tools, or features, for each component to provide examples from different levels (basic to advanced) of capabilities for a web-based information system. A summary of features and tools in GWIS is given Table 3.2.

Table 3.1. Summary of observation campaigns and site descriptions

Site Name and Additional Info	Campaign Start Date	Campaign End Date	Measured Variables
Athens WWTP No.2 - Aeration tank 2 - inner channel [1,2,3]	02/01/98	04/06/98	3
Athens WWTP No.2 - Aeration tank 2 - middle channel [1,2,3]	02/01/98	04/06/98	6
Athens WWTP No.2 - Aeration tank 2 - outer channel [1,2,3]	02/01/98	04/06/98	2
Athens WWTP No.2 - Crude sewage [1,2,3]	02/01/98	04/06/98	7
Athens WWTP No.2 - Secondary clarifier [1,2,3]	02/01/98	04/06/98	1
Athens WWTP No.2 - Secondary clarifier effluent [1,2,3]	02/01/98	04/06/98	4
Athens WWTP No.2 - Aeration tank 2 - inner channel effluent [1,2,3]	04/20/98	04/28/98	6
Athens WWTP No.2 - Returned activated sludge [1,2,3]	04/20/98	04/28/98	6
Athens WWTP No.2 - Secondary clarifier [1,2,3]	04/20/98	04/28/98	6
Athens WWTP No.2 – campaign #3 [1,2,3]	06/12/03	08/17/03	16
Barnett Shoals Dam [4]	07/15/98	08/04/98	7
Barnett Shoals Dam [4]	03/14/99	05/04/99	3
Soque River Site – Upstream [6] (*processed data for model*)	06/20/05	07/11/05	4
Soque River Site – Downstream [6] (*processed data for model*)	06/08/05	07/13/05	3
Soque River Site – Upstream [6]	06/08/05	07/13/05	10
Soque River Site – Downstream [6]	06/08/05	07/13/05	10
USGS Penfield Gauge - HUC: 02218300 [4]	07/15/98	08/04/98	7
USGS Penfield Gauge - HUC: 02218300 [4]	03/14/99	05/04/99	3
Weracoba Creek, Columbus - BMP, Influent [4, 7]	11/01/07	02/29/08	17
Weracoba Creek, Columbus - BMP, Effluent [4, 7]	11/01/07	02/29/08	17
Weracoba Creek, Columbus - Wynnton Rd [4, 7]	11/01/07	02/29/08	10
Whitehall Estate Aquaculture Pond #1 [3, 5]	08/02/98	11/18/98	8
Whitehall Estate Aquaculture Pond #2 [3, 5]	08/02/98	11/18/98	7
Whitehall Estate Aquaculture Pond #3 [3, 5]	05/23/00	10/16/00	18

[1] Liu, 2000; [2] Jiang, 2007; [3] Lin, 2003; [4] Zeng, 2000; [5] Parker, 2004; [6] Feng, 2008; [7] Boner et al., 2008.

Table 3.2. Summary of features and tools in GWIS

Component	Feature
Data Management	GWIS Data Sources
	GWIS Data Web Service
	USGS Real-time Data Integration
	USGS Daily Data Integration
	Data Sources by HUC (Hydrologic Unit Code)
Data Analysis	Descriptive Statistics
	Trend and Seasonal Component Analysis
	Autoregressive Polynomial Interpolation Analysis
	Correlation Analysis with Self Organizing Maps
	Missing Data Analysis with Self Organizing Maps
Modeling	Watershed Water Quality Simulator
Scientific Visualization	Spatial Visualization on Dynamic Maps
	Multi-dimensional Scatter and Histogram Chart
	Dynamic Charts
	Static Charts
Map Services	USGS Dynamic Information Interface
	Maps powered by Google Maps API
	3D Maps powered by Google Earth API
	KML on Google Maps API
	Maps powered by Virtual Earth API
	Maps powered by Yahoo Maps API
Digital Library	GWIS Digital Library
	GWIS Site Description
	GWIS Site Description Web Service
	Meta-data Generator
	GWIS Wiki Fork
	Site Description by USGS Site Number

3.3. GWIS Structure

GWIS is designed as a web-based information system, which runs on any operating system by using a web browser. It has been tested with major web browsers (Internet Explorer, Firefox, Opera, Safari, and Chrome) and operation systems (Windows XP and Vista, Mac OS X Leopard, and Ubuntu). Figure 3.1 presents an overview of the architecture of GWIS, where corresponding layers are positioned according to the components. The details of each component and elements in layers will be explained in the following sections of this chapter.

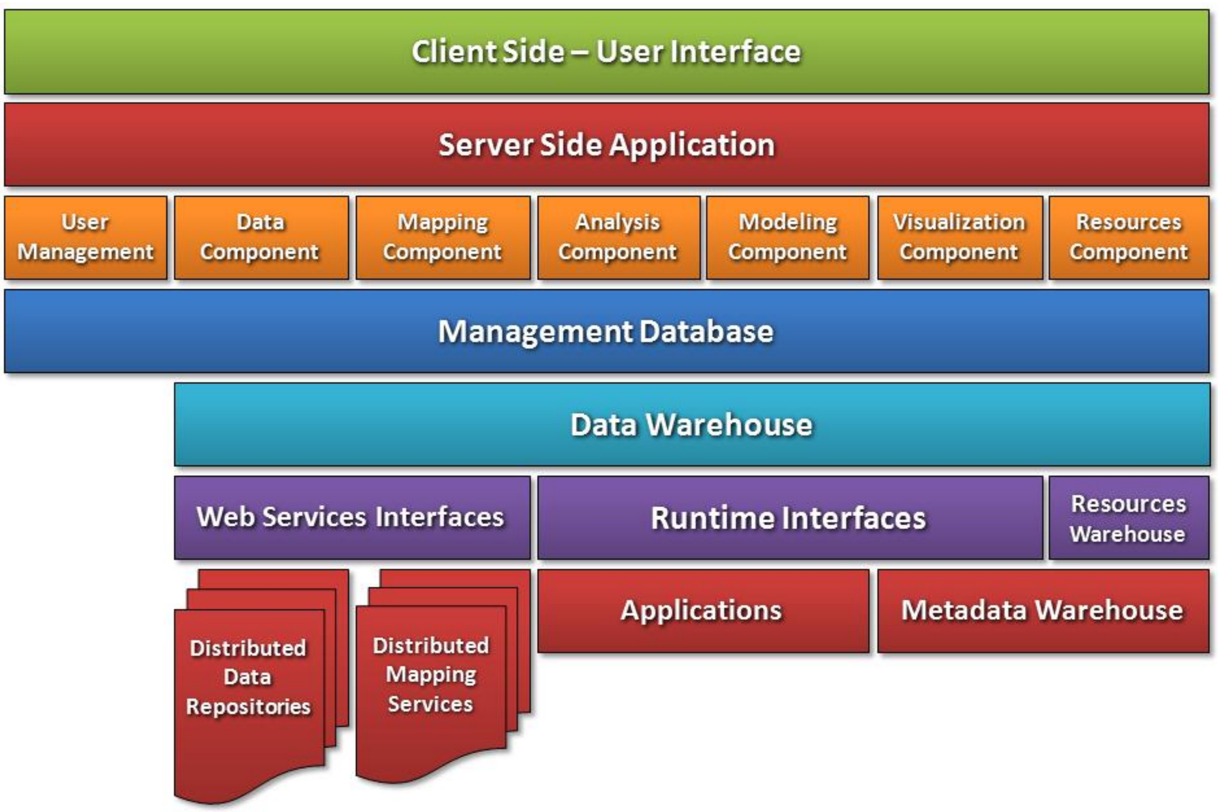


Figure 3.1. Architecture of GWIS

While providing high-volume high-quality water quality data and several sample tools for each component, the scalable and interoperable structure of GWIS also allows developers to easily integrate different data sources and models into the system. The client-side of GWIS is built with HTML, CSS, JavaScript and Macromedia Flash, and the server-side uses PHP and MySQL databases. Models and some of the visualization tools are stand-alone applications created by the MATLAB Compiler, which runs through a Common Gateway Interface (CGI) on the web server. Warehouses support several file formats like XML, CSV, PDF, DOC, etc.

Client-side User Interface: GWIS is designed to provide a better user experience with a simple and robust user interface, providing interactive interfaces with AJAX technology, clear user instructions, information boxes for additional information, auto complete and various controls for easy user input, which will also help consistency of the data throughout the system. The client-side user interface (Figure 3.2) provides a platform for interacting with data, models and analytical tools, visualization and mapping.

The user interface is designed with HTML, Javascript, and Cascading Style Sheets (CSS). Some of the features utilize Macromedia Flash plug-in. The centralized approach for design makes the system easily modifiable and adaptable for other platform requirements. It is enough to change only a couple of CSS files to modify the system for re-branding it in local or commercial use, or adapting it for portable or small screen platforms.

Server Side Application: GWIS depends on PHP scripting language and mySQL databases on the server side. Currently, GWIS runs on Microsoft Internet Information Services, which is freely available with most editions of the Microsoft Windows XP and Vista operating systems. However, the current design of GWIS allows it to run on many freely available web servers (e.g. Apache). Some of the tools in the analysis, modeling and visualization components require CGI (Common Gateway Interface) and MCR (MATLAB Compiler Runtime) facilitates, which require no cost for software purchase and licensing.

Figure 3.2. Home screen of GWIS website

User Management: GWIS allows users to register on the system and manage their data. Registration requires users to provide basic contact information. Registered users can add their projects, observation sites, variables, data and digital resources to the system. All of this information will be accessible anywhere, and shared with other GWIS users. Users can edit their profiles and information submitted by them. Some of the content of the system is modifiable by all users, to provide a collaboration platform.

Components of GWIS: GWIS has six main components for information management. The details of each component and related tools are provided in the next section of this chapter, while the next chapter of the dissertation provides mini-examples of these tools with real data, step by step instructions and screenshots. The basic description and functionality of the main components are given below. Each tool in each component of GWIS is detailed in the next section of this chapter. While GWIS provides easy-to-use analysis and modeling tools, and background information on its methods, it is strongly suggested that the user should consult the original source references and literature to ensure the accuracy, reliability and significance of his/her results.

a. ***Data Component:*** This component provides interfaces to access and transfer data from EPCL projects (24 observation sites), user-submitted data (e.g. high quality data of water quality parameters) and distributed data repositories (e.g. USGS real-time and daily data series). The main functionality of the component is managing data (share and export), web services capabilities, integration interfaces for distributed sources, and connections to site-related sources.

b. ***Visualization Component:*** This component provides static and interactive visualization tools for comparison and statistical evaluation (e.g. line plots, statistical plots, multi-dimensional scatter and histogram charts). It also has a novel interface to visualize time series data on maps, called an Interactive Scientific Visualization Interface.

c. ***Analysis Component:*** This component is an integrated analysis environment, providing descriptive statistics, interpolation analysis, trend and seasonal component analyses, and correlation and missing data analyses with Self Organizing Maps.

d. ***Modeling Component:*** This component demonstrates the incorporation of advanced hydrological models into GWIS with an example model, Watershed Water Quality Simulator.

e. ***Mapping Component:*** This component provides online mapping systems with rich geographical content (powered by Google, Yahoo and Microsoft Maps APIs), Keyhole Markup Language (KML) and overlay support on maps, and an interactive interface for USGS monitoring sites. KML is a file format used to display geographic data on 2D maps and 3D Earth browsers.

f. ***Digital Resources Component:*** This component provides the digital collection of resources such as articles, reports, modeling results, metadata, meeting minutes, audio, video and etc. related with selected observation site. There is also a Web Services support for sharing site description, detailed site description from GWIS and USGS, a meta-data generator and a Wikipedia retriever for watershed-related terms.

Management Database: The management database connects each component to related interfaces, warehouses and web services.

Data Warehouse: The data warehouse includes user submitted data for observation sites. Currently GWIS contains the EPCL time series for 24 observation sites. The data are stored on the server in CSV (Comma Separated Values) file format for each variable of an observation site, and registered on the mySQL server.

Web Services Interfaces: Web Services are utilized in data, mapping and digital resources components of GWIS. The data components use web service to share data with users and access distributed data sources such as those from the USGS. Mapping components use web services to provide a map-server functionality by accessing available Map Services from Google, Yahoo and Microsoft. The digital resources component uses web services to share site description information with users.

Runtime Interfaces: The analysis, modeling and visualization components access stand-alone executable files (compiled by MATLAB) on the server by using CGI, and present results in a report as HTML files with images. Executable files require MATLAB Compiler Runtime on the server side, but users do not need to install any plug-in or MATLAB to run the tools and access results.

Digital Resources Warehouse: The resources warehouse includes user-submitted digital content related with observations sites. Currently GWIS has digital resources from EPCL projects. Files are stored on the server in various formats for each file and registered on the mySQL server.

Metadata Warehouse: The metadata warehouse includes user-submitted metadata files related with observations sites. Currently GWIS has metadata files from EPCL projects. Metadata files are stored on the server in XML format for each file, and registered on the mySQL server.

3.4. GWIS Components

GWIS has six main components with 27 tools and features. Each component provides detailed background information with references about the methods used in the tools, and instructions to fully utilize the system and help users to understand how they will interpret their results.

3.4.1. Data Component

The data component includes several interfaces to access and utilize data located on the GWIS server and distributed data repositories. Users can locate monitoring sites on interactive satellite maps, export data for selected variables, and connect to sources related with the site from EPA and USGS. An interface is designed to transfer data and information from USGS sources to the GWIS server. Users can transfer daily or real-time data for selected parameters and available time ranges to export, visualize or run analyses. GWIS also provides a Data Web Service which makes user-uploaded GWIS data accessible through any platform or software that has web services support. The details and screenshots of the data tools are given below.

GWIS Data Sources Tool: Users can export variable data to their computers by selecting an observation site and a variable from the list provided in the interface (Figure 3.3). The data is provided as a time series in a CSV file format. The first column of the CSV file contains time values, and second column the variable value in its given units.

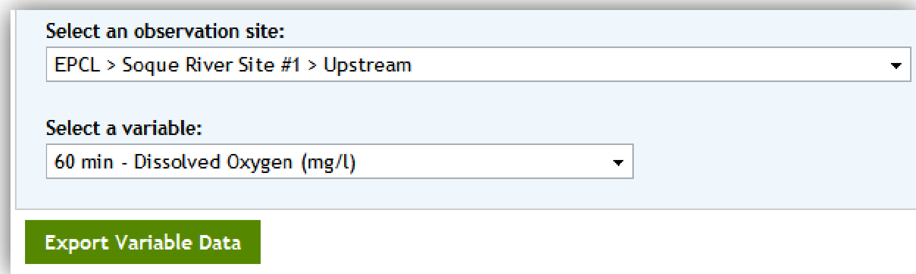


Figure 3.3. Interface of GWIS data sources tool

GWIS Data Web Service Tool: Users can access data uploaded by GWIS users from any tool or system that supports web service integration by using the GWIS Web Service tool. Web service is defined as 'a software system designed to support interoperable machine-to-machine interaction over a network' by the World Wide Web Consortium (W3C, 2009). GWIS uses a REST (Representational State Transfer) web service structure to share data. REST was introduced and defined by Fielding (2000).

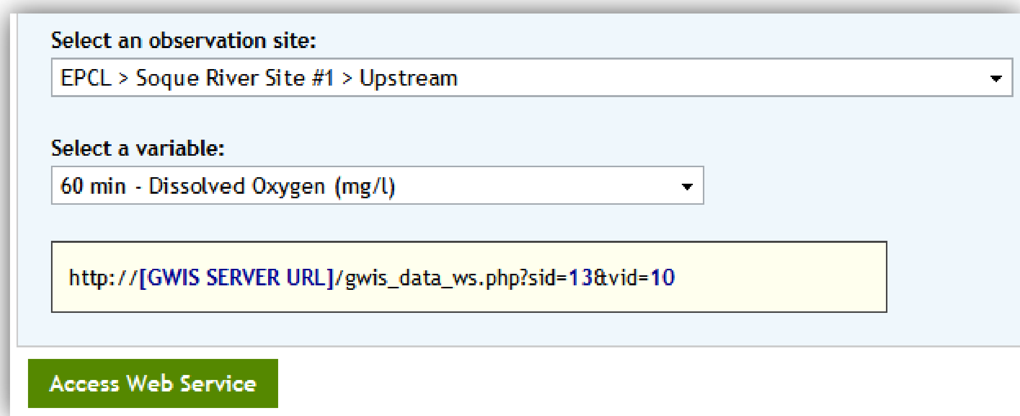


Figure 3.4. Interface of GWIS data web service tool

A REST web service URL can be generated by selecting an observation site and a variable from the list provided in the interface (Figure 3.4). The web Service URL returns the requested variable data in a time series as an XML file (Figure 3.5), depending on the parameters provided in the URL.


```
<gwis_ws>
  <site>
    <project_name>EPCL</project_name>
    <project_website>http://www.modeling.uga.edu</project_website>
    <site_id>12</site_id>
    <site_name>Weracoba Creek, Columbus #1</site_name>
    <site_location>BMP, Influent</site_location>
    <site_huc>03130003</site_huc>
    <site_info>-</site_info>
    <dec_lat>32.4837</dec_lat>
    <dec_lng>-84.96499</dec_lng>
  </site>
  <variable>
    <var_id>23</variable_id>
    <var_name>Dissolved Oxygen</var_name>
    <var_unit>mg/l</var_unit>
    <sampling_frequency>15 min</sampling_frequency>
    <begin_date>2007-11-01</begin_date>
    <end_date>2008-02-29</end_date>
    <time_zone offset="-05:00">EST</time_zone>
  </variable>
  <values>
    <value dateTime="1998-10-05 10:15">20.29</value>
    <value dateTime="1998-10-05 10:30">30.29</value>
    <value dateTime="1998-10-05 10:45">40.29</value>
        ...
  </values>
</gwis_ws>
```


Figure 3.5. GWIS data web service tool sample output

USGS Real-time Data Tool: There are many data sources available (e.g. USGS, EPA) for users in many available web-based information systems. User-uploaded data will be useful for particular studies but users might need data for common variables like flow and weather information. As a prototype IS, GWIS provides not only user-uploaded data, but also interfaces to integrate distributed data sources like USGS. So that users can export the selected time series to their computers, or transfer to GWIS from USGS real-time data source by providing USGS site number, parameter code, and number of days on the interface (Figure 3.6). The data will be exported in XML format with site description and data values. If the user selects to transfer data to GWIS, the data will be available on the GWIS server for further analysis and utilization (Figure 3.7) without any pre-processing.

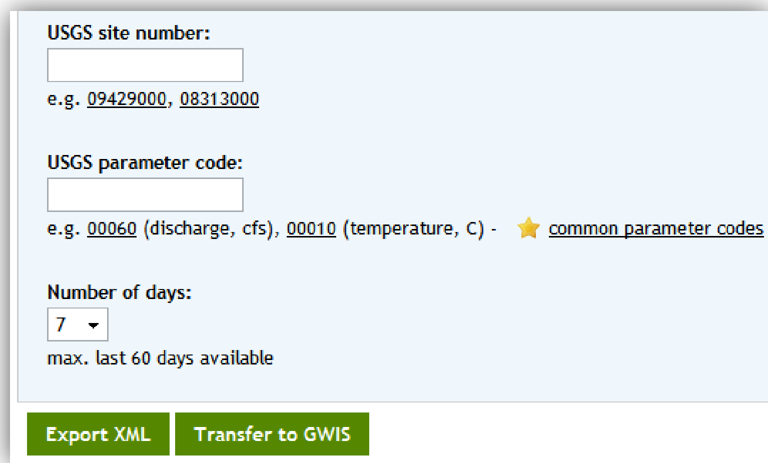


Figure 3.6. Interface of USGS real-time data tool

The tool uses the information provided by the user and accesses the selected data source on the server side using PHP. The data retrieved from the distributed source is processed on the server side and sent to the user's browser to download. If the transfer option is selected by the user, the data are stored on the GWIS server for temporary use.

Figure 3.7. Additional tools to use after transferring real-time USGS data

USGS Daily Data Tool: Users can export data to their computers, or transfer them to GWIS from the USGS daily data source, by providing a USGS site number, parameter code, statistics code, and start and end dates of the sampling range on the interface (Figure 3.8). The data will be exported in XML format with site description and data values. If the user selects to transfer data to GWIS, the data will be available on the GWIS server for further analysis and utilization (Figure 3.7) without any pre-processing.

Data Sources by HUC (Hydrologic Unit Code) Tool: Users can access information sources (Figure 3.9) from the USGS and EPA by providing a HUC number. The tool connects users to related websites for sub-basin information and data sources (real-time and daily).

Figure 3.8. Interface of USGS daily data tool

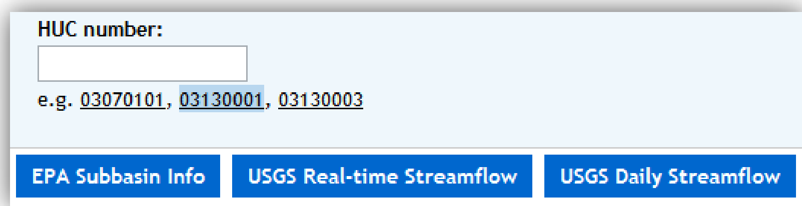


Figure 3.9. Interface of data sources by HUC tool

3.4.2. Analysis Component

GWIS has an integral analysis environment, with statistical analysis tools and expert systems methods. The user can generate statistical summaries of the data and histograms, carry out trend and seasonal component analyses using Dynamic Harmonic Regression (DHR; Young, 2004). Autoregressive polynomial interpolation analysis, and correlation and missing data analysis with Self Organizing Maps (SOMs) are other available tools in the system. The details and screenshots of the analysis tools are given below.

Descriptive Statistics Tool: Descriptive Statistics are used to describe the basic features of the data (Fisher, 2009). They form the basis of quantitative analysis of the data together with simple statistical graphics. The GWIS Descriptive Statistics tool (Figure 3.10) provides values of statistical parameters (minimum, maximum, mean, standard deviation values, and percentage of missing values in the data series) and histogram plots of the data for a selected observation site and any number of variables. The tool runs on CGI by accessing a stand-alone executable file. CGI is a standard interface between external gateway programs and information servers (e.g. HTTP servers). The output from the executable file is sent to the client's browser as an HTML file with images in a report format (Figure 3.11). The reports can be exported as PDF files or printed. Images and results obtained by the analysis can be saved for later use.

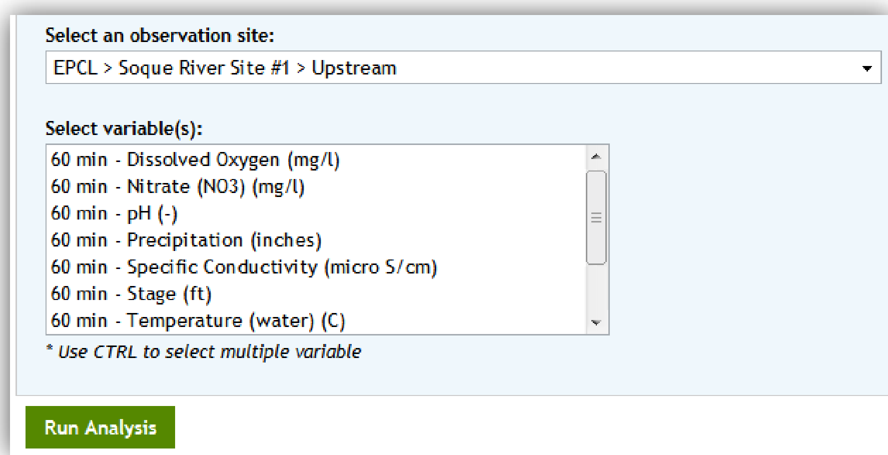


Figure 3.10. Interface of Descriptive Statistics tool

Users can get a better understanding of their data by using descriptive statistics values. A histogram is a display of tabulated frequencies. It shows what proportion of the observation variable values fall into each value range (Histogram, Wikipedia, 2009). It can be used for understanding the distribution and quality of the data.

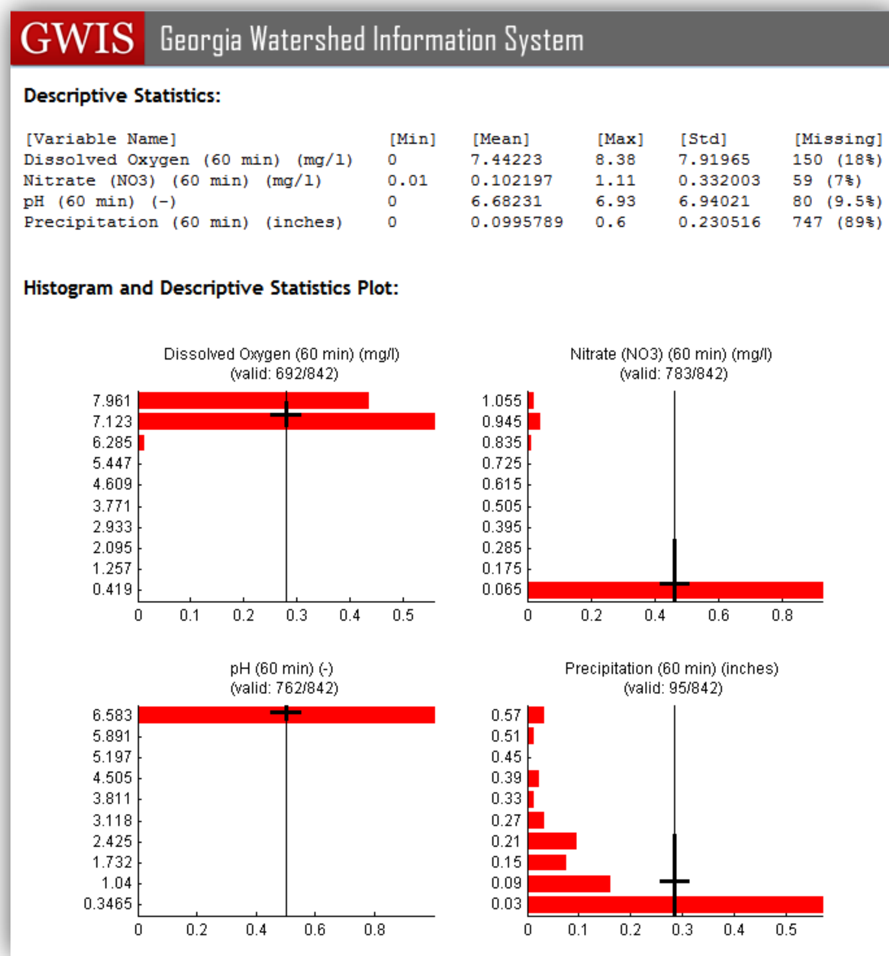


Figure 3.11. Sample output of the Descriptive Statistics tool

Trend and Seasonal Component Analysis Tool: Trends can be extracted from the time series of an observation site variable data and a seasonal component about the trend can be modeled with an Auto-Regressive (AR) model by using this analysis tool (Figure 3.12). In statistics, an autoregressive model is a type of random process which is often used to model and predict various types of data behavior. An autoregressive model depends on the statistical properties of the past behavior of a variable (its derivatives in time) to predict future values of the variable. This analysis tool is built with the MATLAB Captain Toolbox. The Captain Toolbox is a MATLAB compatible toolbox (Young, 2004) for time series analysis (Romanowicz, 2006), forecasting (Taylor, 2007), system identification, signal processing interpolation and outlier smoothing (Beck et al., 2010; Lin and Beck, 2006; Lin, 2003).

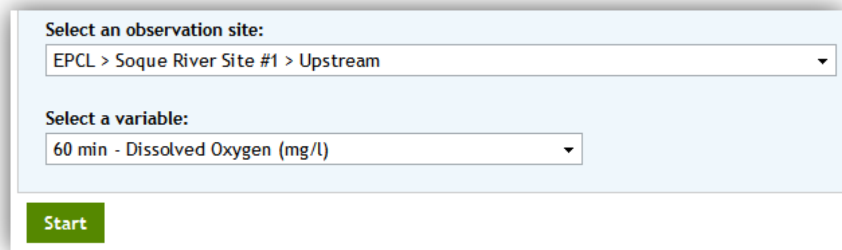


Figure 3.12. Interface of Trend and Seasonal Component Analysis tool

The tool runs on CGI by accessing a stand-alone executable file. The output from the executable file is sent to client's browser as an HTML file with images in a report format (Figure 3.13 and Figure 3.14). The reports can be exported as a PDF file or printed. Images and results obtained by the analysis (trend and seasonal component time series) can be saved for later use. The report provides plots and a data file of the trend and seasonal component of the time series for the selected variable. The trend is a long term pattern, showing the direction and rate of change in a time series. The seasonal component is the variation in a time series that describes any regular fluctuations.

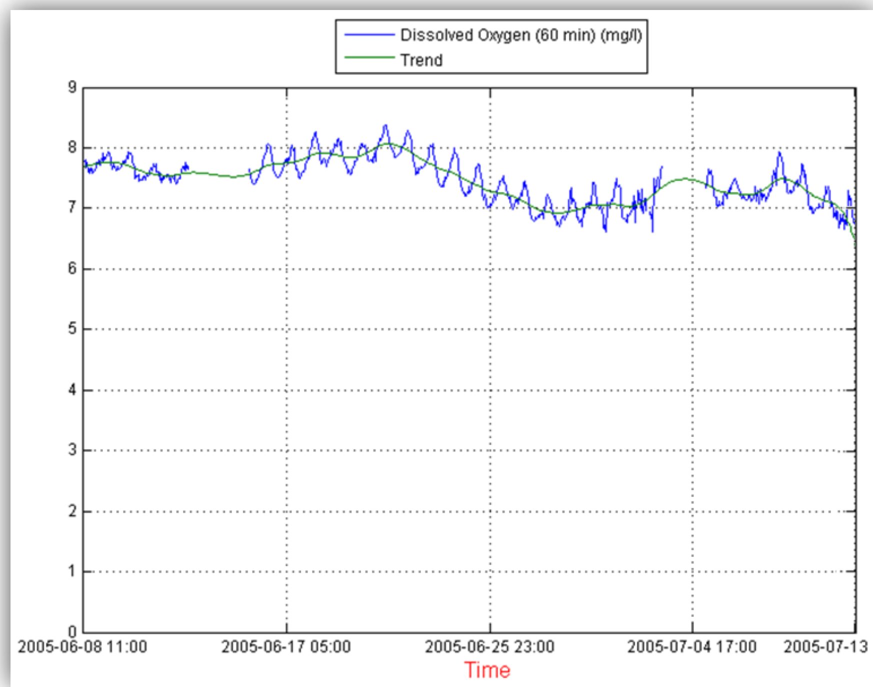


Figure 3.13. Sample output of the Trend Analysis

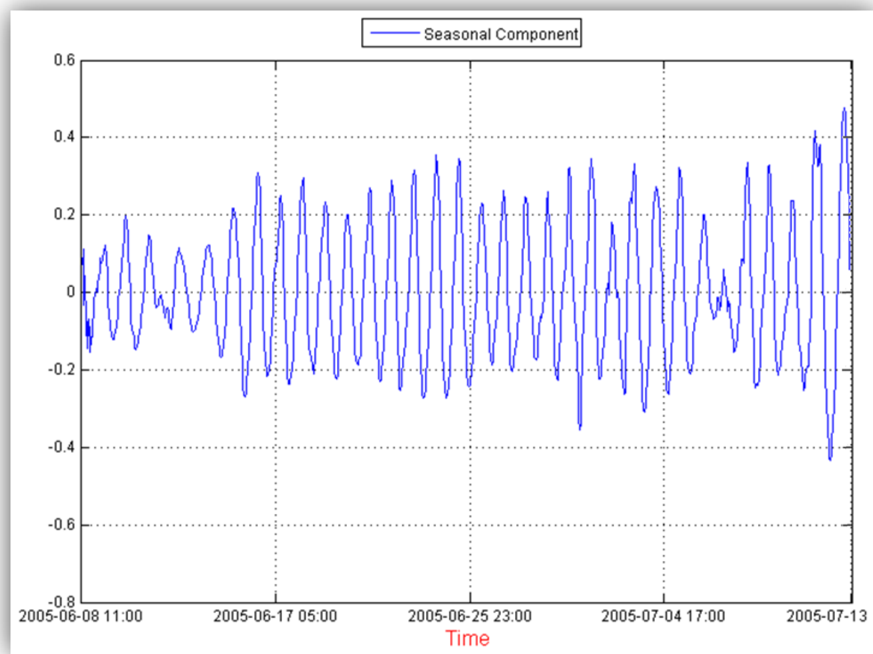


Figure 3.14. Sample output of the Seasonal Component Analysis

Autoregressive Polynomial Interpolation Analysis Tool: This analysis tool extracts the trend and seasonal components of a time series, and generates a new time series to fit the actual observations by using Autoregressive (AR) Polynomial Interpolation. Polynomial interpolation is used to interpolate a data set by using a polynomial function. It can be used to approximate functions and their derivatives. More information is available in a review by Mateescu (2008) for polynomial interpolation and its applications to autoregressive models. This analysis tool (Figure 3.15) is built with MATLAB Captain Toolbox (Young, 2004).

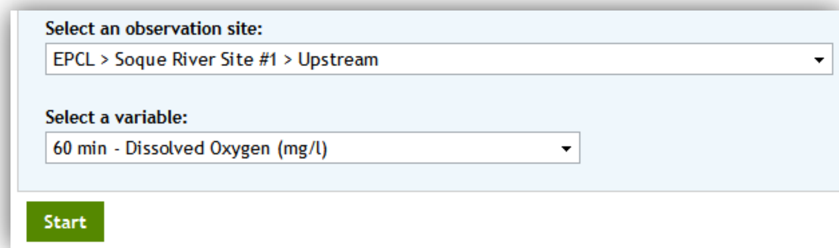


Figure 3.15. Interface of Autoregressive Polynomial Interpolation Analysis tool

Users can generate an interpolated time series for data sets that have missing values by using this tool. The tool runs on CGI by accessing a stand-alone executable file. The output from the executable file will be sent to client browser as an HTML file with images in a report format (Figure 3.16). The red line in Figure 3.16 shows the predicted interpolation and the blue dots are the actual time series values of the selected variable. Images and results obtained by the analysis (time series of interpolation) can be saved for later use.

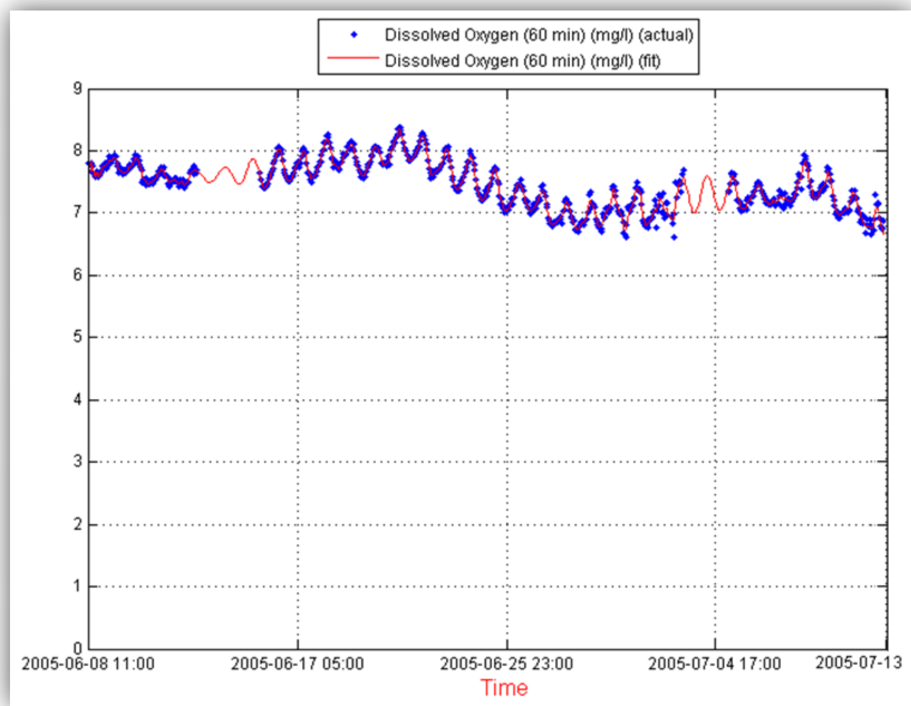


Figure 3.16. Sample output of the AR Polynomial Interpolation Analysis tool

Correlation Analysis with Self Organizing Maps Tool: This analysis tool uses Self Organizing Maps to generate colorful visualizations of observation data to reduce dimensionality of the data and helps users see the pattern of the time series and extract hidden correlations between variables. Correlation analysis is very important in selecting variables for modeling applications. Correlation analysis can be used to determine which variables have close correlation and as a model preparation tool for environmental models. However, for some data sets correlation cannot be exposed with simple correlation analysis, hence the introduction of SOMs.

'Self Organizing Maps' (SOMs) is a data visualization technique (Kohonen, 1995) which reduces the dimensions of data through the use of self-organizing neural networks. Neural networks are used in many environmental science studies (Demir, 2003; Aydiner et al., 2009; Sunlu et al., 2009; Bayar, 2009; Onkal-Engin et al., Sunlu et al., 2007; 2005a; Onkal-Engin et al., 2005b; Aydiner et al., 2005). It is not easy to visualize and understand high dimensional data with traditional visualization approaches. The SOMs reduce dimensions by generating a two-dimensional map which shows the similar patterns of the data by grouping them together. SOMs are utilized in many studies for pattern recognition (Alhoniemi, 2002), data mining, visualization, user interfaces (Himberg, 2004), data exploration (Kaski, 1997), variable selection and feature extraction (Laine, 2003). This analysis tool (Figure 3.17) is built with the MATLAB SOM Toolbox (Vesanto, 1999). Self organizing maps are utilized successfully in many environmental studies (Lee and Scholz, 2006; Manolakos et al., 2007; Lee et al, 2007; Zhang et al., 2008; Zhang et al., 2009).

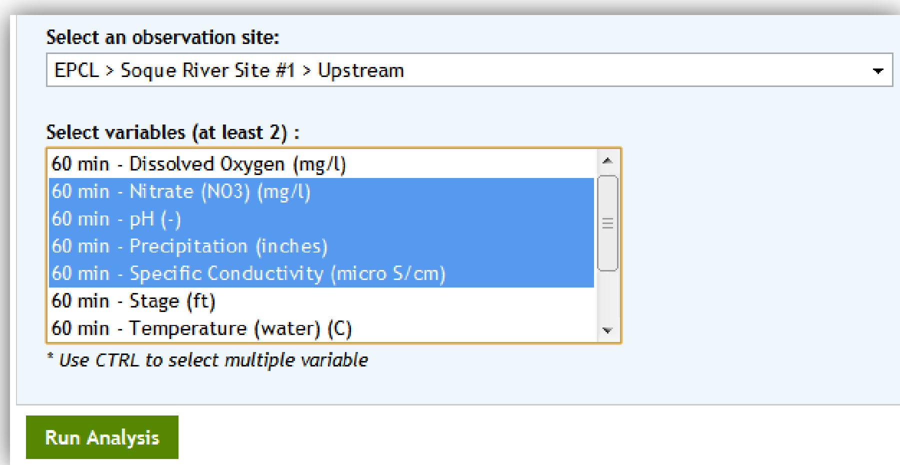


Figure 3.17. Interface of Correlation Analysis with SOM tool

The tool runs on CGI by accessing a stand-alone executable file. The output from the executable file will be sent to the client browser as an HTML file with images in a report format (Figure 3.18). The output includes descriptive statistics of the variables, optimized internal model parameters and the SOM components maps graph. Images and results obtained by the analysis (pattern maps of the variables) can be saved for later use.

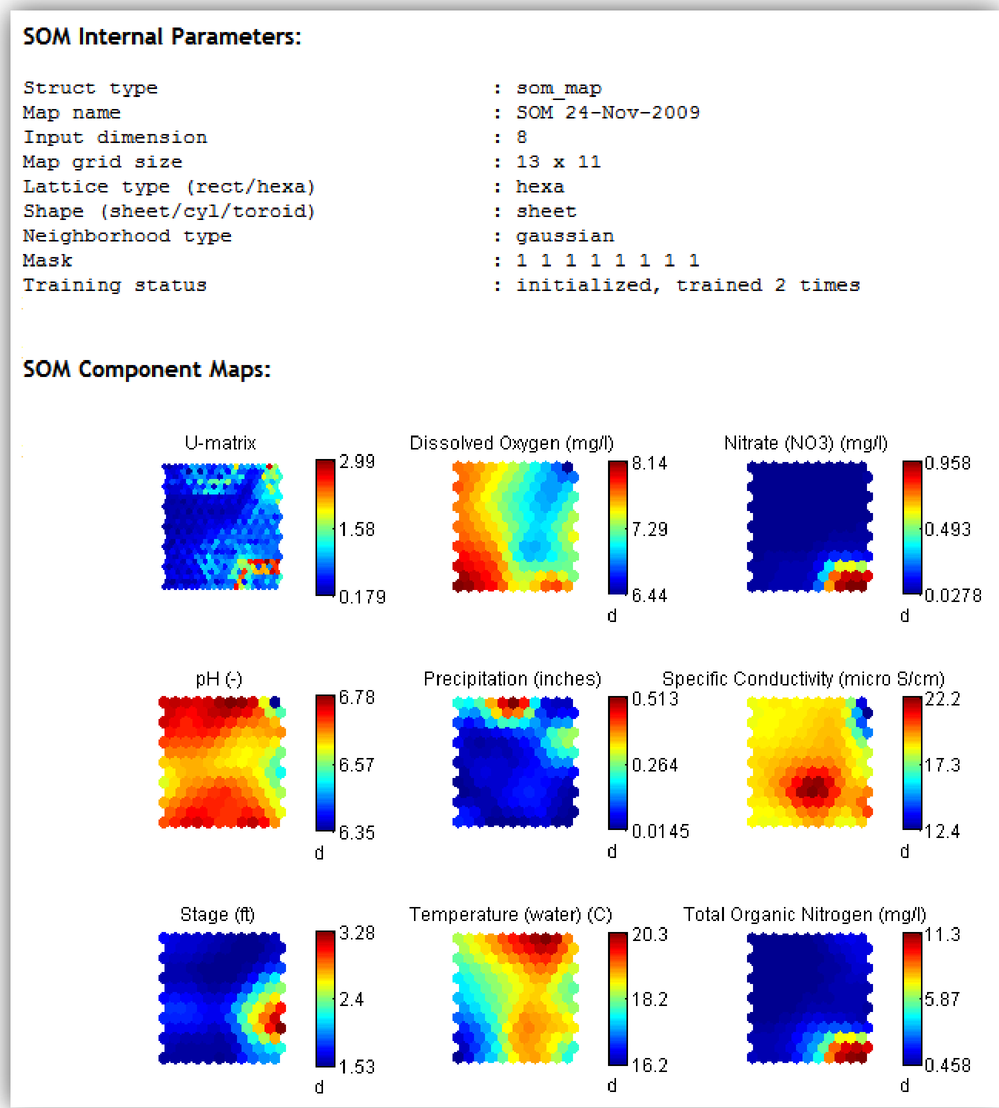


Figure 3.18. Sample output of the Correlation Analysis with SOM tool

Missing Data Analysis with Self Organizing Maps Tool: This tool can be used for analyzing and completing a data set with missing values. Instead of using a univariate method for filling each missing value, a multivariate approach by using SOM replaces missing data with the best matching value of an n-dimensional vector. Univariate analysis techniques utilize only a single dependent variable. Details and the background theory of SOMs are given above. Multivariate analysis methods provide a combination of dependent variables that meets a selected statistical criterion. This analysis tool (Figure 3.19) is built with MATLAB SOM Toolbox.

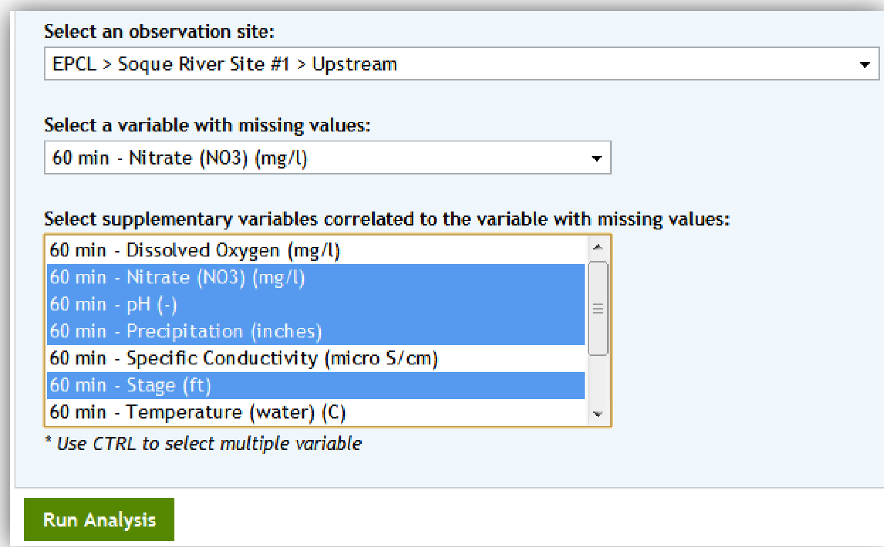


Figure 3.19. Interface of Missing Data Analysis with Self Organizing Maps tool

The user needs to select a variable with missing values from an observation site, and any number of supplementary variables that might be correlated with the target variable. The tool runs on CGI by accessing a stand-alone executable file. The output from the executable file will be sent to client browser as an HTML file with images in a report format (Figure 3.20). The output includes descriptive statistics of the variables, scatter and line plots of actual and predicted time series values of the variable, and link to download the predicted time series values. Images and results obtained by the analysis (completed data series) can be saved for later use.

3.4.3. Modeling Component

Even as a web-based Information System, GWIS demonstrates advanced capabilities to run a complex mechanistic model. As an example of such a mechanistic model, the Watershed Water Quality Simulator (WWQS) has been incorporated into GWIS for water quality simulations.

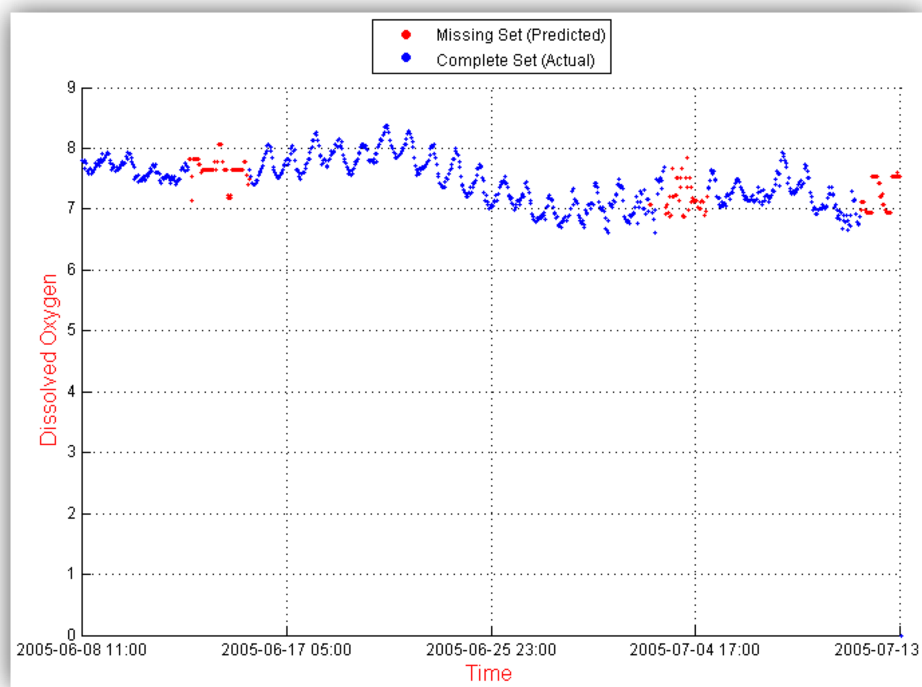


Figure 3.20. Sample output of the Missing Data Analysis with SOM tool

Watershed Water Quality Simulator: WWQS is new-generation watershed model developed by the Environmental Informatics and Control Program in the University of Georgia (Feng, 2008; Feng and Beck, 2008). It is designed for continuously simulating the watershed hydraulics, sediment and pollutant behavior, for evaluating the effects of pollution discharges and water quality conditions. The development of WWQS is favored by a dual concept of mechanistic and data-based modeling. It features a fully interactive basis to describe the hydrological process for both the land and water compartment of a watershed. This feature makes it able to simulate transient, unsteady conditions. The model is applied in several basins in North Georgia, U.S., and has achieved good performances. The main interface (Figure 3.21) with five steps (Figure 3.22 - Figure 3.26), and sample results (Figure 3.27) are given below with screenshots.

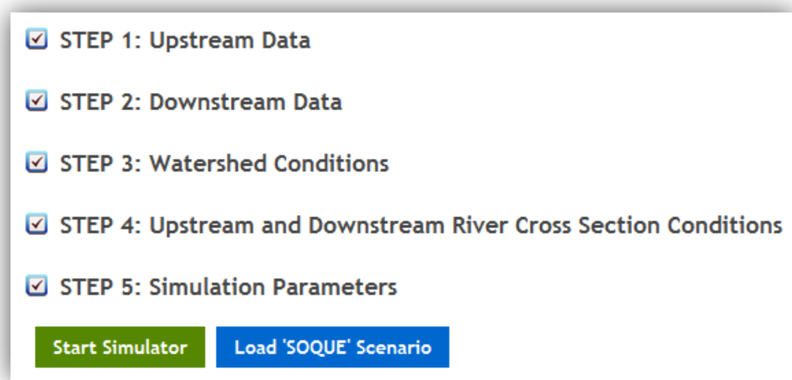


Figure 3.21. Interface of Watershed Water Quality Simulator tool

In the first step, users need to select an observation site and time series for flow, suspended solid, total reactive phosphorus and precipitation. The second step requires for three of these variables (flow, suspended solid, and total reactive phosphorus) for the downstream location. Third step is where users provide watershed conditions like drainage area, potential labile phosphorus content in soil and potential phosphorus content in soil. In the fourth step, users need to enter upstream and downstream river cross-section conditions like elevation, bottom and side slope, and distance between upstream and downstream locations. The fifth step requires values of 20 simulation parameters [Curve Number (CN) value for dry condition, CN value for moderate condition, CN value for wet condition, groundwater discharge rate, Manning's roughness coefficient, sediment entrainment coefficient, representative particle dimension, coefficient in erosion equation, coefficient in erosion equation, coefficient in erosion equation, average sediment concentration in base-flow, suspended solid dispersion coefficient, phosphate adsorption coefficient, phosphate adsorption equilibrium coefficient, phosphorus soil partitioning coefficient, phosphorus mobilization factor, phosphorus eluviation factor, rainfall wash off lasting time, soil evapotransporation rate, phosphate dispersion coefficient] from users.

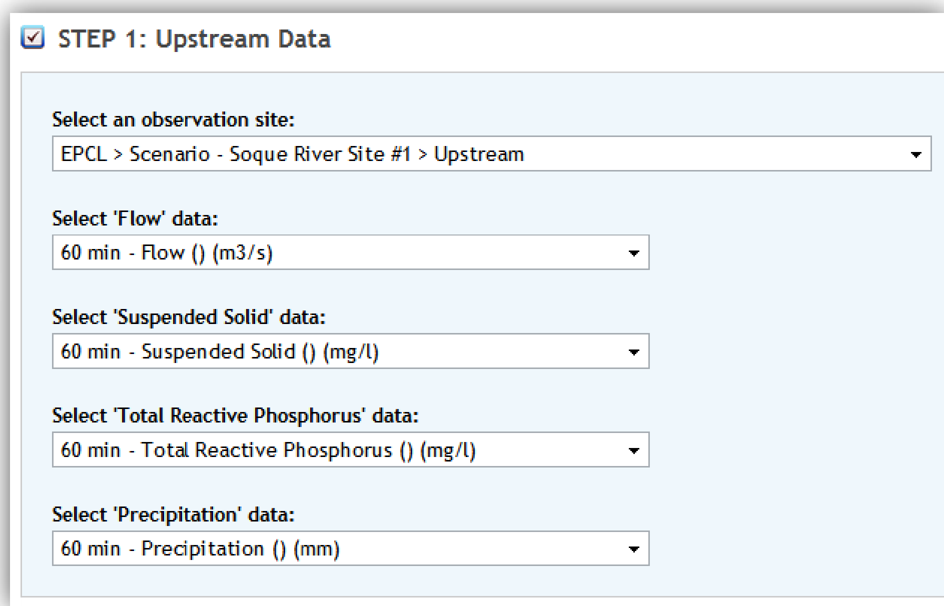


Figure 3.22. Step 1 of WWQS tool for upstream data inputs

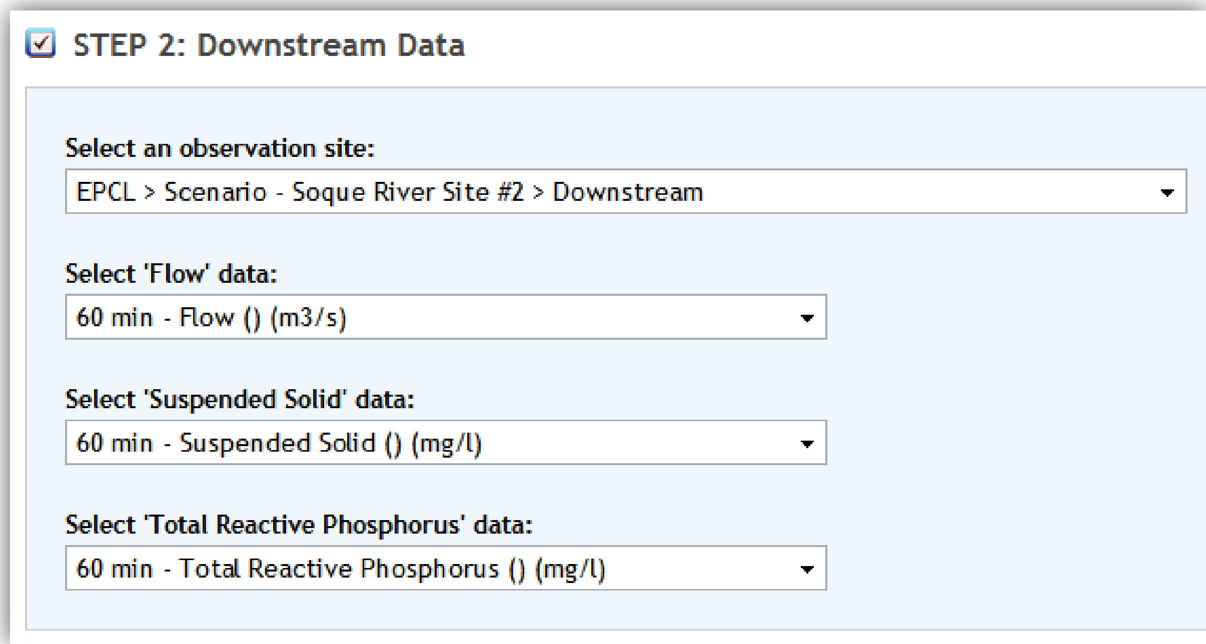


Figure 3.23. Step 2 of WWQS tool for downstream data inputs

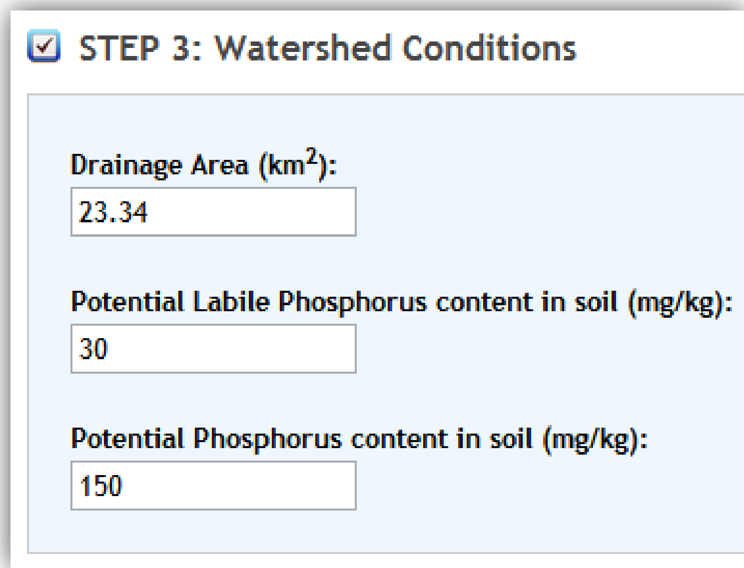


Figure 3.24. Step 3 of WWQS tool for watershed conditions input

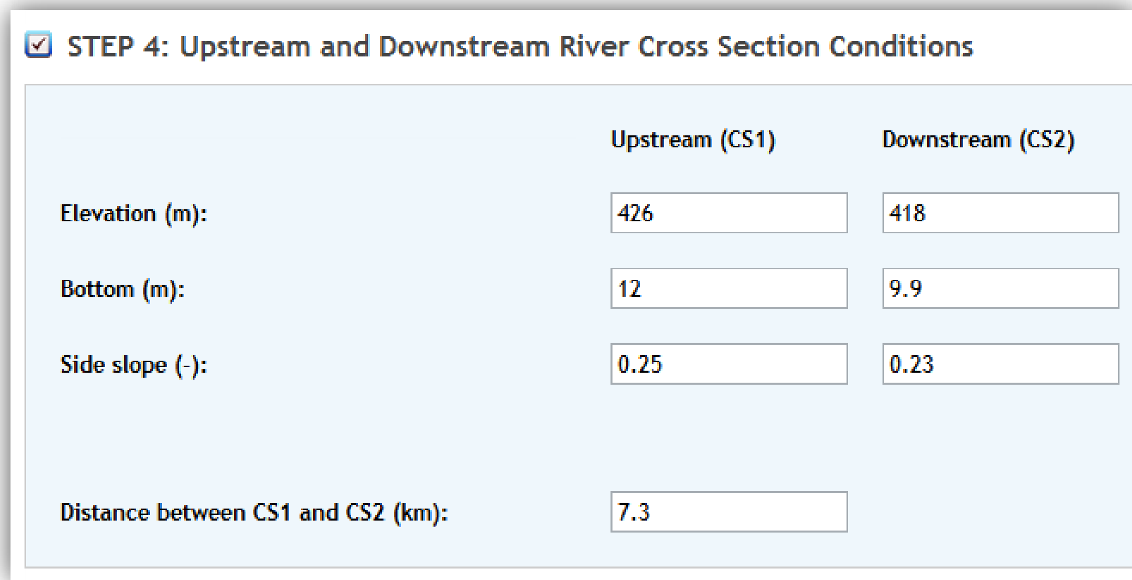


Figure 3.25. Step 4 of WWQS tool for river cross section conditions input

The modeling tool runs on CGI by accessing a stand-alone executable file. The output from the executable file will be sent to client browser as an HTML file with images in a report format (Figure 3.27). The output includes line plots of actual and simulated time series values of the variables flow, suspended solids and total reactive phosphorus concentrations, and confidence intervals as an area plot around the simulated time series. The confidence interval is a common way of quantifying the error in an estimate. Related literature and examples on confidence intervals can be found on many web sources (Wikipedia, 2010). The report page also provides a link to download the simulated time-series values. Images and results obtained by the analysis (simulated data series) can be saved for further analysis.

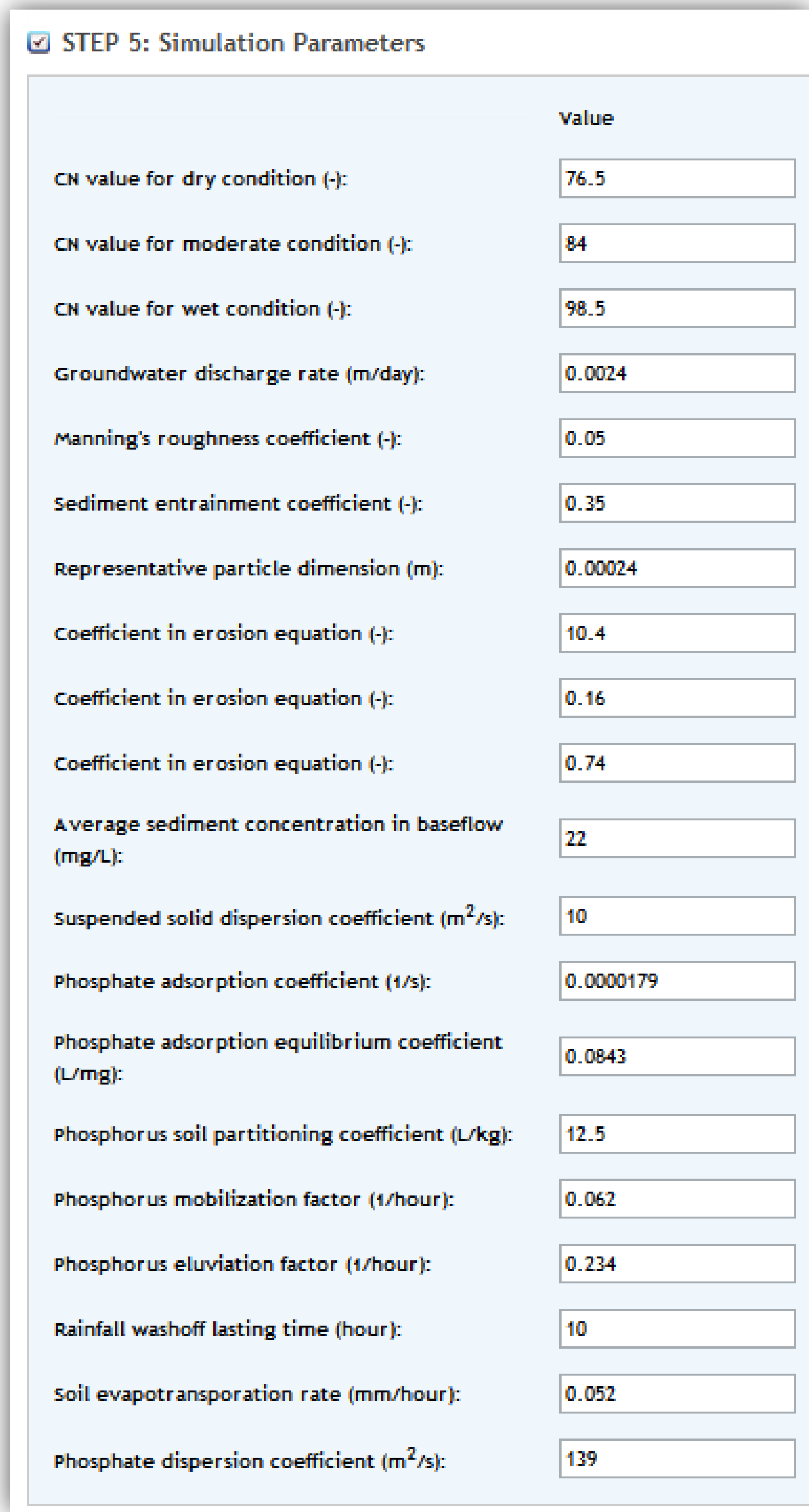


Figure 3.26. Step 5 of WWQS tool for simulation parameters input

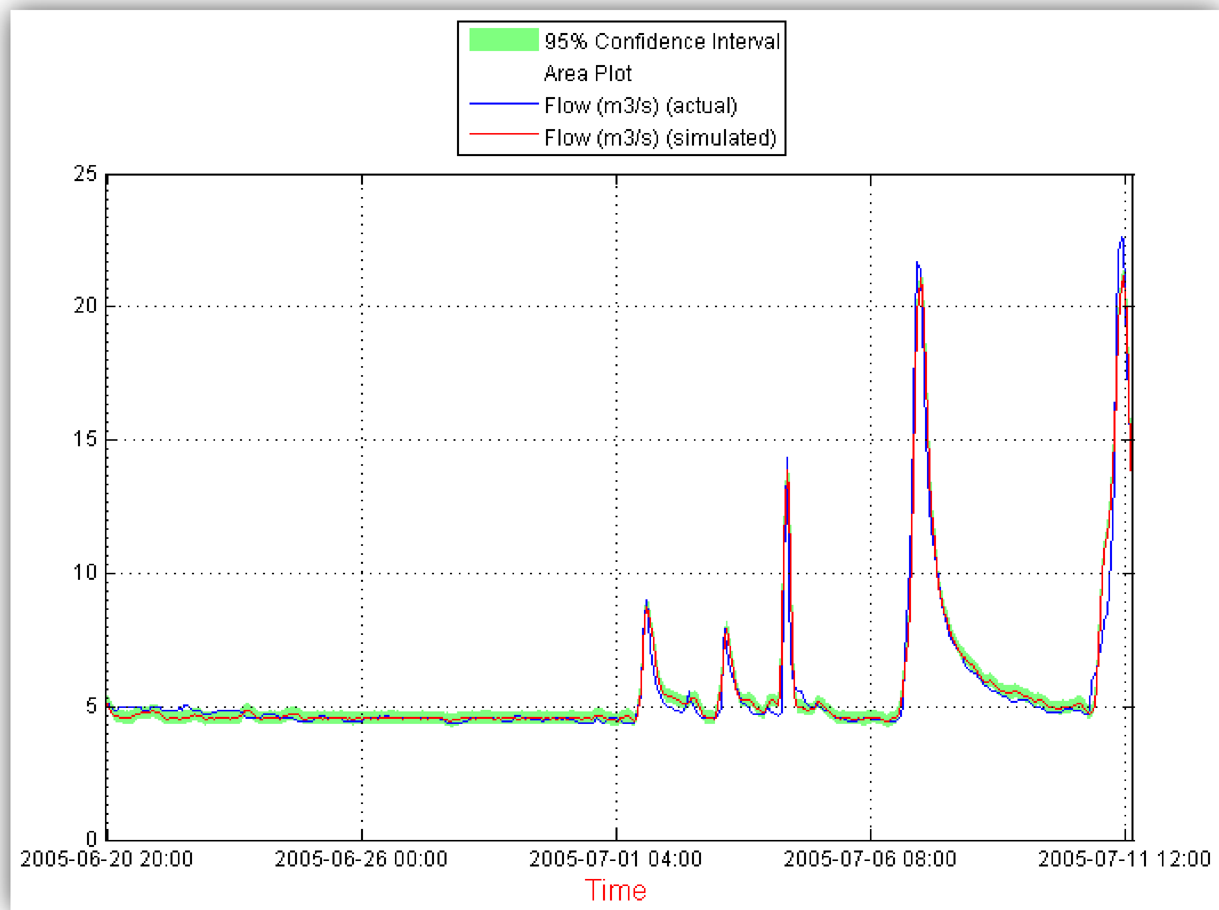


Figure 3.27. Sample simulation results for flow by WWQS tool

3.4.4. Visualization Component

The visualization component includes tools to visualize water quality variables, compare variables from different monitoring sites, export graphs, generate reports, carry out statistical visualizations, and overlap meta-data on graphs. GWIS also provides an advanced visualization interface integrating the mapping, analysis and visualization capabilities of GWIS for visualizing time series variables on interactive maps. All the resulting graphs can be exported as PDF or image files. The details and screenshots of the visualization tools are given below.

Spatial Visualization on Interactive Maps Tool: Even the advanced technical solutions to environmental issues can be challenged with the need to engage with the broader society and its institutions. It is important to provide high quality environmental information in a timely and cost-effective manner to the general public and decision makers. Environmental Information Systems show promise of contributing to the social component of the triple-bottom-lines of the sustainability with its novel scientific visualization capabilities (Demir, 2009c). GWIS provides an advanced scientific visualization tool that integrates mapping, visualization and analysis features on a web-based system.

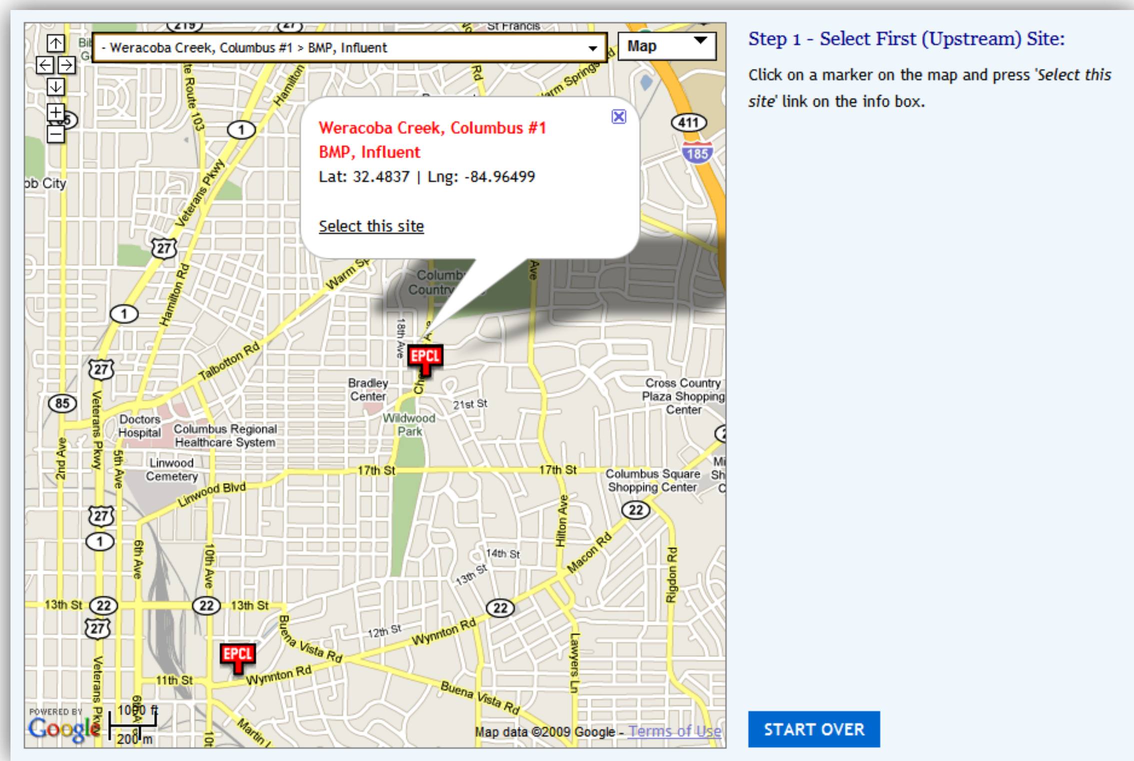


Figure 3.28. Interface of Spatial Visualization on Interactive Maps tool

The interface (Figure 3.28) allows users to create an animation of the time series overlaid on the physical domain (e.g. river segment). Users can select any number of observation sites with one variable to create an interpolated time series for each internal segment in between the observation sites. The flow structure of the domain can be modified by users (e.g. river structure or process flow in a WWTP).

Several web technologies are used to build the interface, such as PHP, mySQL, AJAX, Javascript, JSON, JQuery Javascript library, Google Map Services, and HTML. Having novel data and mapping tools in GWIS, makes it easier to build such a scientific visualization interface. The interface communicates with the data and mapping services components of GWIS to visualize data with rich geographic content. Much time and effort would be required to build a similar visualization tool without an information system like GWIS.

Steps to carry out a general case study by using the visualization interface are as follows:

- Selecting observation sites and variables (Figure 3.29) by navigating on the interactive satellite map (minimum two sites, e.g. upstream and downstream locations).
- Adding additional markers on the map (Figure 3.30) to identify the structure of the physical domain (e.g. river channel or process flow for a WWTP). *GWIS automatically creates additional segments for a river channel and generates observation data by interpolation for internal sections for smooth visualization at this step.*

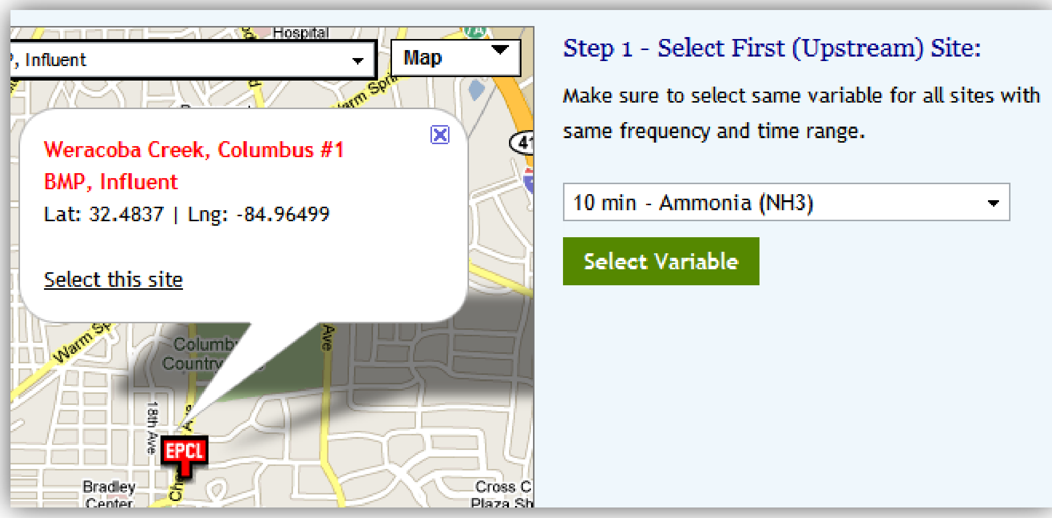


Figure 3.29. Step for selecting observation site and variable

The interface is ready for visualization (Figure 3.31) after the last step. Animation controls, or a slider (Figure 3.32), can be used to visualize data on maps. Slider allows users to go back and forth in time through the time series data for the observed water quality values. An animation can be played and paused from any point in time. Color scale corresponds to the concentration value of the selected variable where dark red refers to the highest concentration and dark green to lowest concentration in the time series.

A centralized approach, using a comprehensive information system to manage data, carry out complex hydrological simulations, and visualize results on interactive maps on a web-based interface, will increase the involvement of non-expert stakeholders and the general public in addressing and resolving environmental issues.

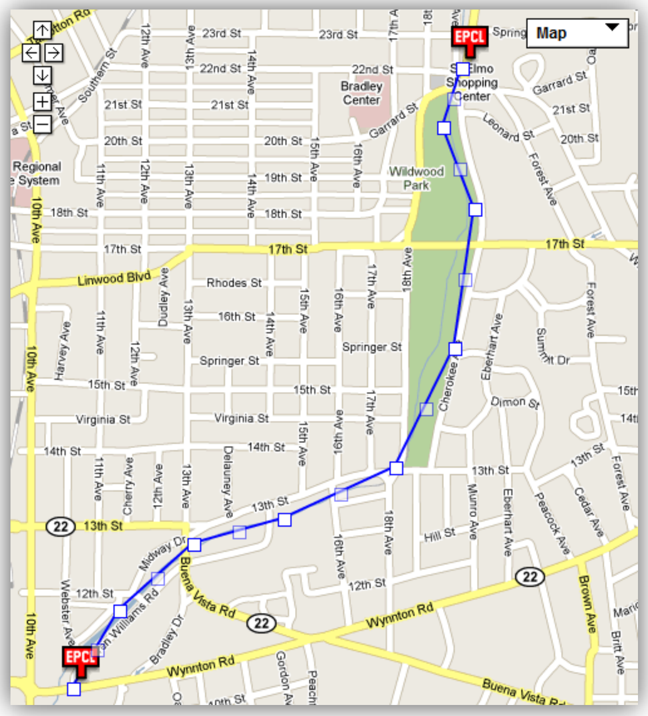


Figure 3.30. Step for identifying structure of the physical domain (e.g. river channel)

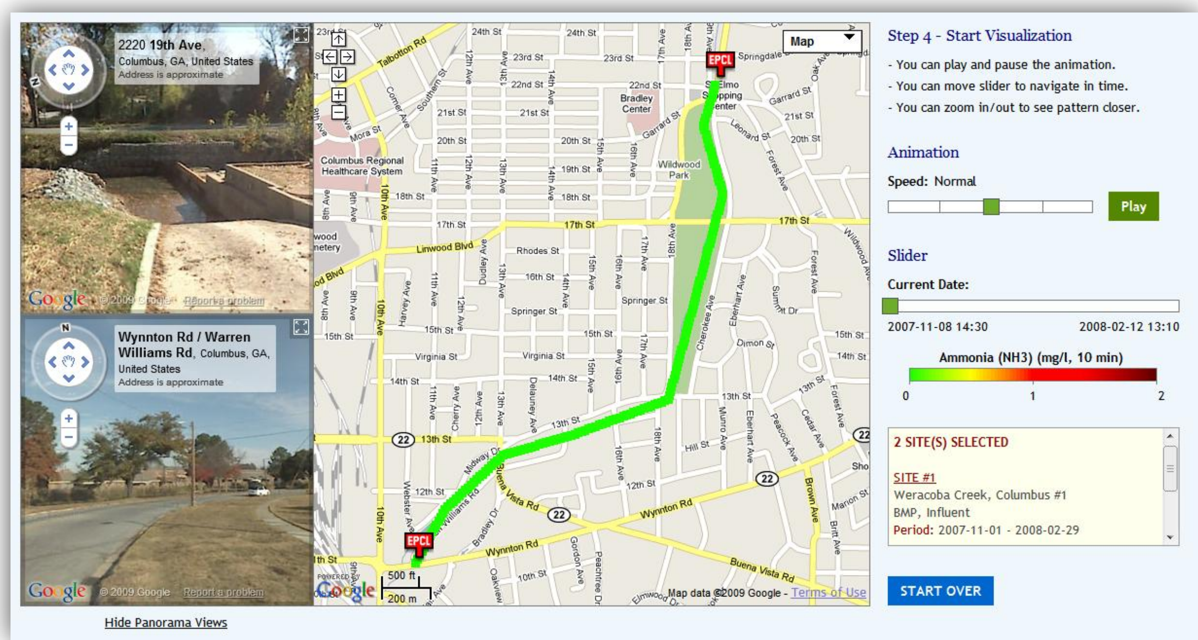


Figure 3.31. Visualization tool is ready to play animation

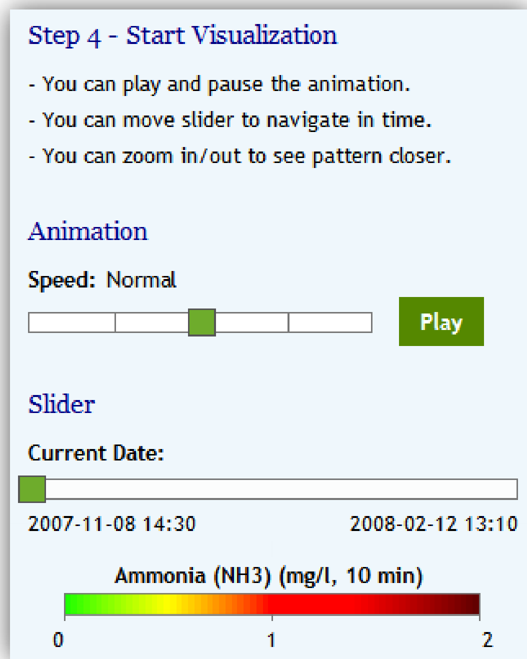


Figure 3.32. Animation controls of the spatial visualization tool

Visualization and animation provide information that the non-expert user can easily understand and allows them to compare results for different management scenarios. This novel scientific visualization can be used as a tool for involving stakeholders in environmental decision making, engaging the general public in understanding the natural and built environment, and thus furthering the social legitimacy of the proposed technical solution (Demir, 2009c).

Multi-dimensional Scatter and Histogram Charts: The interface (Figure 3.33) allows users to generate multi-dimensional scatter and histogram graphs combined in a [n x n] matrix (Figure 3.34). It can be used to compare any number of variables, extract correlations between variables, and get statistical information (e.g. density estimation, frequency of values) from histogram charts.

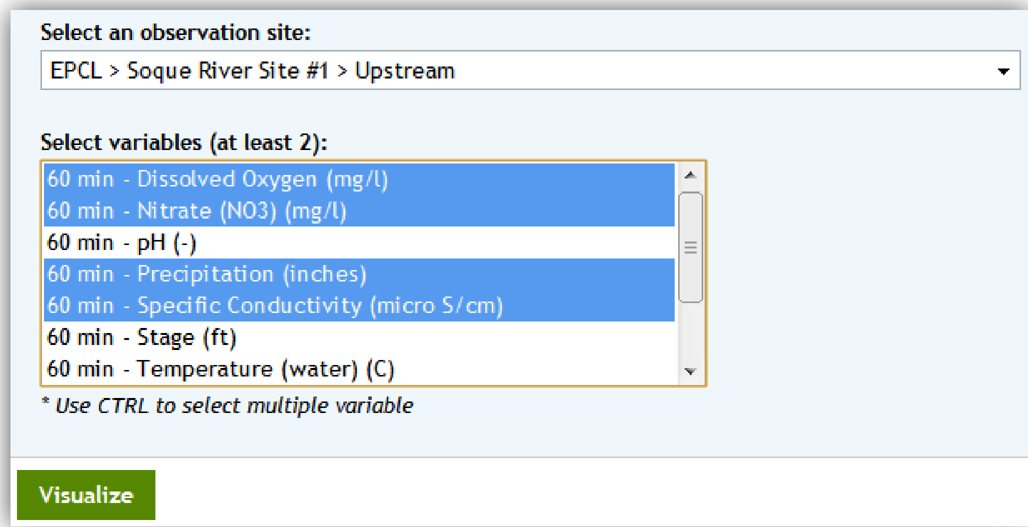


Figure 3.33. Interface of Multi-dimensional Scatter and Histogram tool

Interactive Charts: The interface (Figure 3.35) allows users to visualize a variable, or compare a variable from two separate observation sites (e.g. upstream and downstream) through an interactive tool. The tool provides interactive navigation through time series, various zoom options, and meta-data overlays (Figure 3.36). Meta-data often include information about the sampling campaign and data. It is helpful to understand the anomalies in the data series.

Static Charts: The visualization interface (Figure 3.37) allows users to visualize any number of variables from one observation site in one graph (Figure 3.38). It can be used to compare time series of variables, and help model preparation.

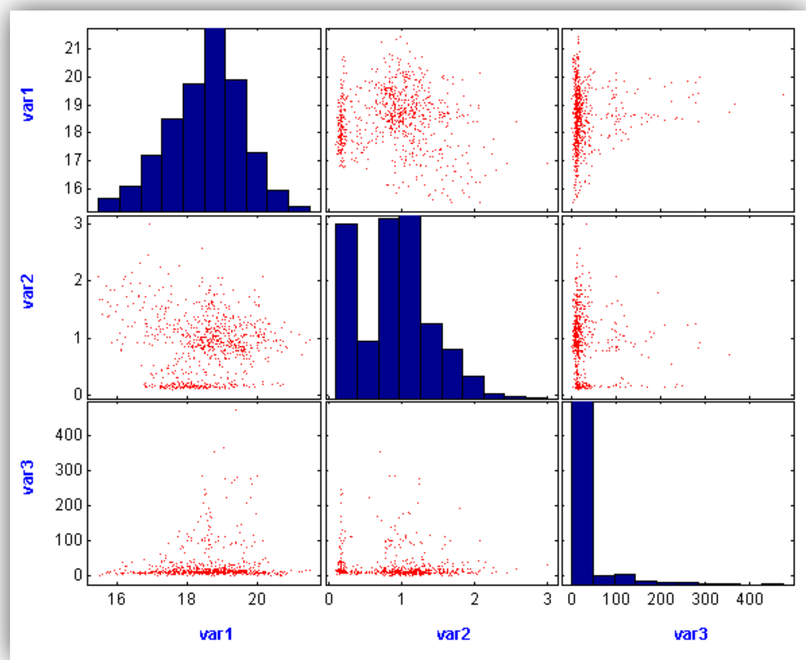


Figure 3.34. Multi-dimensional scatter and histogram graphs

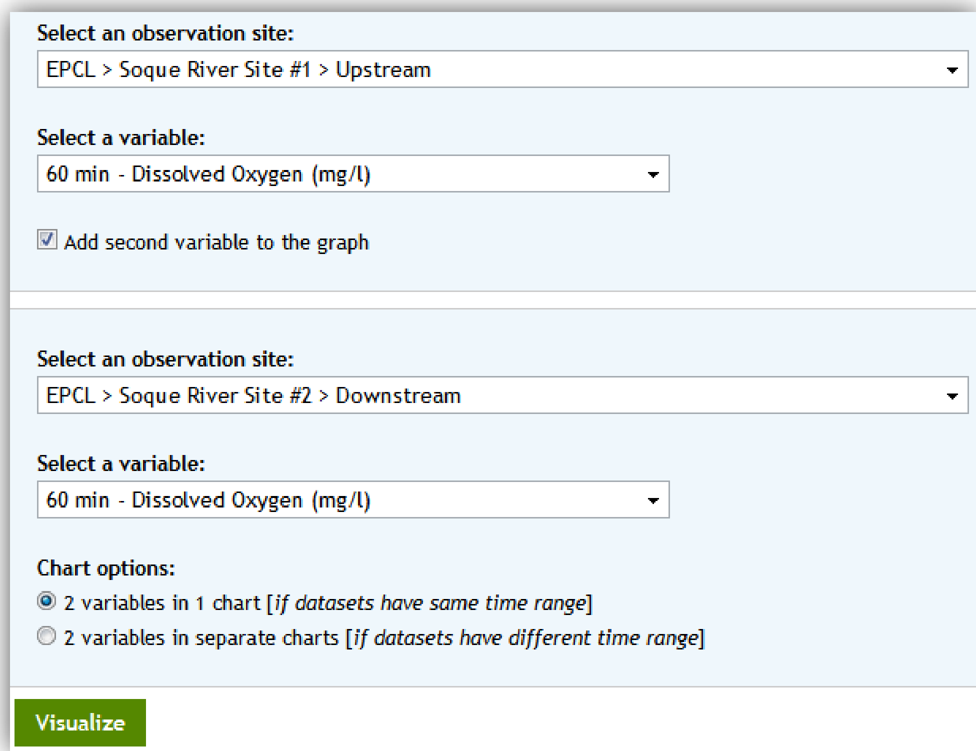


Figure 3.35. Interface of the Interactive Charts tool

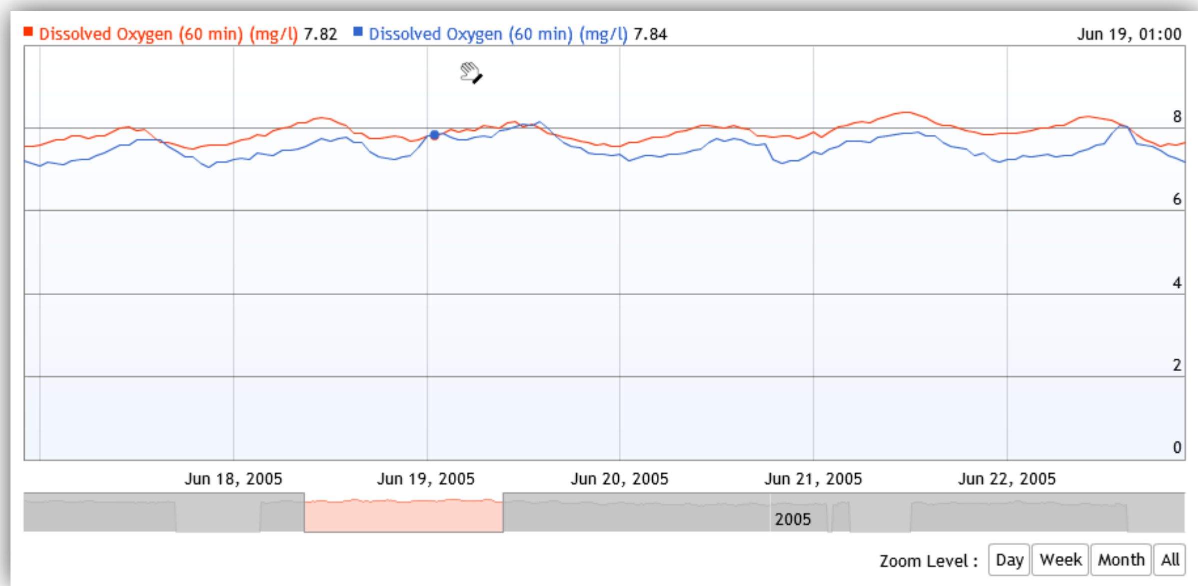


Figure 3.36. Comparison of upstream and downstream values of dissolved oxygen

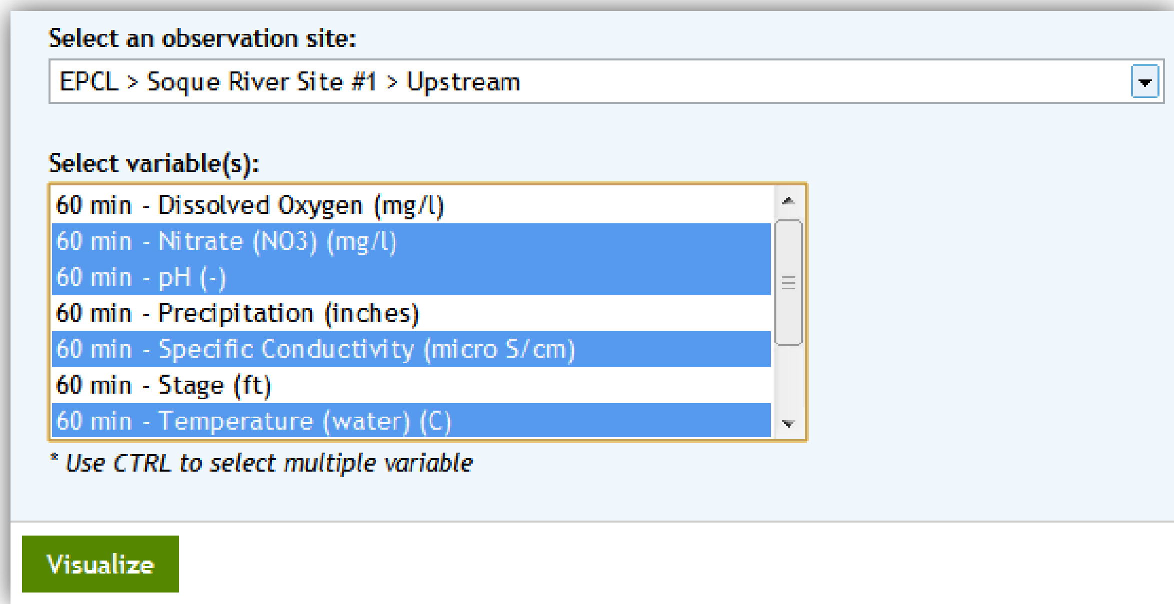


Figure 3.37. Interface of the Static Charts tool

Figure 3.38. Sample static visualization of four water quality variables

3.4.5. Mapping Component

Being integrated with several mapping services (e.g. Google Maps API, Google Earth API, Microsoft Virtual Earth API, and Yahoo Maps API), GWIS provides online mapping features to locate monitoring sites, with then a site description and access to related tools and services. The mapping component supports KML (Keyhole Markup Language) to overlay rich geographic content on interactive maps. As a novel and advanced feature, GWIS provides a USGS Interactive Information Interface.

GWIS Mapping Services: While providing high quality map types (e.g. street, satellite, hybrid, 2D/3D terrain, 360° panoramic, bird's-eye view and 3D building overlays), GWIS web-based mapping distributes network and storage loads to third-party web service providers. The interface is integrated with Google, Yahoo and Microsoft Map Services (Figure 3.39 - Figure 3.42) to provide mapping capabilities. Users can navigate on interactive maps with various map types to access monitoring sites, data and information available on GWIS servers. Most of the mapping services (e.g. Google Maps) provide an Application Programming Interface (API) to embed maps to other websites and information systems by using JavaScript. API is an interface to provide interaction between a software program and other programs or information systems. JavaScript is a scripting language, integrated to the web browser that allows the development of interactive user interfaces and advanced websites. The mapping APIs provide several utilities to modify maps and add information to the map through different services.

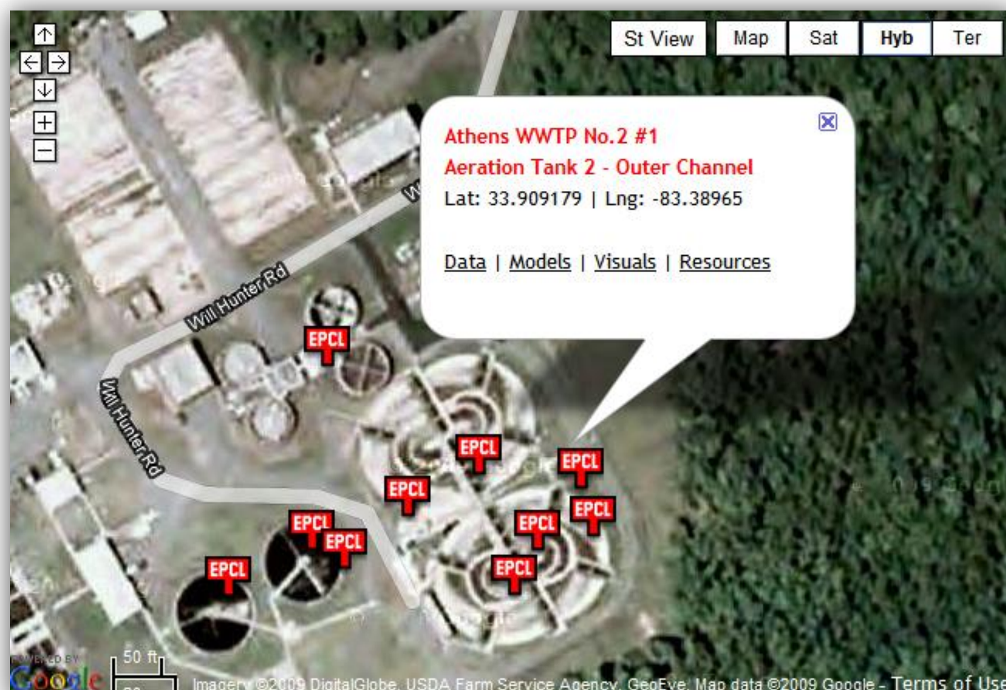


Figure 3.39. Interface of the map services powered by Google Maps API

Figure 3.40. Interface of the map services powered by Yahoo Maps API

Figure 3.41. Interface of the map services powered by Microsoft Virtual Earth API

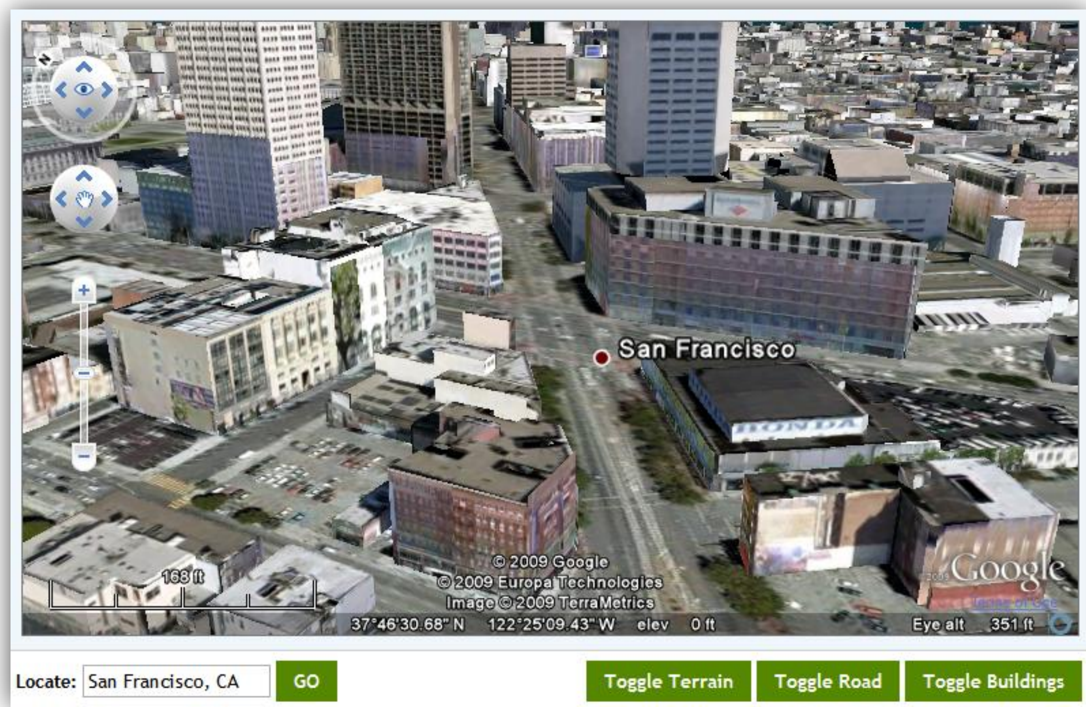


Figure 3.42. Interface of the map services powered by Google Earth API

KML Support on GWIS Maps: KML is an XML-based language schema (Figure 3.43) for expressing geographic annotation and visualization. GWIS Maps have support for the KML (Figure 3.44) standard with the following elements: placemarks, icons, folders, descriptive HTML, KMZ (compressed KML, including attached images), polylines and polygons, including color, fill, and opacity, network links to import data interactively, ground overlays and screen overlays. KML files can be shared by e-mail, hosting them locally for sharing within a private internet, or hosting them publicly on a web server. GWIS provides several example KML files from various internet sources for embedding environmental information on maps like per capita CO_2 emissions of states, EPA AirNow national air quality information, USGS real-time stream-gages, protected areas in Bolivia, EPA air emission sources and many others.

```
<?xml version="1.0" encoding="UTF-8"?>
<kml xmlns="http://www.opengis.net/kml/2.2">
<Placemark>
  <name>New York City</name>
  <description>New York City</description>
  <Point>
    <coordinates>-74.006393,40.714172,0</coordinates>
  </Point>
</Placemark>
</kml>
```

Figure 3.43. Sample KML document for New York city location

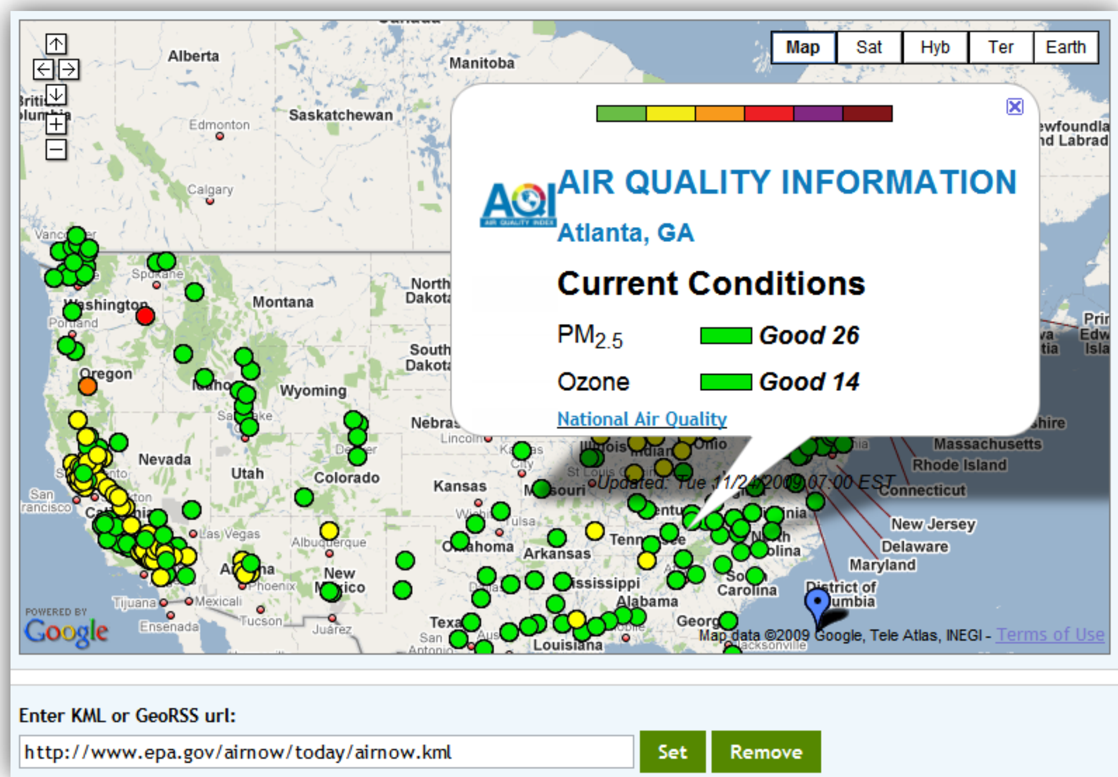


Figure 3.44. Sample screenshot of KML support on GWIS Maps

USGS Interactive Information Interface: GWIS has a novel information interface (Figure 3.45) combining mapping and integration capabilities to access USGS monitoring sites, and site related data and information. The interface was developed in 2008 and presented on many occasions (Demir, 2009a) as a part of GWIS, including the Georgia Water Resources Conference (Demir, 2009b) in April 2009. Most of the web-based information systems (e.g. USGS, EPA) provide complex interfaces for data queries. Few of these systems provide alternative and easy interfaces for low-end users. It is noteworthy that the USGS developed a similar online mapping interface for data access only in July 2009, which is easy to use but lacks direct data access and IS tools like visualization and analysis.

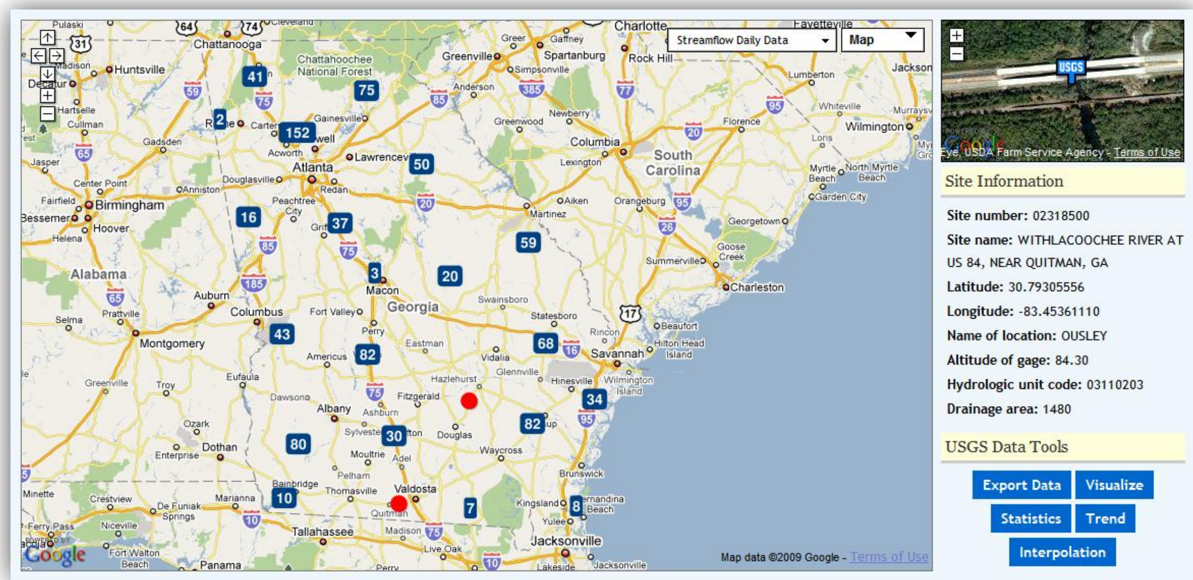


Figure 3.45. Interface of the USGS Interactive Information tool

Users can navigate through USGS monitoring sites using various types of interactive maps. USGS has more than thousands of monitoring sites for Georgia, but it is not easy to visualize thousands of markers on a map and navigate amongst them. The interface clusters monitoring sites in real time on the map to improve performance and user experience. Users can access site description, export, visualize and analyze data easily on the information panel. The interface uses a real-time clustering algorithm, providing new cluster centers for every zoom level. Numbers is blue boxes on the map show the number of observation sites in the cluster center, and red circles are individual observation sites. Clicking on the numbers zooms into the map to show smaller clusters or individual sites. Red circles load real-time information from USGS, such as a satellite view of the site, site details, and transfer the 60-day discharge values to GWIS for further analysis. The discharge data will be available after the transfer to download, visualize, and carry out statistical analysis without any pre-processing of the data.

3.4.6. Digital Resources Component

The digital resources component contains tools for user-posted digital resources, observation site descriptions generated by GWIS users, integration of USGS site description, a web service for GWIS site description information, meta-data generator tool, and a Wikipedia retriever.

GWIS Digital Library: The GWIS digital library (Figure 3.46) provides user-posted media and information such as articles, dissertations, modeling results, reports, meta-data, web-links, images, maps, audios, videos and other media types linked to related monitoring sites. Users can submit their digital resources by using a submission form (Figure 3.47). All of these contents can be downloaded by users in various file formats.

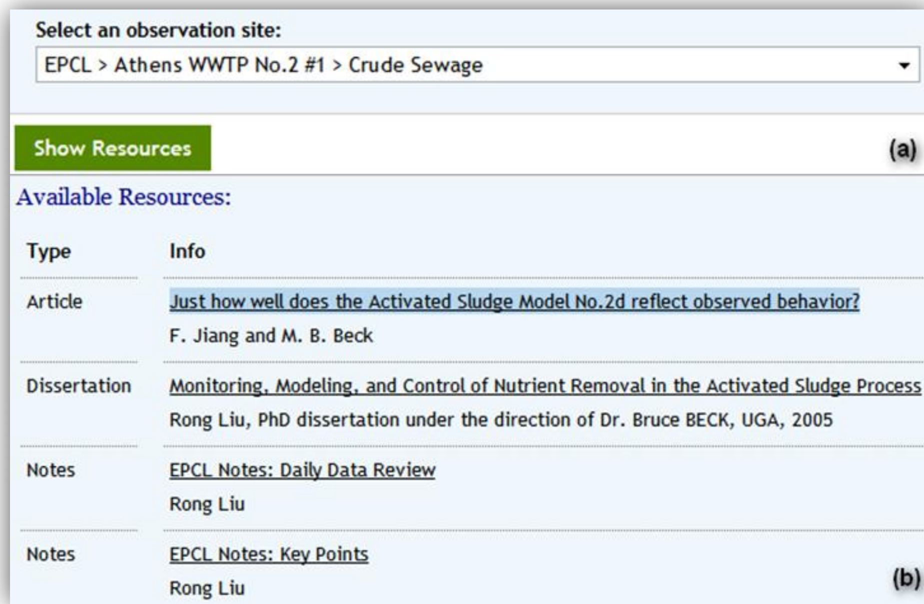


Figure 3.46. Interface of GWIS Digital Library, (a) selection, (b) results

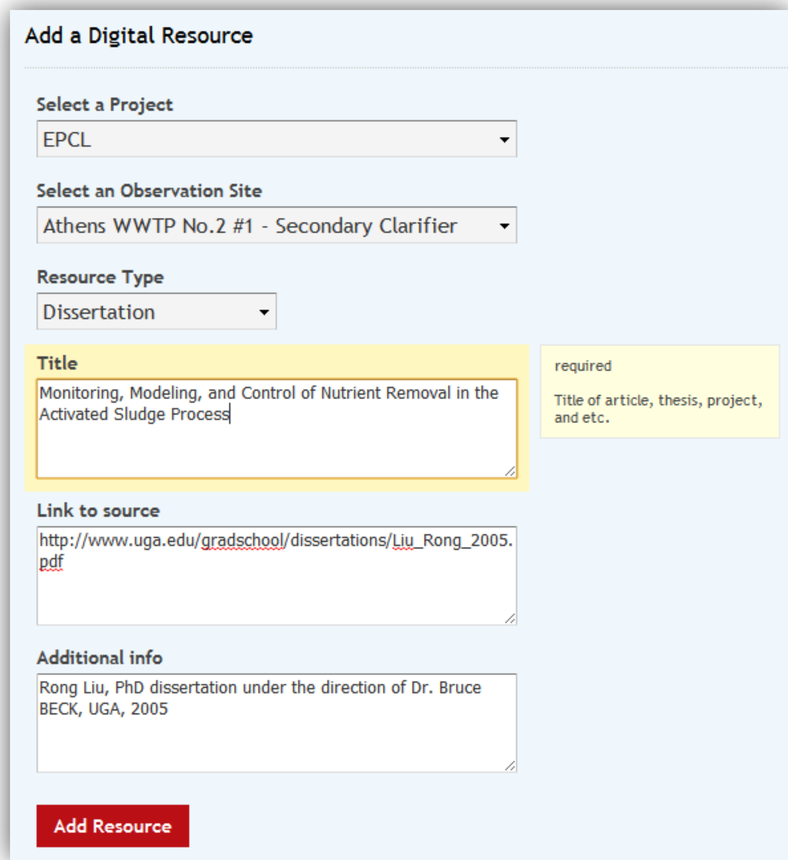


Figure 3.47. Submission form for digital resources

GWIS Site Description Tool: This tool (Figure 3.48 - Figure 3.49) provides a detailed site description provided by GWIS users with the physical location of the site. Users are also provided with links to outside sources related with the HUC of the site (EPA Sub-basin Info, USGS Real-time Streamflow, USGS Daily Streamflow).

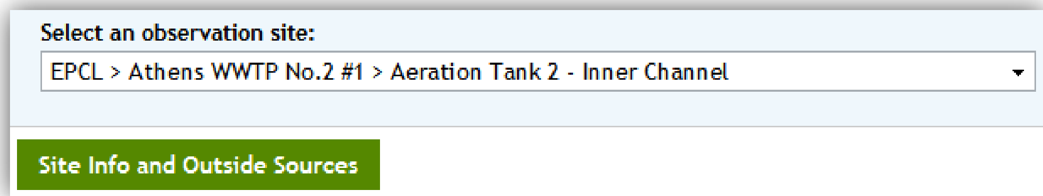


Figure 3.48. Selection interface of GWIS Site Description Tool

USGS Site Description Tool: Detailed site descriptions of the 370,933 USGS monitoring sites is accessible by using the USGS site number value through the GWIS interface (Figure 3.50 - Figure 3.51).

Site Description Web Service: GWIS features a *Site Description Web Service* to share user-provided site information through a web service, which will make it accessible through any supporting platform or software. Users can generate a web service URL by using the interface (Figure 3.52) and accessing a detailed site description in an XML file format (Figure 3.53).

Wikipedia Retriever Tool: GWIS provides a tool (Figure 3.54) to retrieve information (Figure 3.55) from Wikipedia by using a predefined list of terms and keywords. It can be used as a learning tool to build a list of keywords or terms for a course or special topic as a package of learning material. The content will be automatically retrieved from Wikipedia with texts, images and links.

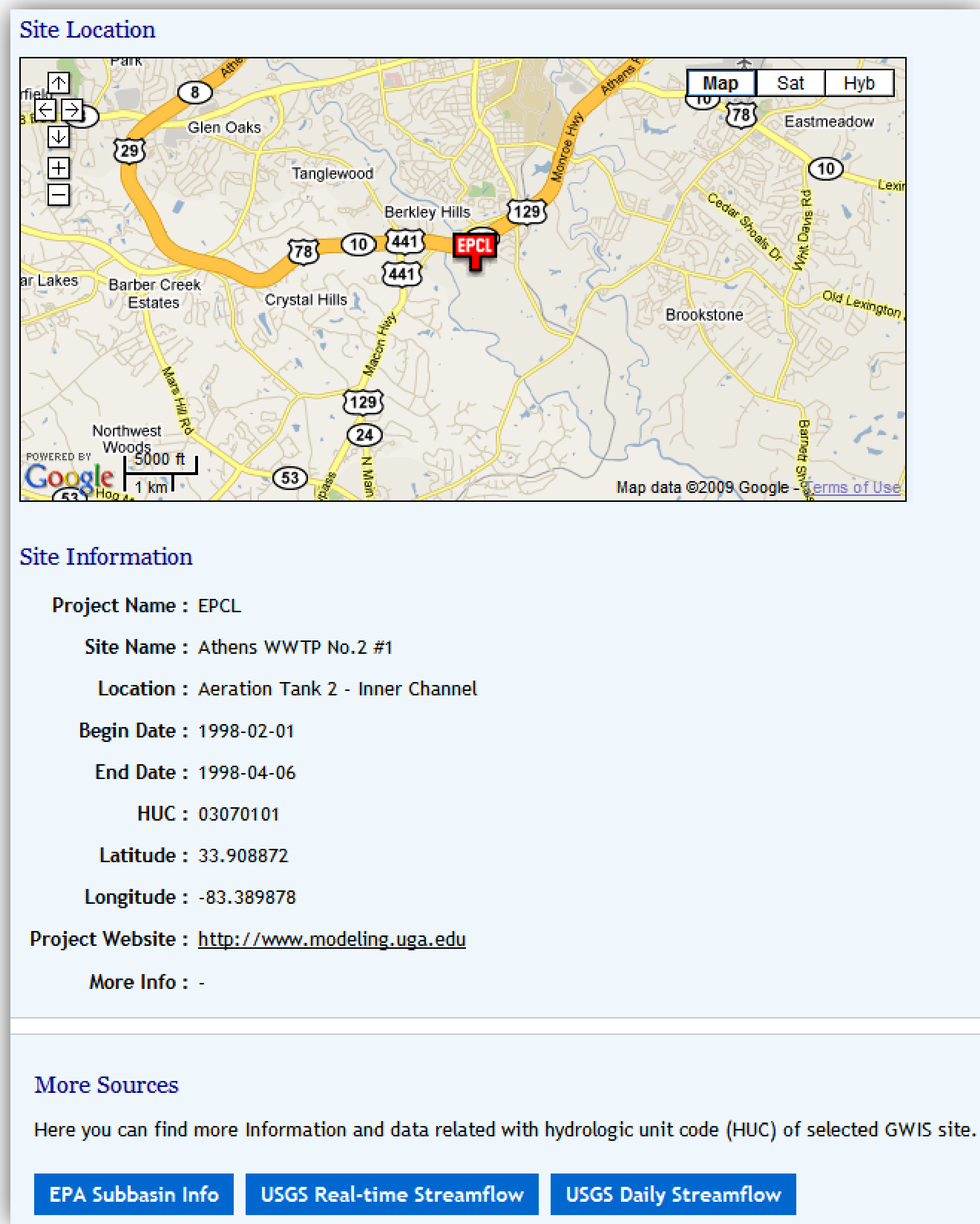


Figure 3.49. Results page of GWIS Site Description Tool

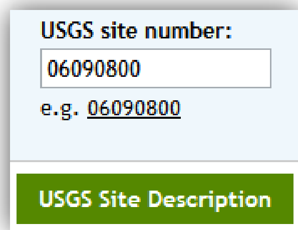


Figure 3.50. USGS site number input for USGS Site Description Tool

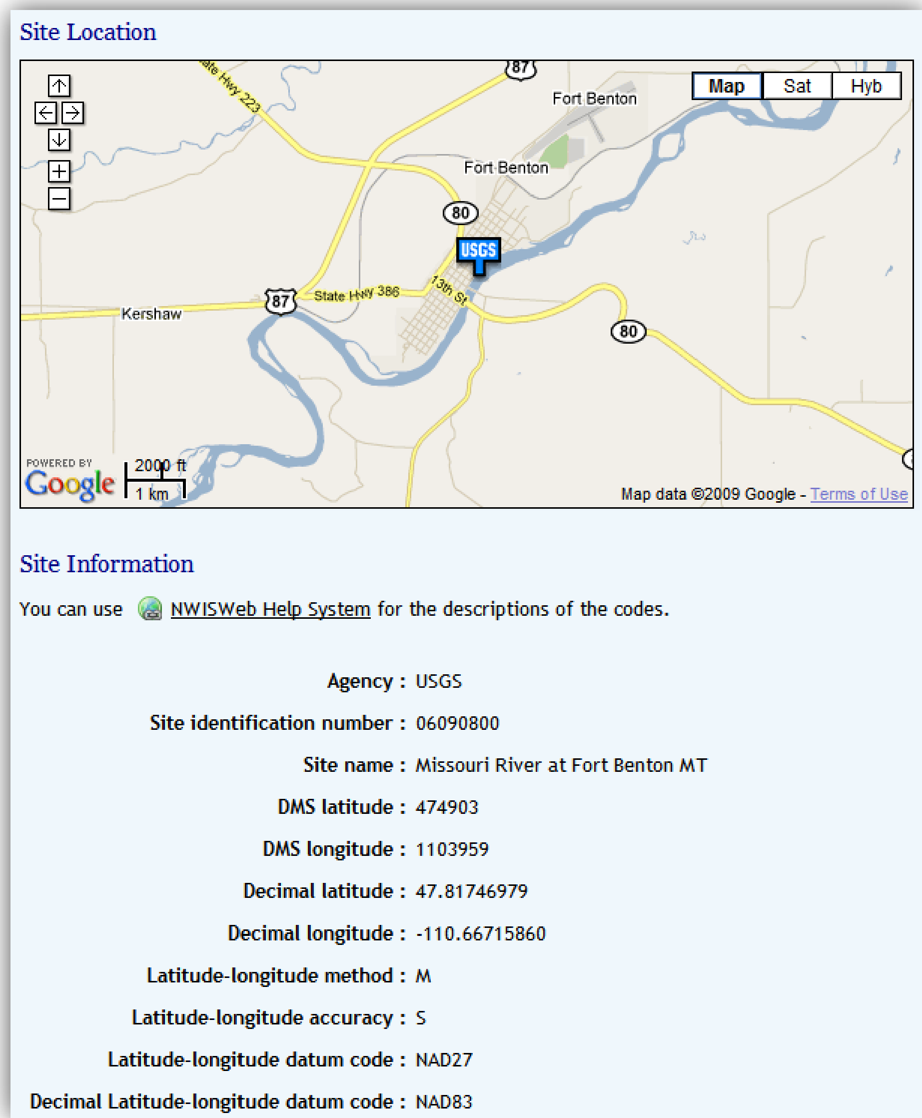


Figure 3.51. Partial results from USGS Site Description Tool

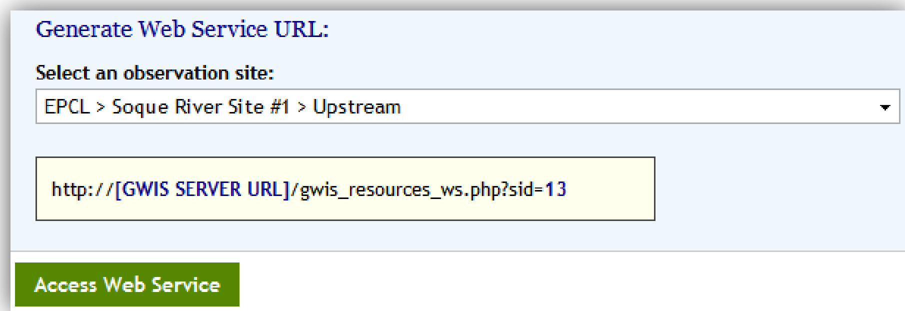
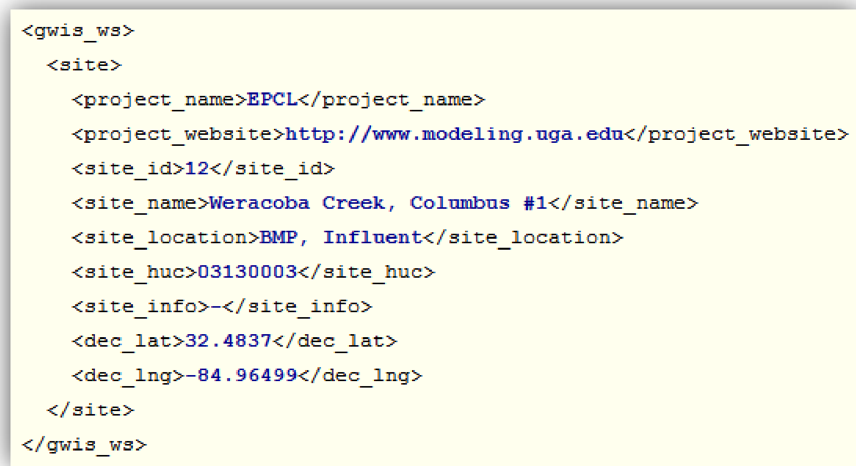


Figure 3.52. Interface of Site Description Web Service



```xml
<gwis_ws>
  <site>
    <project_name>EPCL</project_name>
    <project_website>http://www.modeling.uga.edu</project_website>
    <site_id>12</site_id>
    <site_name>Weracoba Creek, Columbus #1</site_name>
    <site_location>BMP, Influent</site_location>
    <site_huc>03130003</site_huc>
    <site_info>-</site_info>
    <dec_lat>32.4837</dec_lat>
    <dec_lng>-84.96499</dec_lng>
  </site>
</gwis_ws>
```


Figure 3.53. Sample XML output from Site Description Web Service

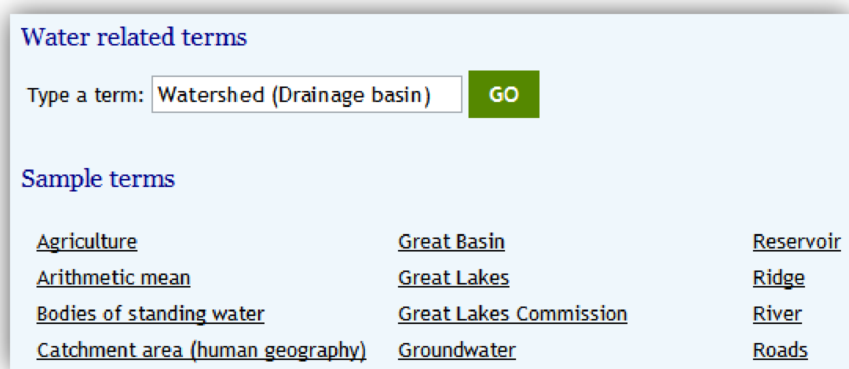
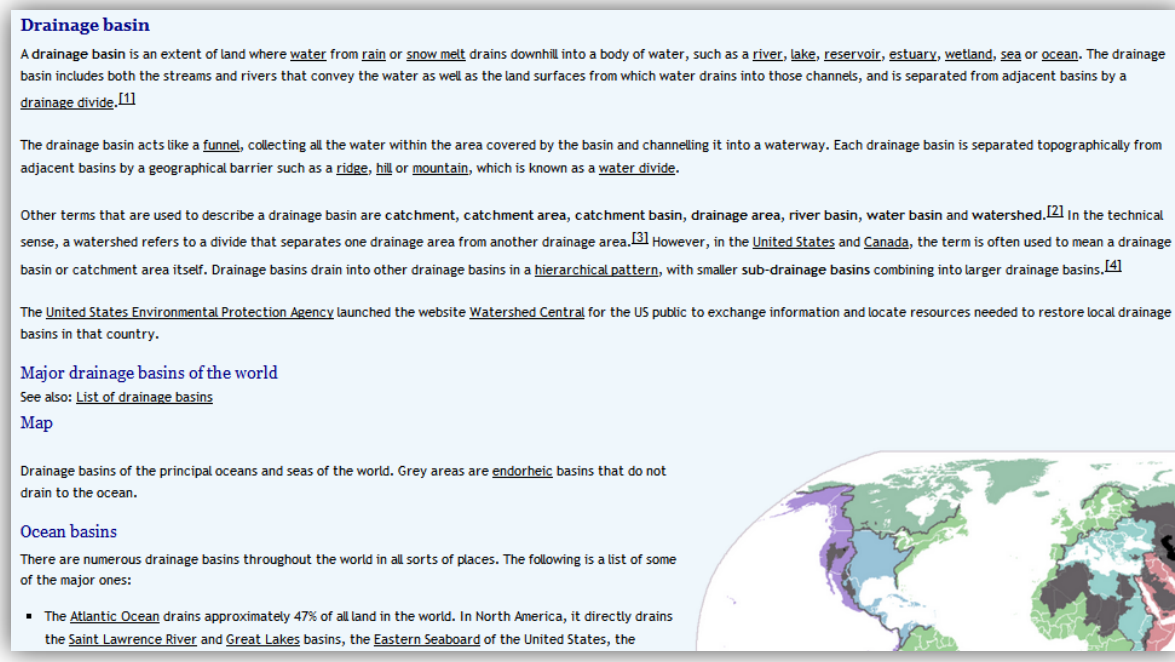


Figure 3.54. Input interface of Wikipedia Retriever Tool



Drainage basin

A **drainage basin** is an extent of land where water from rain or snow melt drains downhill into a body of water, such as a river, lake, reservoir, estuary, wetland, sea or ocean. The drainage basin includes both the streams and rivers that convey the water as well as the land surfaces from which water drains into those channels, and is separated from adjacent basins by a drainage divide.[1]

The drainage basin acts like a funnel, collecting all the water within the area covered by the basin and channelling it into a waterway. Each drainage basin is separated topographically from adjacent basins by a geographical barrier such as a ridge, hill or mountain, which is known as a water divide.

Other terms that are used to describe a drainage basin are **catchment, catchment area, catchment basin, drainage area, river basin, water basin** and **watershed.**[2] In the technical sense, a watershed refers to a divide that separates one drainage area from another drainage area.[3] However, in the United States and Canada, the term is often used to mean a drainage basin or catchment area itself. Drainage basins drain into other drainage basins in a hierarchical pattern, with smaller **sub-drainage basins** combining into larger drainage basins.[4]

The United States Environmental Protection Agency launched the website Watershed Central for the US public to exchange information and locate resources needed to restore local drainage basins in that country.

Major drainage basins of the world

See also: List of drainage basins

Map

Drainage basins of the principal oceans and seas of the world. Grey areas are endorheic basins that do not drain to the ocean.

Ocean basins

There are numerous drainage basins throughout the world in all sorts of places. The following is a list of some of the major ones:

- The Atlantic Ocean drains approximately 47% of all land in the world. In North America, it directly drains the Saint Lawrence River and Great Lakes basins, the Eastern Seaboard of the United States, the

Figure 3.55. Resulting page of Wikipedia Retriever Tool

Meta-data Generator Tool: GWIS has a meta-data generator (Figure 3.56) to create a site-specific file including information from monitoring campaigns, which can be visualized on graphs. Meta-data are very important to understand the behavior of the basic data. The tool allows users to create, modify, view and export meta-data by using a web interface. A meta-data file consists of a date, time and description (Figure 3.56) of each event that occurred during the sampling process. GWIS allows users to create meta-data files for new and already available monitoring campaigns. Already available meta-data files can be modified by users to correct or extend the information. Users need to provide exact date and time of information (Figure 3.57) to create or modify meta-data files. GWIS provides calendar controls (Figure 3.57) to help users enter date input easily.

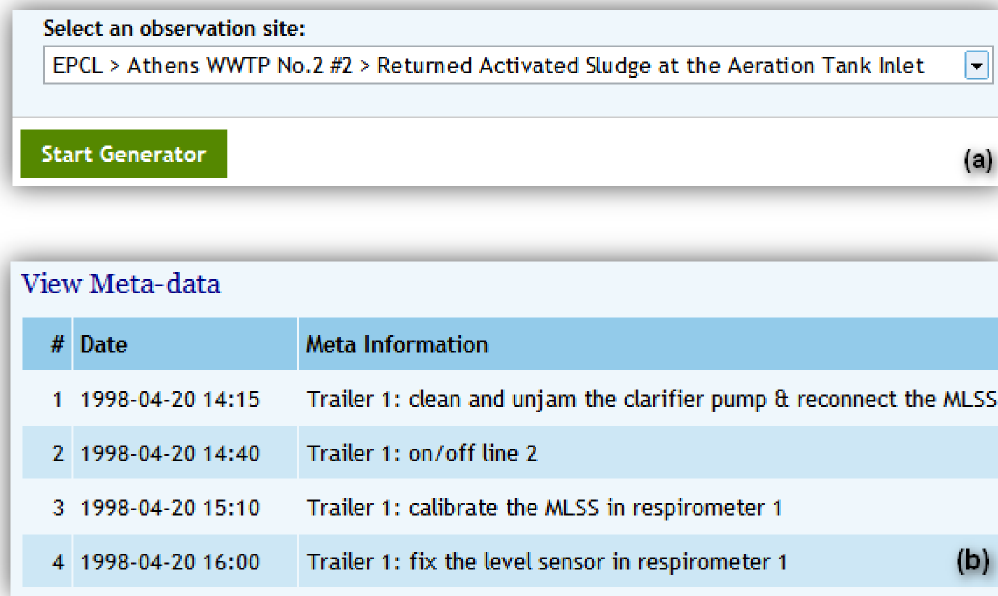


Figure 3.56. (a) Selection and (b) viewing interface of Meta-data Generator Tool

Figure 3.57. Creating or modifying an event on the interface

3.5. Comparison of GWIS

Contemporary watershed and water related information systems (for 15 sites) are summarized with respect to the technical capabilities in Table 3.3. The architectures of some of the advanced web-based environmental information systems (Figure 3.58) are visualized in parallel with that of GWIS (Figure 3.1) for easy comparison. Dark sections on the figures show that the EIS does not have the corresponding components or features. 'Limited' labels mean that capabilities are limited for the labeled components.

The Arkansas Watershed Information System (Figure 3.58a) provides limited capabilities for mapping and digital resources management components. The CUAHSI Hydrologic Information System (Figure 3.58b) has powerful data management component but provides limited mapping capabilities. The Bear River Watershed Information System (Figure 3.58c) has basic visualization tools with limited data management, mapping and digital resources management capabilities. WebH2O (Figure 3.58d), a commercial Environmental Information System, provides some analysis and visualization tools according to the client's needs, but with limited data management and mapping capabilities. As an overall evaluation of the figure, current EISs either have limited or no capabilities for some of the components when compared to GWIS. The reason for this difference is either that these EISs have a focus only on specific components, or that they have different research and development objectives. GWIS, as a prototype of a comprehensive Information System, tries to enable from basic to advanced capabilities for each component on a web-based user interface. Also these systems (other than GWIS) have limitations on usability, interoperability, linking and querying capabilities.

Table 3.3. Summary of technical capabilities of contemporary information systems

Technical Features / Information System and Focus	Data Management	Analysis Capabilities	Modeling	Scientific Visualization	Map Services Integration	Digital Resources	User Management	Web Services Integration	UI Design and Instructions
Arkansas WIS [1] *watershed digital resources*					*limited*	*yes*			*low*
CUAHSI HIS [2] *hydrological information*	*yes*			*limited*	*limited*		*yes*	*yes*	*middle*
Bear River WIS [3] *watershed data and information*	*limited*			*limited*	*limited*	*yes*	*yes*	*limited*	*middle*
WebH$_2$O EIS [4] *environmental information*	*yes*	*limited*	*limited*	*limited*	*limited*	*yes*	*yes*	*limited*	*high*
Willapa WIS [5] *watershed data and maps CD*	*limited*					*yes*			*low*
Tualatin River WIS [6] *watershed data and maps CD*	*limited*					*yes*			*low*
Montana Water IS [7] *water related data*	*yes*				*limited*	*yes*	*yes*		*middle*
Mattole WIS [8] *watershed maps and documents*						*limited*			*low*
San Gregorio WIS [9] *watershed maps and documents*						*limited*			*low*
EPA Waters WIN [10] *watershed information network*	*yes*				*yes*	*yes*			*middle*
Russian River IS [11] *watershed digital resources*	*limited*					*yes*	*yes*		*middle*
Gulf of Maine WIS [12] *watershed data and information*	*limited*				*limited*	*yes*			*low*
Alabama Water IS [13] *watershed data and information*	*limited*				*limited*	*yes*			*low*
Klamath Resource IS [14] *watershed data and information*	*limited*					*yes*			*low*
GWIS *watershed information system*	*yes*	*yes*	*yes*	*yes*	*yes*	*yes*	*yes*	*yes*	*high*

[1] watersheds.cast.uark.edu
[2] his.cuahsi.org
[3] www.bearriverinfo.org
[4] info.webh2o.net
[5] www.inforain.org/willapa
[6] www.inforain.org/tualatin
[7] nris.mt.gov/wi.asp
[8] www.mattole.org
[9] sgreg.stillwatersci.com
[10] www.epa.gov/waters
[11] www.russianriverwatershed.net
[12] www.gm-wics.sr.unh.edu
[13] www.aces.edu/waterquality
[14] www.krisweb.com

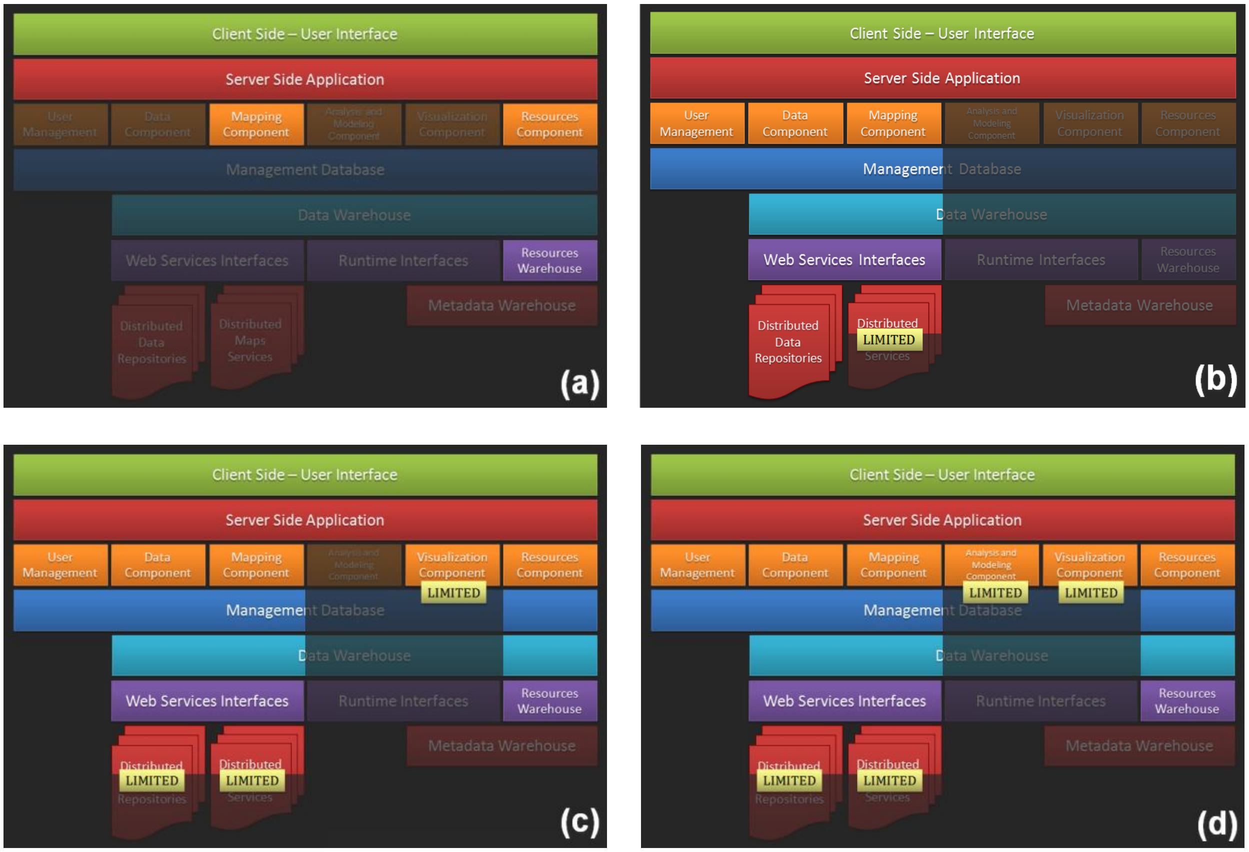


Figure 3.58. Architectures of some modern web-based EIS.

(a) Arkansas Watershed Information System, (b) CUAHSI Hydrologic Information System, (c) Bear River Watershed Information System, (d) WebH$_2$O Environmental Information System.

3.6. Conclusion

All of the tools and features of the six main components of GWIS have been introduced and explained with reference to related methods and literature, and screenshots of the interfaces and sample results are given in this chapter. While providing extensive functionality in these selected components, contemporary information systems provide limited or no capabilities for other components. With over 30 information system features, GWIS, as a prototype, provides basic-to-advanced tools for each component. As a comprehensive information system, having core components implemented in the system, GWIS provides users with an integrated environment to create novel visualization and information interfaces (Figure 3.31 and Figure 3.45). The next chapter provides mini case studies for each tool with real water quality data and presents the resulting graphics and reports.

CHAPTER 4

GWIS: HOW TO USE

4.1. Introduction

This chapter illustrates mini examples with real water quality data for each tool and presents results as graphics or reports. It also includes instructions for tools provided by all the GWIS components, a user and content management section, new features suggestions and special design features of GWIS. A novel help tool, GWIZARD, will be presented at the end of the chapter.

4.2. GWIS User Interface

The GWIS user interface (Figure 4.1) consists of four main sections. Whereas the header section contains the GWIS logo and navigation menu, the footer section contains program information and quick links to the main GWIS sections. The middle section is divided into two parts: the left panel contains GWIZARD, the GWIS description, and the featured project; and right panel is reserved for the main content.

The navigation menu has two levels, separated for site functions and components. The site function level has quick links to the home page, the GWIS summary page, contact information, site navigation guide (GWizard), login and registration forms. Some of the site function page can be accessed through the footer section. The component navigation level has links to the list of the site and six main component pages (data management, analysis, modeling, visualization, mapping and digital resources management).

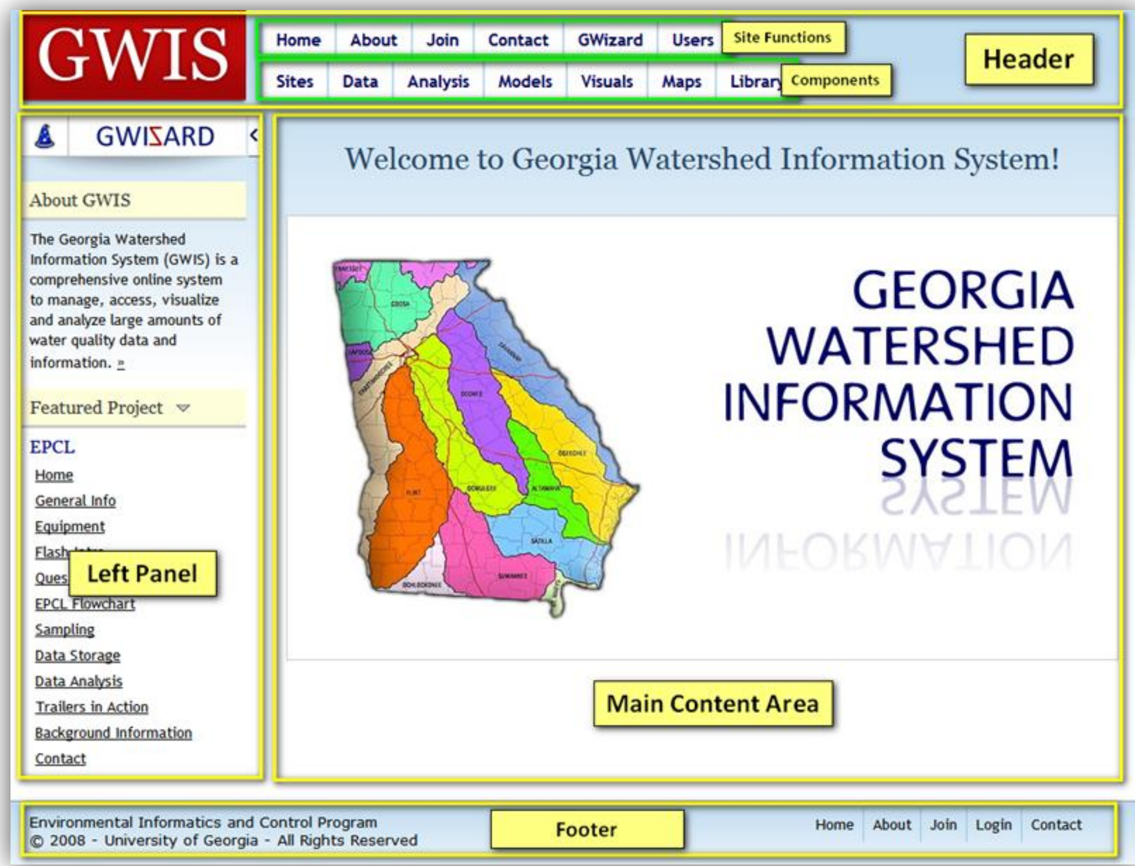


Figure 4.1. Interface of GWIS with sections outlined

4.3. GWIS Components

GWIS has six components with 27 tools and features. Each component provides the necessary background information, with references about the methods used in the tools, and instructions to fully utilize the system and help users to understand how they will understand the results. This section provides mini case studies for each tool as a scenario and instructions to complete this scenario. In most of the scenarios, the Soque River monitoring campaign data are used either from the upstream location, or both the upstream and downstream locations, where the analysis requires two separate observation sites.

4.3.1. Data Component

4.3.1.1. GWIS Data Sources Tool

Scenario: User wants to export all available time series to his computer for the dissolved oxygen variable of the Soque River upstream observation site

Instructions: Go to 'Data' component and then 'GWIS Data Sources' section

Click 'Start' button

Select 'Soque River #1 > Upstream' from the list (Figure 4.2)

Select 'Dissolved Oxygen' variable

Click 'Export Variable Data' button

Save the file to your computer

Results: Time series data are downloaded to the computer and can be used for further analysis and visualization. File format is CSV, which can be opened by using Microsoft Excel software. The first column in the file is date and time values for each sample, and the second column is the variable concentration in given units.

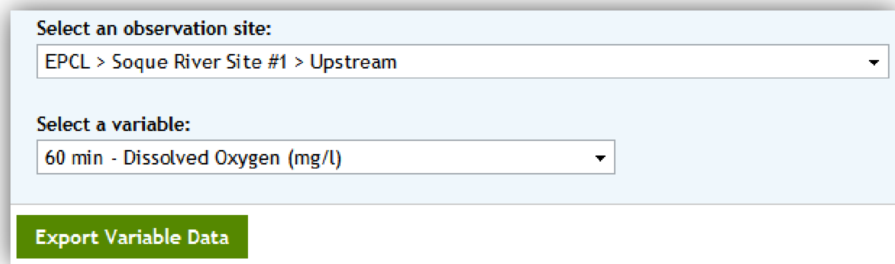


Figure 4.2. GWIS data source tool screenshot

4.3.1.2. GWIS Data Web Service Tool

Scenario #1: User wants to generate a web service URL to access dissolved oxygen data for the upstream Soque River monitoring campaign and use the time series directly within his data access software or model

Instructions: Go to 'Data' component and then 'GWIS Data Web Service' section

Click 'Start' button

Select 'Soque River #1 > Upstream' from the list (Figure 4.3)

Select 'Dissolved Oxygen' variable

Copy the automatically generated URL from the yellow box and paste it into your application

Results: Time series data will be transferred directly to the software or model, sending a request over the web service in XML file format (Figure 4.4). The XML file contains basic information about the observation site and variable, and concentrations of the variable in the time series.

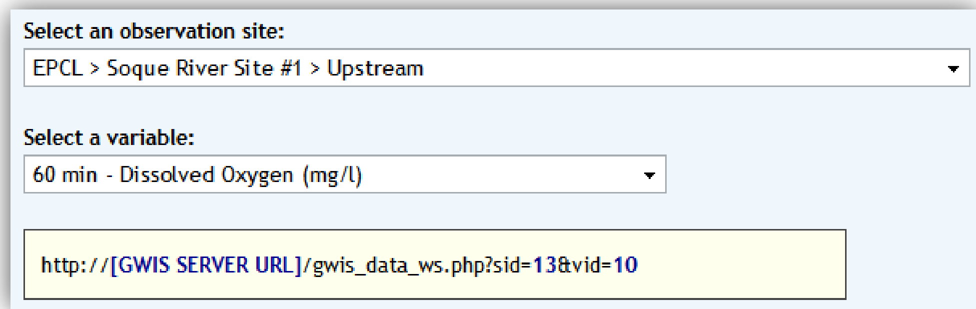


Figure 4.3. Generating a web service url


```xml
<gwis_ws>
  <site>
    <project_name>EPCL</project_name>
    <project_website>http://www.modeling.uga.edu</project_website>
    <site_id>12</site_id>
    <site_name>Weracoba Creek, Columbus #1</site_name>
    <site_location>BMP, Influent</site_location>
    <site_huc>03130003</site_huc>
    <site_info>-</site_info>
    <dec_lat>32.4837</dec_lat>
    <dec_lng>-84.96499</dec_lng>
  </site>
  <variable>
    <var_id>23</variable_id>
    <var_name>Dissolved Oxygen</var_name>
    <var_unit>mg/l</var_unit>
    <sampling_frequency>15 min</sampling_frequency>
    <begin_date>2007-11-01</begin_date>
    <end_date>2008-02-29</end_date>
    <time_zone offset="-05:00">EST</time_zone>
  </variable>
  <values>
    <value dateTime="1998-10-05 10:15">20.29</value>
    <value dateTime="1998-10-05 10:30">30.29</value>
    <value dateTime="1998-10-05 10:45">40.29</value>
        ...
  </values>
</gwis_ws>
```


Figure 4.4. Sample XML file format

Scenario #2: User wants to access GWIS web service to download dissolved oxygen data for the upstream Soque River Basin in XML format

Instructions: Go to 'Data' component and then 'GWIS Data Web Service' section

Click 'Start' button

Select 'Soque River #1 > Upstream' from the list (Figure 4.5)

Select 'Dissolved Oxygen' variable

Click 'Access Web Service' button

Save the file to your computer

Results: Time series data are downloaded to the computer and can be used for further analysis and visualization. File format is XML, which can be opened by using a basic text editor or accessed by a model or software. The XML file contains basic information about the observation site and variable, and concentrations of the variable in the time series.

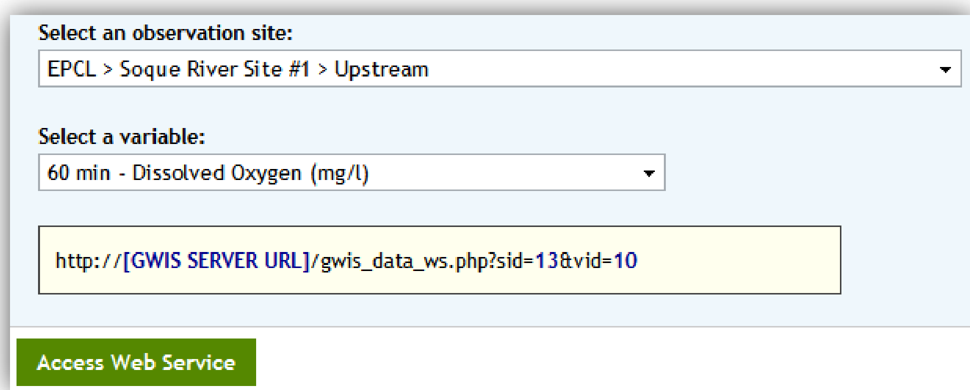


Figure 4.5. Accessing web service directly within GWIS

4.3.1.3. USGS Real-time Data Tool

Scenario #1: User wants to download real-time discharge data for last 60 days to his computer for the observation site Withlacoochee River, Quitman, GA, '02318500' (site number) from USGS servers

Instructions: Go to 'Data' component and then 'USGS Real-time Data' section

Click 'Start' button

Enter site number '02318500' to the input box (Figure 4.6)

Enter parameter code '00060' for discharge to the input box

Select '60' days from selection box

Click 'Export XML' button and save the file to your computer

Results: Time series data are downloaded to computer and can be used for further analysis and visualization. File format is XLS, which can be opened by using a basic text editor or Microsoft Excel. The data file contains basic information about the observation site and variable, and concentrations of the variable in the time series.

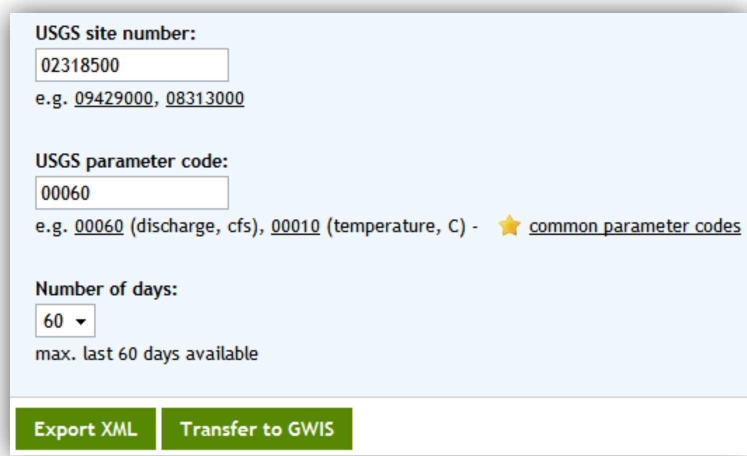


Figure 4.6. USGS real-time data integration tool

Scenario #2: User wants to transfer real-time discharge data for the last 60 days to the GWIS server for the observation site Withlacoochee River, Quitman, GA, ‘02318500’ (site number) from USGS servers, and utilize GWIS tools for visualization and analysis of the data set

Instructions: Go to ‘Data’ component and then ‘USGS Real-time Data’ section

Click ‘Start’ button

Enter site number ‘02318500’ to the input box (Figure 4.6)

Enter parameter code ‘00060’ for discharge to the input box

Select ‘60’ days from selection box

Click ‘Transfer to GWIS’ button

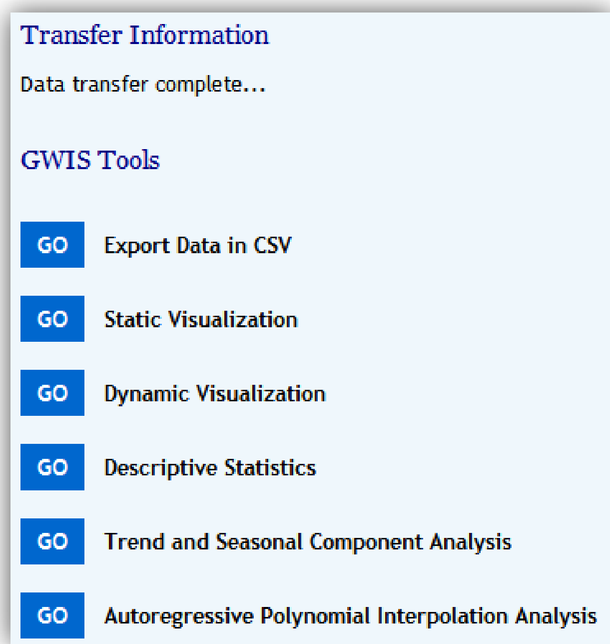


Figure 4.7. Data are transferred to GWIS for further analysis and visualization

Results: Data are downloaded to the GWIS servers. The following screen provides information about the observation site and the tools for utilizing data (Figure 4.7)

Click any of the 'GO' buttons for your analysis and visualization preference

4.3.1.4. USGS Daily Data Tool

Scenario #1: User wants to download daily discharge data for July 2008 to his computer for observation site Withlacoochee River, Quitman, GA, '02318500' (site number) from USGS servers

Instructions: Go to 'Data' component and then 'USGS Daily Data' section

Click 'Start' button

Enter site number '02318500' to the input box (Figure 4.8)

Enter parameter code '00060' for discharge to the input box

Enter statistic code '00003' for mean values to the input box

Click the calendar icon to select July-01-2008 as 'Campaign start date', and July-31-2008 as 'Campaign end date'

Click 'Export XML' button and save the file to your computer

Results: Time series data are downloaded to the computer and can be used for further analysis and visualization. File format is XLS, which can be opened by using a basic text editor or Microsoft Excel. The data file contains basic information about the observation site and variable, and concentrations of the variable in the time series.

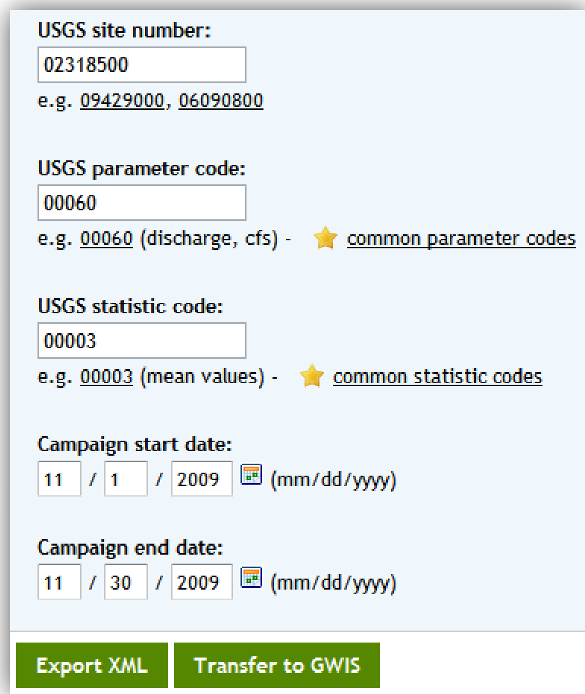


Figure 4.8. USGS daily data integration tool

Scenario #2: User wants to transfer daily discharge data for July 2008 to the GWIS server for the observation site Withlacoochee River, Quitman, GA, '02318500' (site number) from USGS servers, and utilize GWIS tools for visualization and analysis of the data set

Instructions: Go to 'Data' component and then 'USGS Daily Data' section

Click 'Start' button

Enter site number '02318500' to the input box (Figure 4.8)

Enter parameter code '00060' for discharge to the input box

Enter statistic code '00003' for mean values to the input box

Click the calendar icon to select July-01-2008 as 'Campaign start date', and July-31-2008 as 'Campaign end date'

Click 'Transfer to GWIS' button

Results: Data are downloaded to the GWIS servers. The following screen provides information about the observation site and tools for utilizing the data (Figure 4.7)

Click any of the 'GO' buttons for your analysis and visualization preference

4.3.1.5. Data Sources by HUC (Hydrologic Unit Code) Tool

Scenario: User wants to access EPA Sub-basin information for the Upper Oconee Watershed with HUC of '03070101'

Instructions: Go to 'Data' component and then 'Data Sources by HUC' section

Click 'Start' button

Enter HUC number '03070101' to the input box (Figure 4.9)

Click 'EPA Subbasin Info' button

Results: An information window will appear on the screen as a popup without leaving the page (Figure 4.10)

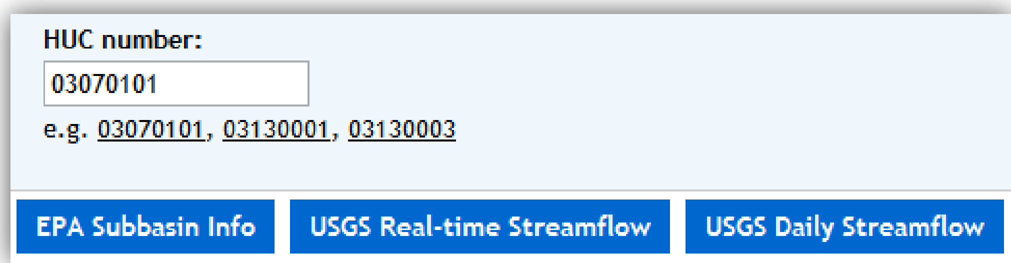


Figure 4.9. Interface of data sources by HUC tool

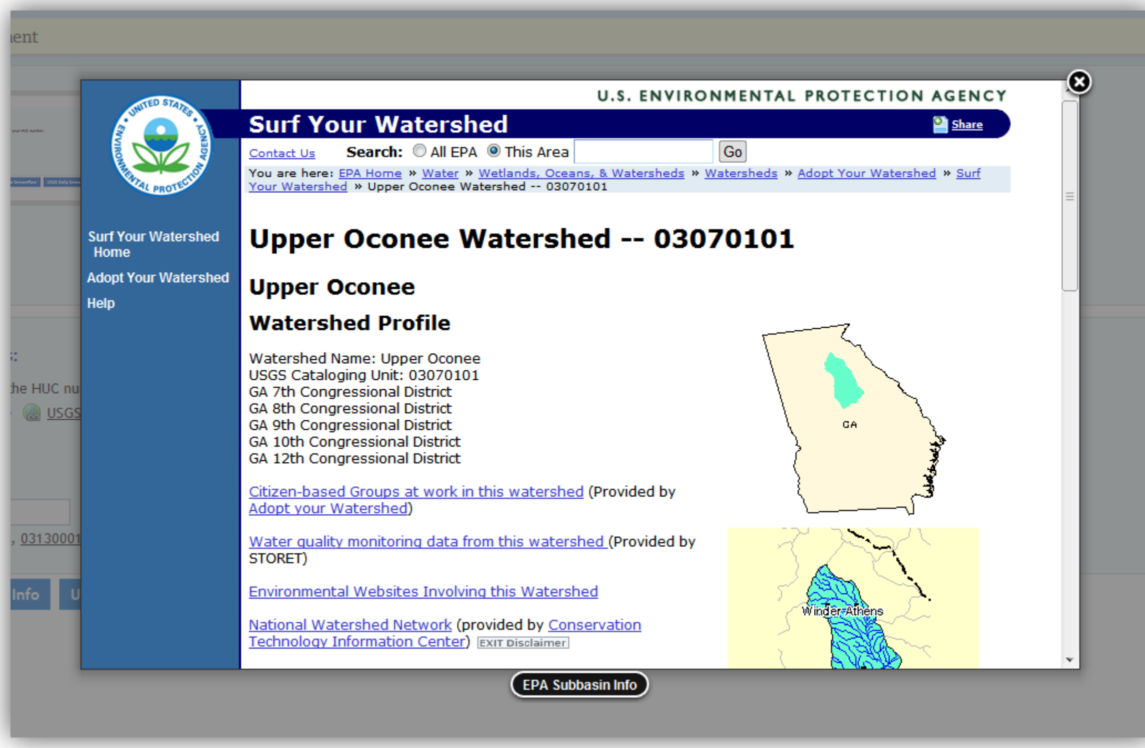


Figure 4.10. EPA Sub-basin information page opens without leaving GWIS

4.3.2. Analysis Component

4.3.2.1. Descriptive Statistics Tool

Scenario: User wants to generate descriptive statistics for various variables of an observation site (Soque River Site #1 - Upstream) uploaded to GWIS

Instructions: Go to ‘Analysis’ component and then ‘Descriptive Statistics’ section

Click ‘Start’ button

Select the observation site ‘Soque River Site #1 - Upstream’ from the selection box (Figure 4.11)

Select any number of variables from the selection box

Click ‘Run Analysis’ button

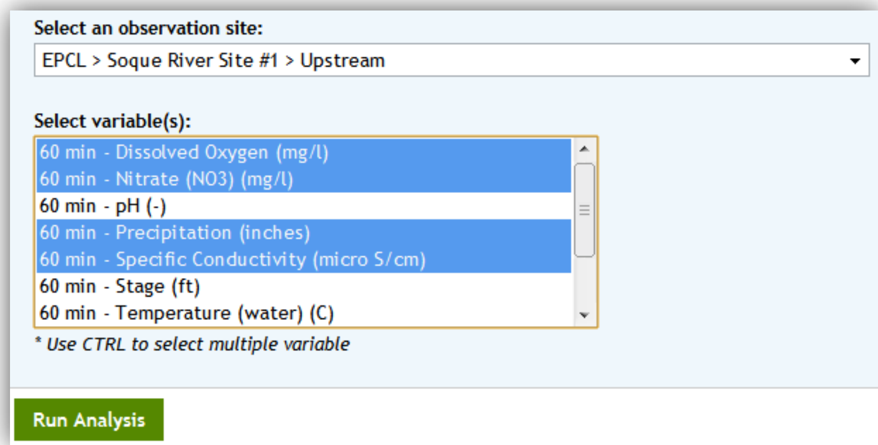


Figure 4.11. Descriptive statistics analysis tool

Results: The following screen provides basic information about the observation site, and a report will appear with descriptive statistics information (Figure 4.12) and a histogram plot (Figure 4.13) of the variables

Click the 'Print Report' button to print the report as a PDF file or on paper from any printer

Descriptive Statistics:

[Variable Name]	[Min]	[Mean]	[Max]	[Std]	[Missing]
Dissolved Oxygen (60 min) (mg/l)	0	7.44223	8.38	7.91965	150 (18%)
Nitrate (NO3) (60 min) (mg/l)	0.01	0.102197	1.11	0.332003	59 (7%)
Precipitation (60 min) (inches)	0	0.0995789	0.6	0.230516	747 (89%)
Specific Conductivity (60 min) (micro S/cm)	0	18.725	26.1	20.4791	97 (12%)

Figure 4.12. Statistical properties of the selected variables

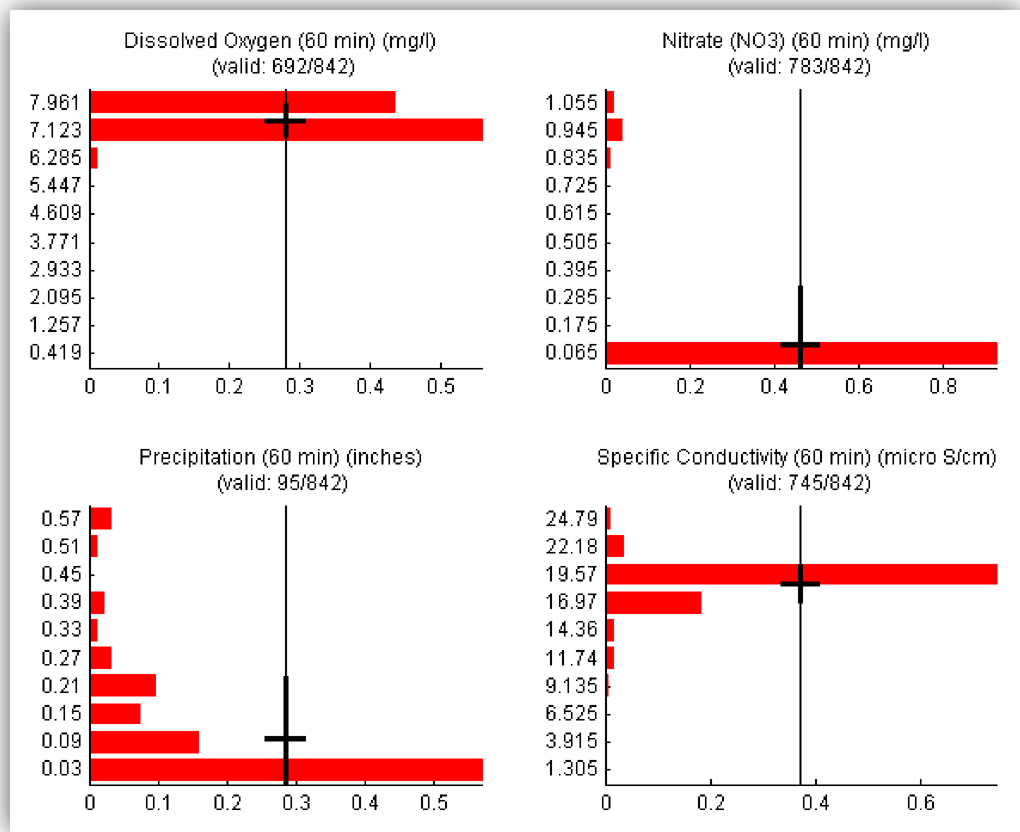


Figure 4.13. Histogram plot of the selected variables

4.3.2.2. Trend and Seasonal Component Analysis Tool

Scenario: User wants to generate trend and seasonal components of dissolved oxygen for an observation site (Soque River Site #1 - Upstream) uploaded to GWIS by using an auto-regressive model

Instructions: Go to 'Analysis' component and then 'Trend and Seasonal Component Analysis' section

Click 'Start' button

Select the observation site 'Soque River Site #1 - Upstream' from the selection box (Figure 4.14)

Select the variable 'Dissolved Oxygen' from the selection box

Click 'Start' button

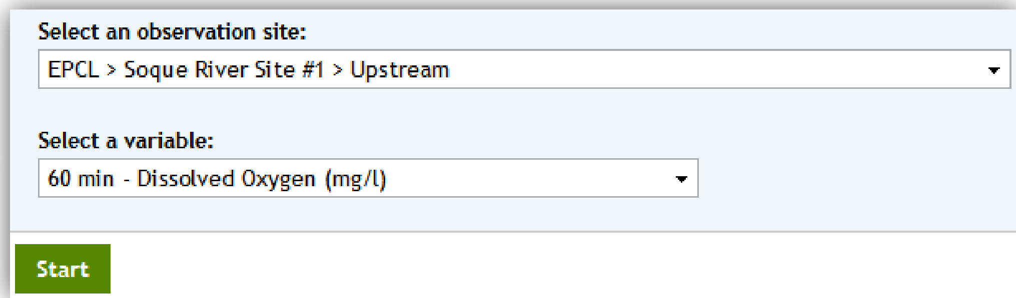


Figure 4.14. Trend and seasonal component analysis screen

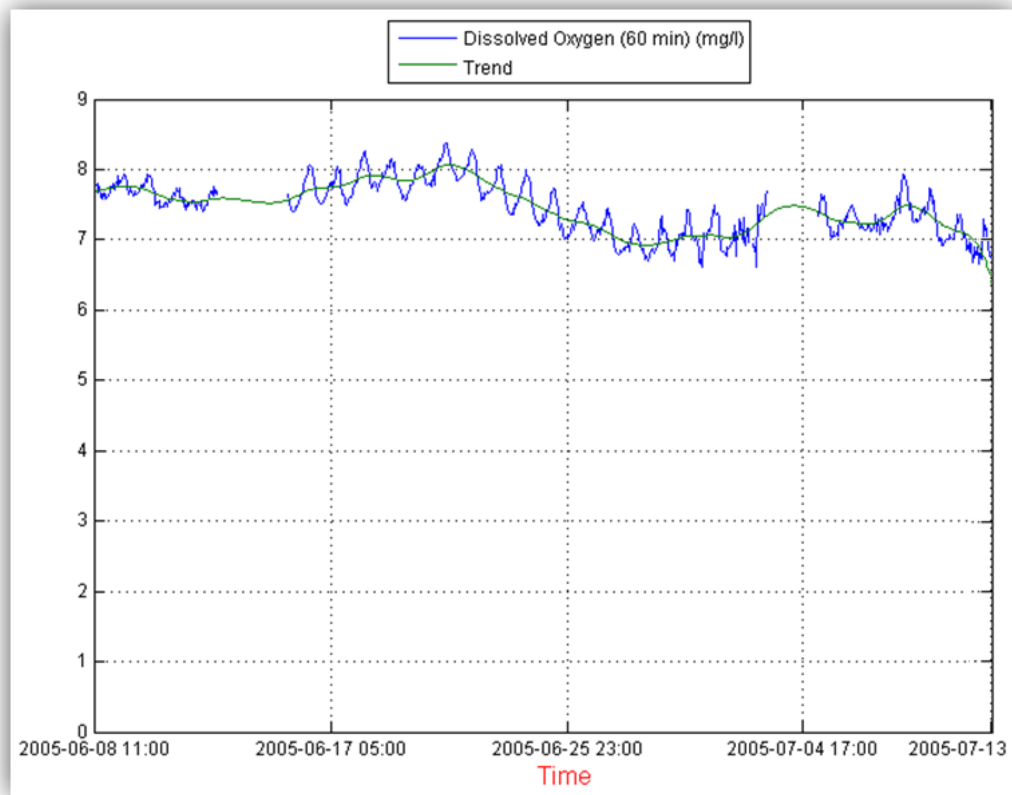


Figure 4.15. Plot for trend and actual values of dissolved oxygen

Results: The following screen provides basic information about the observation site, and a report will appear with the trend (Figure 4.15) and seasonal component (Figure 4.16) plots

Click the 'Download Data' link in the report to export the trend and seasonal component values as a CSV file to your computer for further analysis

Click the 'Print Report' button to print the report as a PDF file or on paper from any printer

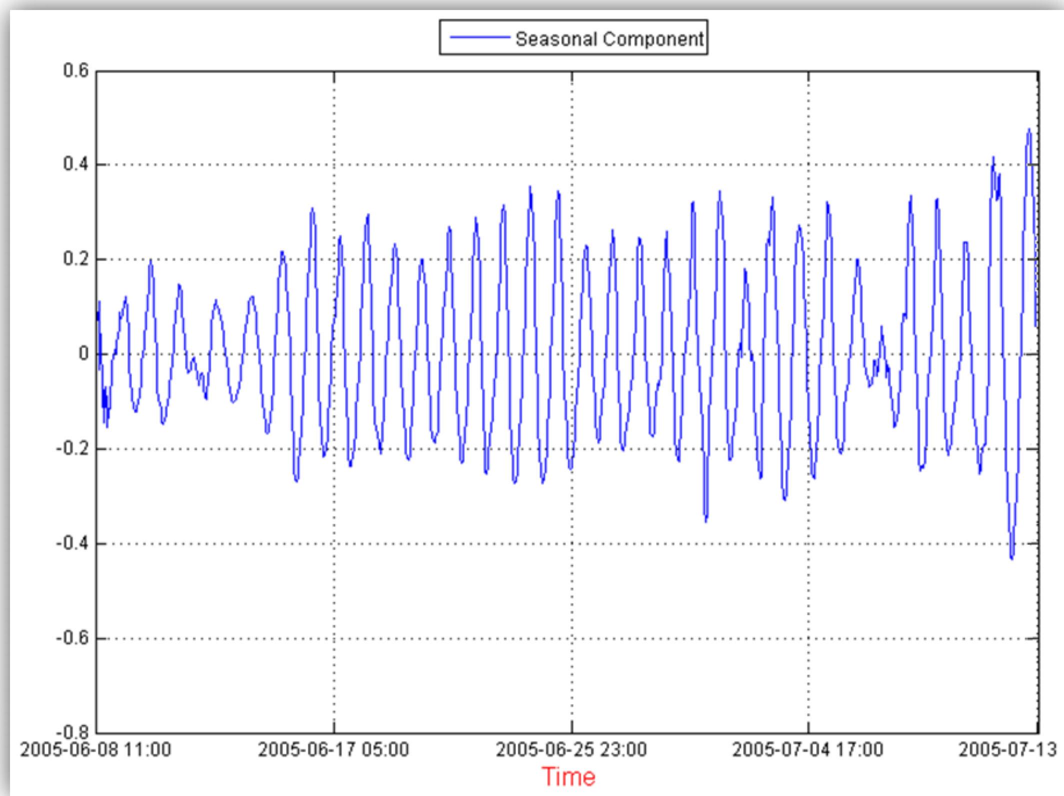


Figure 4.16. Plot for seasonal component values of dissolved oxygen

4.3.2.3. Autoregressive Polynomial Interpolation Analysis Tool

Scenario: User wants to carry out an interpolation analysis for the dissolved oxygen variable of an observation site (Soque River Site #1 - Upstream) uploaded to GWIS by using an auto-regressive model

Instructions: Go to the 'Analysis' component and then the 'Autoregressive Polynomial Interpolation Analysis' section

Click 'Start' button

Select the observation site 'Soque River Site #1 - Upstream' from the selection box (Figure 4.17)

Select the variable 'Dissolved Oxygen' from the selection box

Click 'Start' button

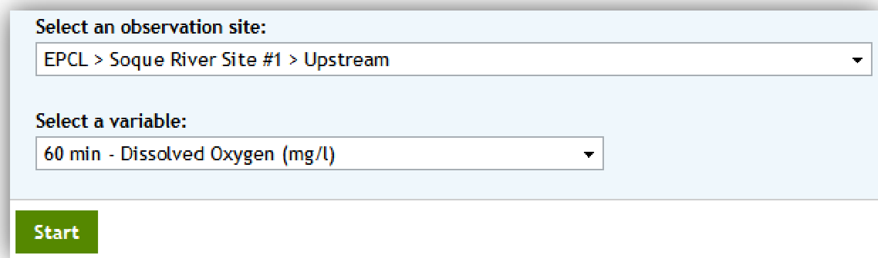


Figure 4.17. Autoregressive polynomial interpolation analysis interface

Results: The following screen provides basic information about the observation site, and a report will appear with interpolation (Figure 4.18) and residual plots

Click the 'Download Data' link in the report to export interpolated values as a CSV file to your computer for further analysis

Click the 'Print Report' button to print the report as a PDF file or on paper from any printer

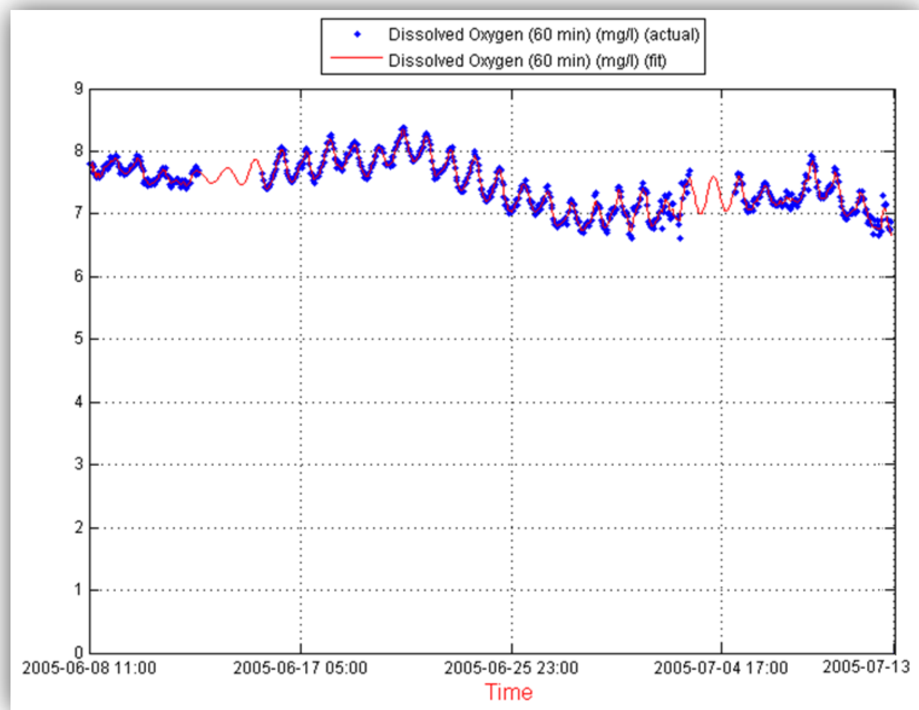


Figure 4.18. Actual and simulated values of dissolved oxygen

4.3.2.4. Correlation Analysis with Self Organizing Maps Tool

Scenario: User wants to carry out a correlation analysis for specific variables of an observation site (Soque River Site #1 - Upstream) uploaded to GWIS by using Self Organizing Maps

Instructions: Go to the 'Analysis' component and then the 'Correlation Analysis with Self Organizing Maps' section

Click 'Start' button

Select the observation site 'Soque River Site #1 - Upstream' from the selection box (Figure 4.19)

Select any number of variables from the selection box

Click 'Run Analysis' button

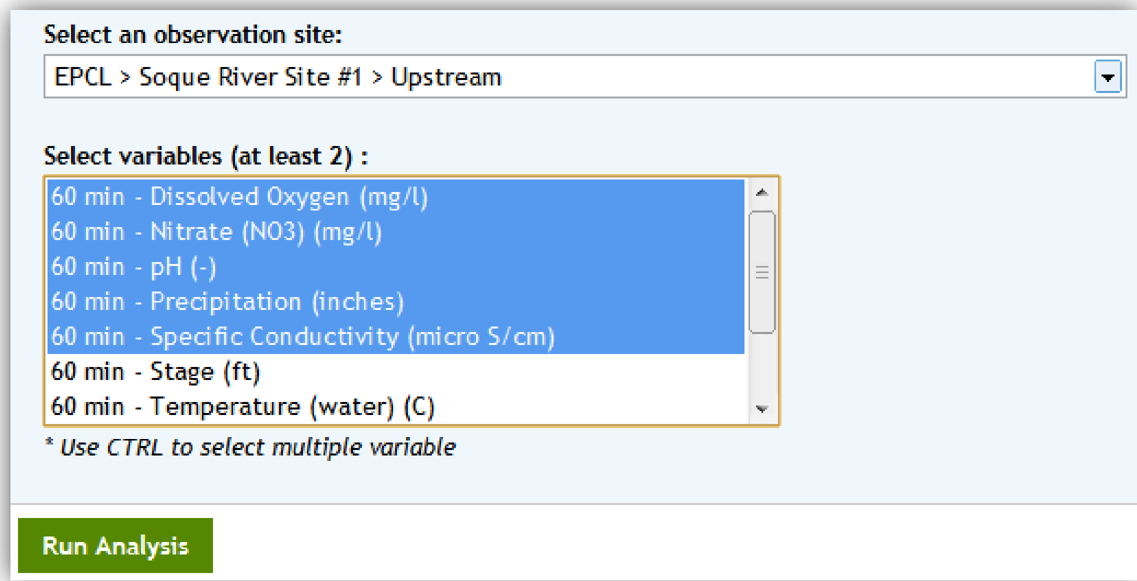
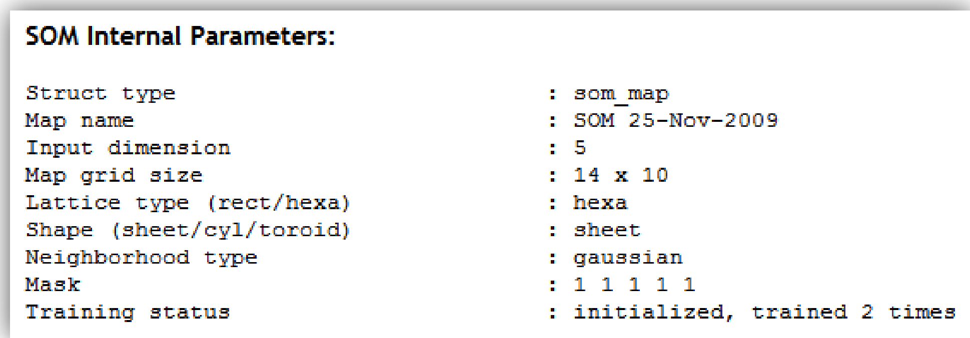


Figure 4.19. Correlation analysis with Self Organizing Maps tool

Results: The following screen provides basic information about the observation site, and a report will appear with descriptive statistics information (Figure 4.12), internal model parameters (Figure 4.20), and component maps plot (Figure 4.21) of the variables

Click the 'Print Report' button to print the report as a PDF file or on paper from any printer



SOM Internal Parameters:

Struct type	:	som_map
Map name	:	SOM 25-Nov-2009
Input dimension	:	5
Map grid size	:	14 x 10
Lattice type (rect/hexa)	:	hexa
Shape (sheet/cyl/toroid)	:	sheet
Neighborhood type	:	gaussian
Mask	:	1 1 1 1 1
Training status	:	initialized, trained 2 times

Figure 4.20. Internal model parameters for Self Organizing Maps

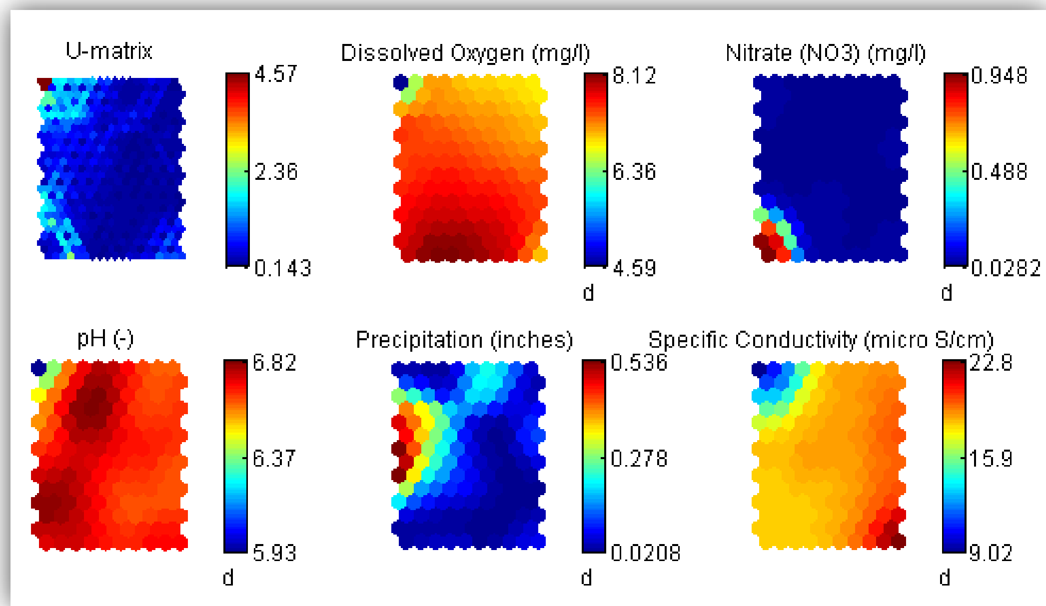


Figure 4.21. SOM component maps plot for selected variables

4.3.2.5. Missing Data Analysis with Self Organizing Maps Tool

Scenario: User wants to complete a data series (dissolved oxygen) with missing values for an observation site (Soque River Site #1 - Upstream) by using Self Organizing Maps

Instructions: Go to the 'Analysis' component and then 'Missing Data Analysis with Self Organizing Maps' section

Click 'Start' button

Select the observation site 'Soque River Site #1 - Upstream' and the variable with missing values as 'Dissolved Oxygen' from the selection box

Select any number of supplementary variables correlated to the variable with missing values from the selection box (Figure 4.22)

Click 'Run Analysis' button

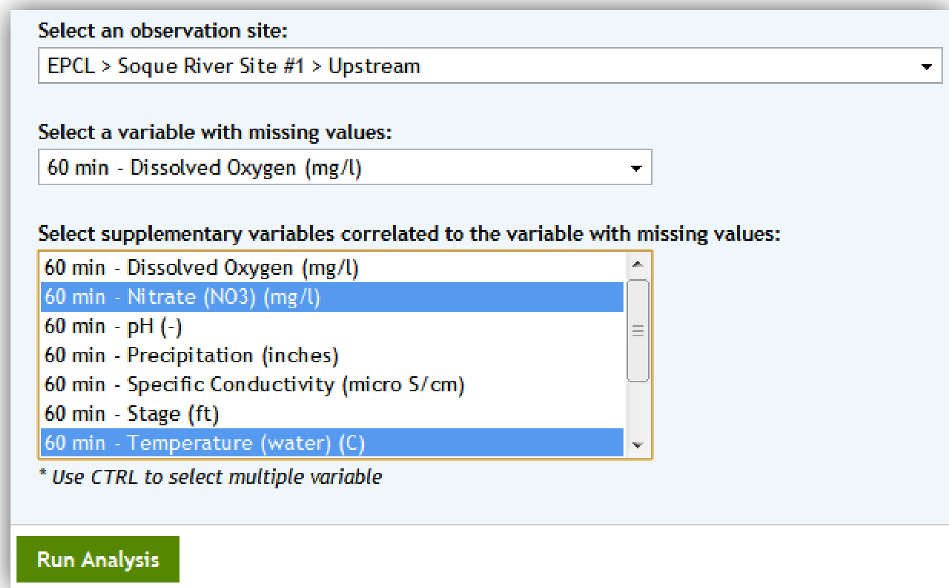


Figure 4.22. Missing data analysis with Self Organizing Maps tool interface

Results: The following screen provides basic information about the observation site, and a report will appear with the descriptive statistics information (Figure 4.12), scatter and line plots of missing and complete sets (Figure 4.23)

Click the 'Download Data' link in the report to export completed values of the time series as a CSV file to your computer for further analysis

Click the 'Print Report' button to print the report as a PDF file or on paper from any printer

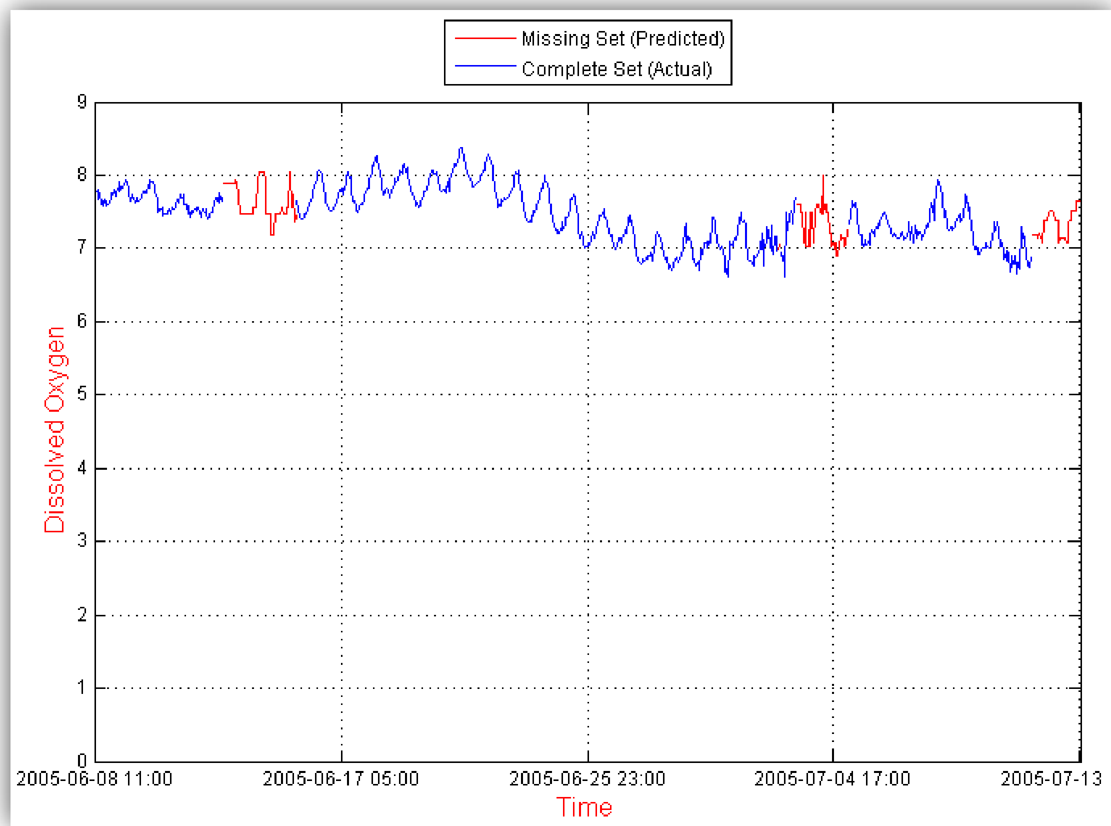


Figure 4.23. Actual and predicted values of dissolved oxygen

4.3.3. Modeling Component

4.3.3.1. Watershed Water Quality Simulator

Scenario: User wants to simulate a watershed's hydraulics, sediment and pollutant behavior to evaluate pollution discharges and water quality conditions.

Data series will be used from the upstream and downstream observation sites at the Soque River Basin.

Instructions: Go to 'Models' component and then 'Watershed Water Quality Simulator' section

Click 'Start' button

STEP 1

Select the observation site 'Soque River Site #1 - Upstream' from the selection box

Select the variable 'Flow', 'Suspended Solid', 'Total Reactive Phosphorus', and 'Precipitation' from the selection boxes (Figure 4.24)

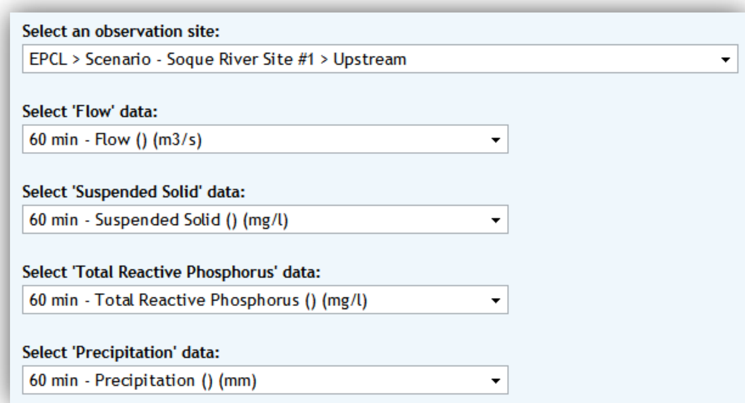


Figure 4.24. Step 1 of WWQS for upstream conditions

STEP 2

Select the observation site 'Soque River Site #2 - Downstream' from the selection box

Select the variable 'Flow', 'Suspended Solid', and 'Total Reactive Phosphorus' from the selection boxes (Figure 4.25)

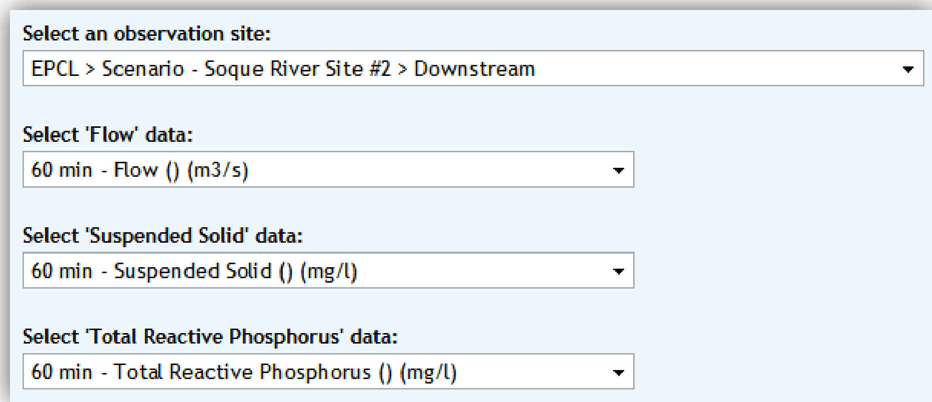


Figure 4.25. Step 2 of WWQS for downstream conditions

STEP 3

Enter the 'Drainage Area' value as '23.34' to the input box (Figure 4.26)

Enter the 'Potential Labile Phosphorus content in soil' value as '30'

Enter the 'Potential Phosphorus content in soil' value as '150'

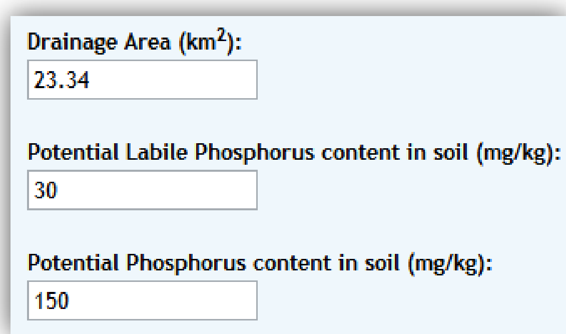


Figure 4.26. Step 3 of WWQS for watershed conditions

STEP 4

Enter the 'Elevation' value for upstream and downstream to the input box

Enter the 'Bottom' value for upstream and downstream to the input box

Enter the 'Side slope' value for upstream and downstream to the input box

Enter the 'Distance between CS1 and CS2' value to the input box (Figure 4.27)

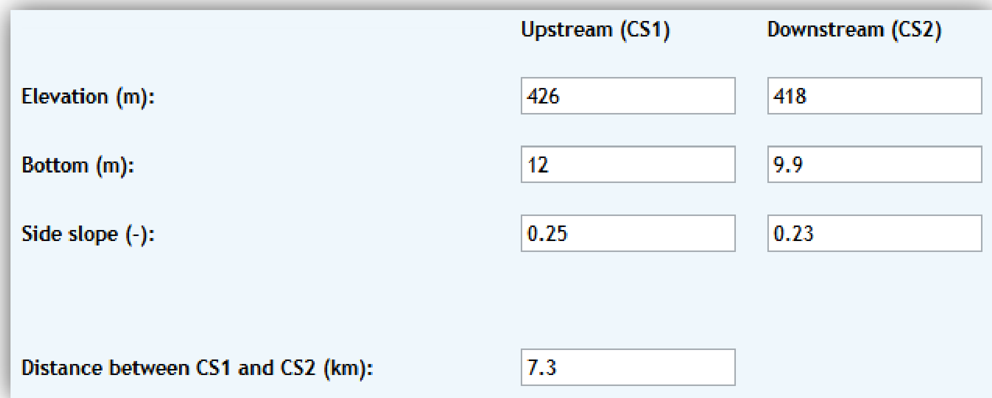


Figure 4.27. Step 4 of WWQS for river cross section conditions

STEP 5

Enter each simulation parameter value into 20 input boxes (Figure 4.28)

Click 'Start Simulator' button

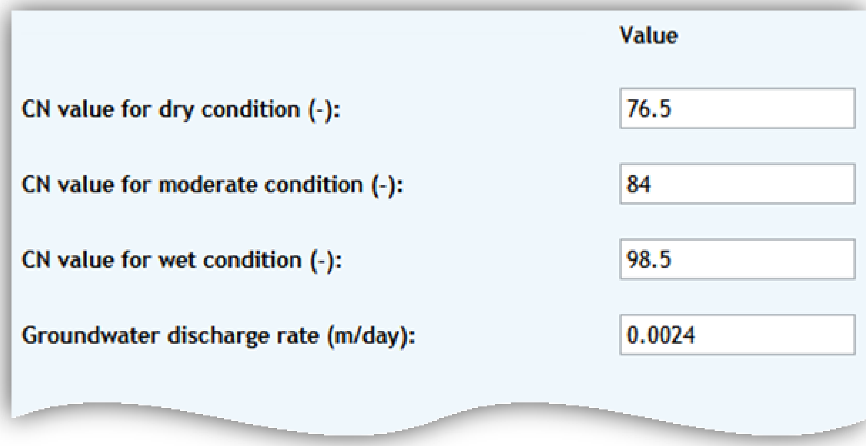


Figure 4.28. Step 5 of WWQS for entering 20 simulation parameters

Results: The following screen provides basic information about the observation sites (upstream and downstream), and a report will appear with plots of the simulation results for flow (Figure 4.29), suspended solids (in log scale, Figure 4.30) and total reactive phosphorus concentration (Figure 4.31), and confidence interval as an area plot around the simulated time series.

Click the 'Download Data' link in the report to export simulated values of the time series as a CSV file to your computer for further analysis

Click the 'Print Report' button to print the report as a PDF file or on paper from any printer

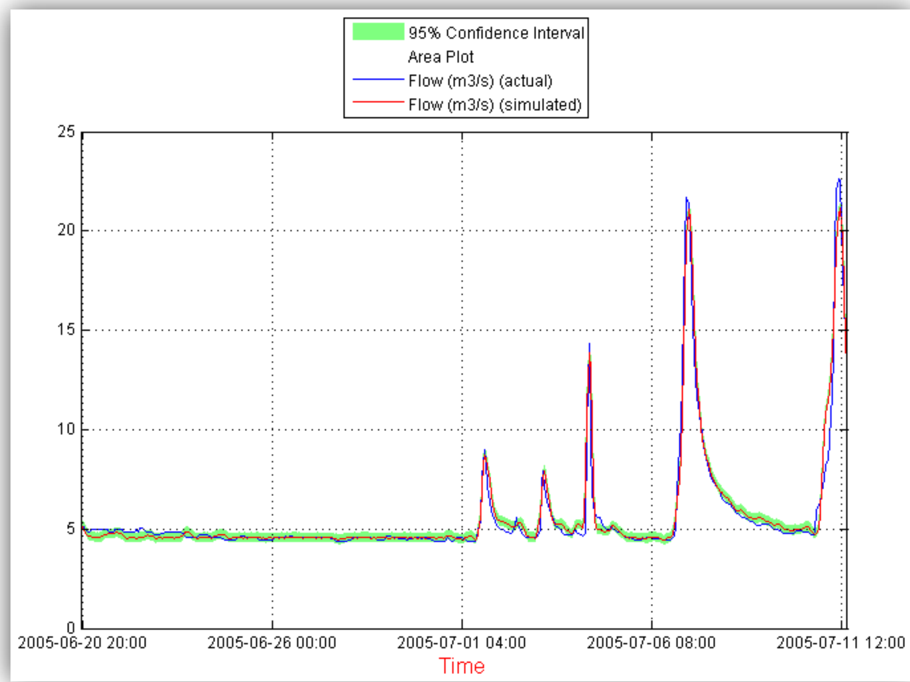


Figure 4.29. Actual and simulated values for flow by WWQS tool

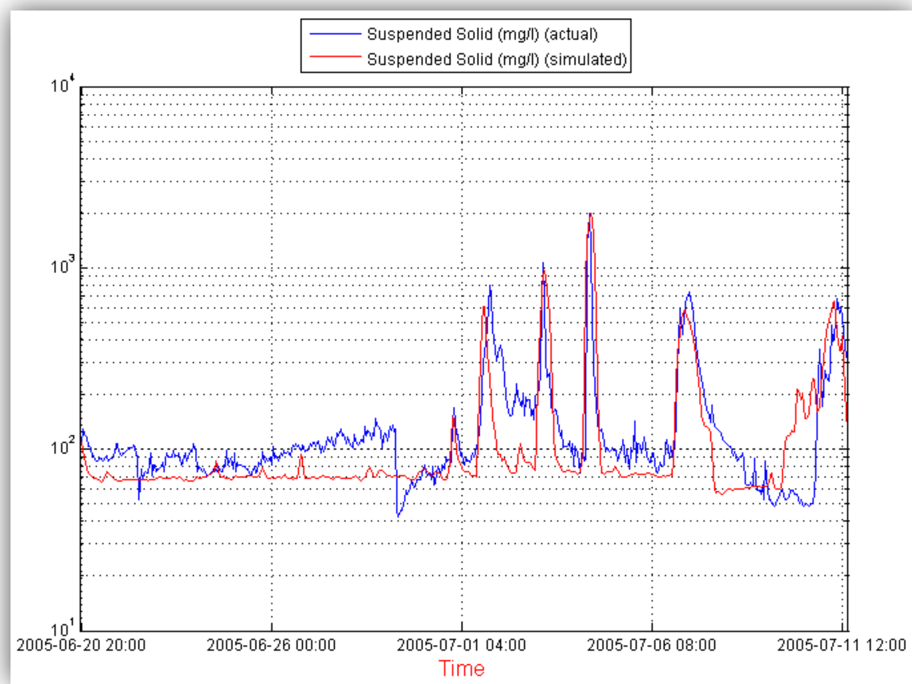


Figure 4.30. Actual and simulated values for suspended solids by WWQS tool

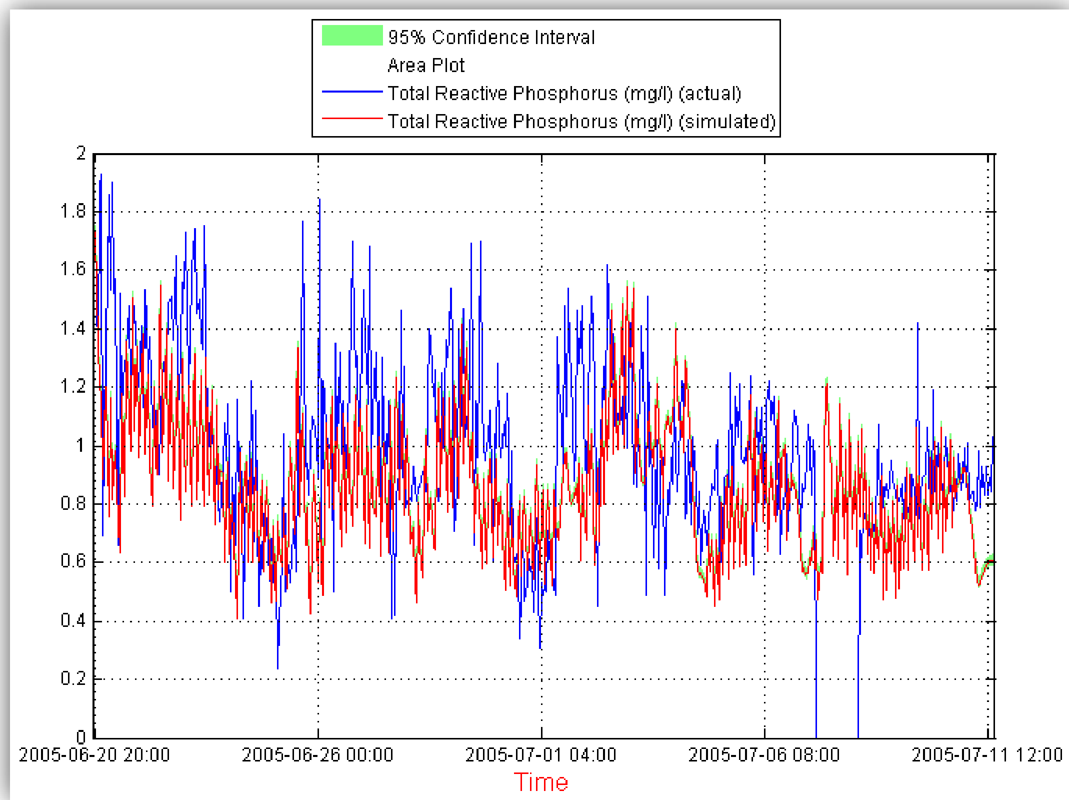


Figure 4.31. Actual and simulated values for total reactive phosphorus by WWQS tool

4.3.4. Visualization Component

4.3.4.1. Spatial Visualization on Interactive Maps Tool

Scenario: User wants to visualize the time series of dissolved oxygen from upstream to downstream locations at the observation sites (Weracoba Creek, Colombus) uploaded to GWIS, and animate the time series through time along the river structure

Instructions: Go to 'Visuals' component and then 'Spatial Visualization on Interactive Maps' section

Click 'Start' button

Select observation site 'Weracoba Creek, Columbus #1 BMP, Influent' from the selection box on the map (Figure 4.32)

Click 'Select this site' on the information box

Select the variable 'Ammonia (NH3)' from the selection box on the right panel, and click 'Select Variable' button (Figure 4.32)

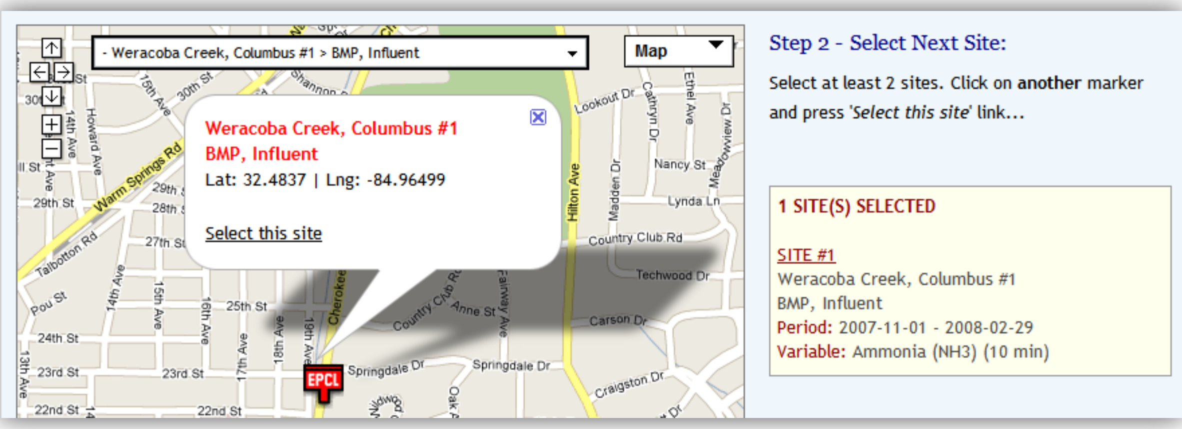


Figure 4.32. Selecting upstream site and variable from the interface

Select observation site 'Weracoba Creek, Columbus #3 Wynnton Rd Bridge' from the selection box on the map

Click 'Select this site' on the information box (Figure 4.33)

Select the variable 'Ammonia (NH3)' from the selection box on the right panel, and click 'Select Variable' button

Click the 'End Selection' button on the right panel (Figure 4.33)

Results: The next screen shows panoramic views of the upstream and downstream observation sites, and river channel structure generator (Figure 4.34)

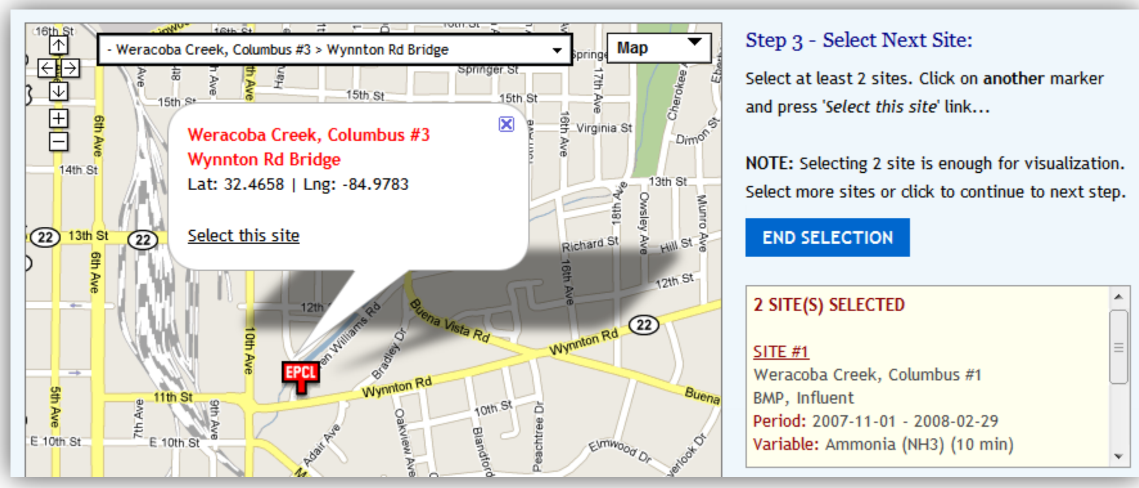


Figure 4.33. Selecting downstream site and variable to finalize selections

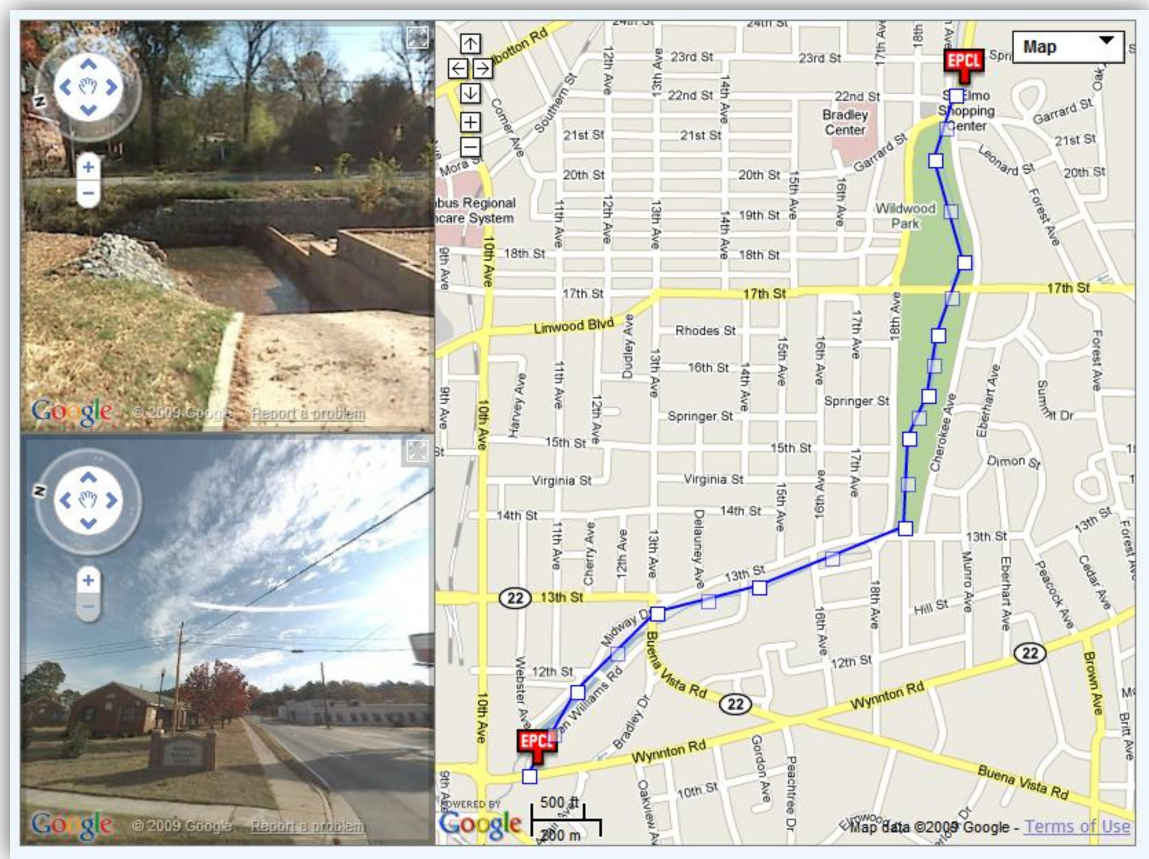


Figure 4.34. Panoramic view of the Weracoba Creek and river structure generator

Define the river channel structure on the map by moving blue squares along the river (Figure 4.34)

Click 'Next Step' button on the right panel

Results: The next screen provides animation controls (Figure 4.35) on the right panel. The user can adjust the speed of the simulation, and navigate by using the slider through time

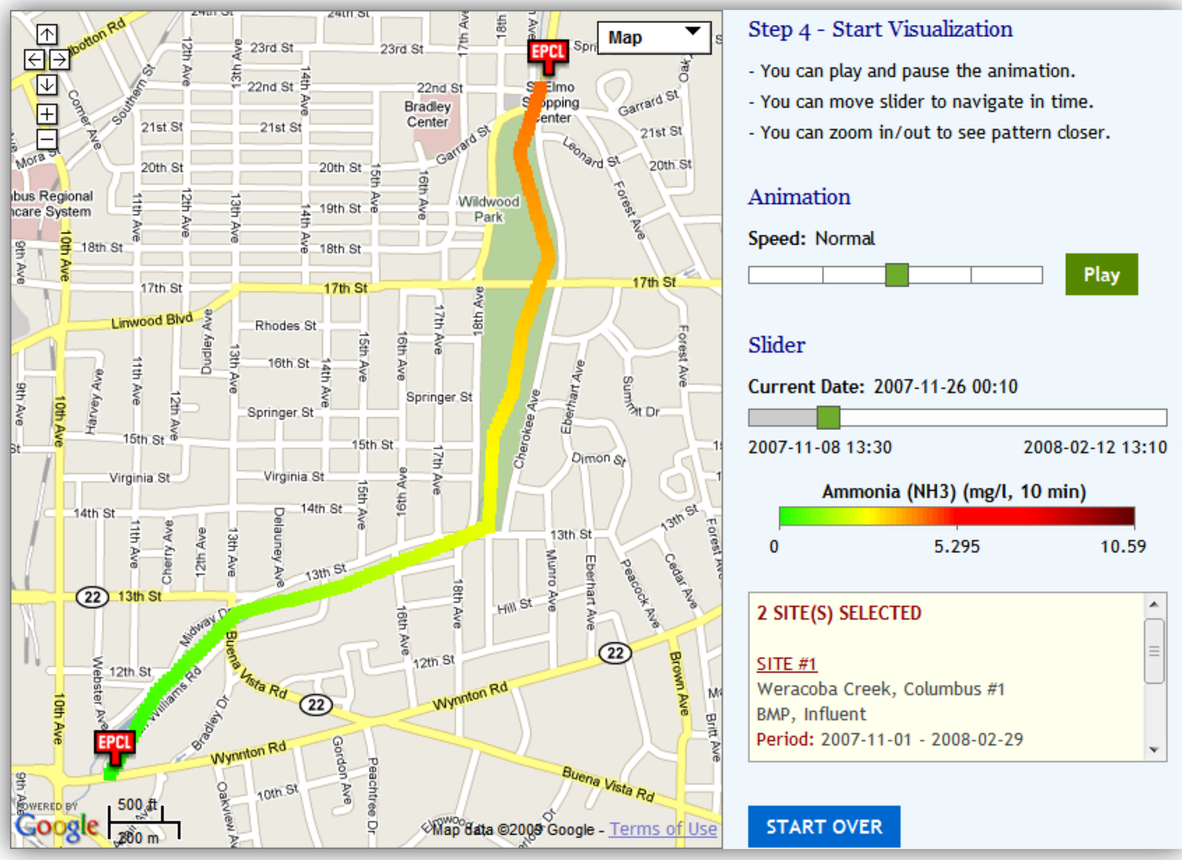


Figure 4.35. Animation controls and slider to visualize time series

4.3.4.2. Multi-dimensional Scatter and Histogram Charts

Scenario: User wants to generate multi-dimensional scatter and histogram charts of variables for the observation site (Soque River Site #1 - Upstream) uploaded to GWIS, and understand the correlations between variables

Instructions: Go to 'Visuals' component and then 'Multi-dimensional Scatter and Histogram Chart' section

Click 'Start' button

Select observation site 'Soque River Site #1 - Upstream' from the selection box on the map (Figure 4.36)

Select any number of variables from the selection box

Click 'Visualize' button

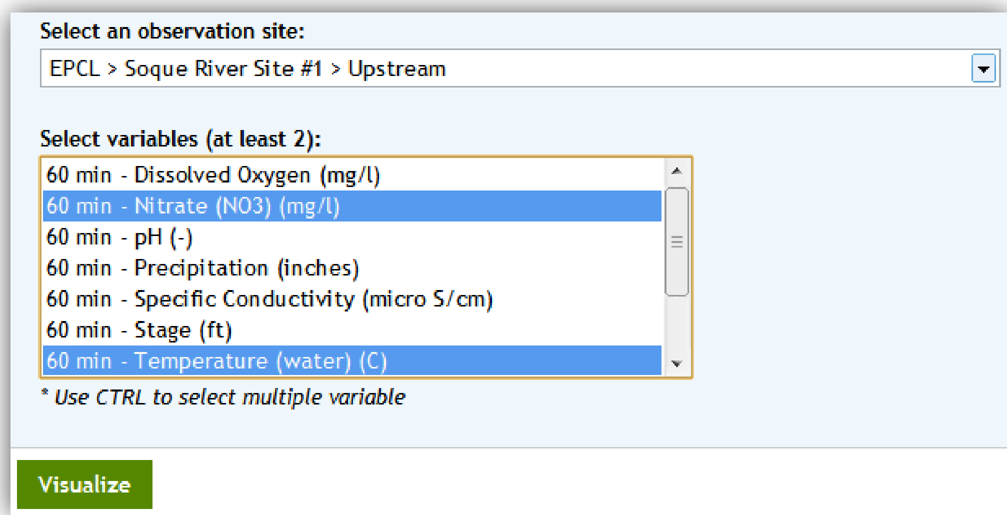


Figure 4.36. Multi-dimensional scatter and histogram tool interface

Results: The next screen provides basic information about the observation site, and a report will appear with scatter and histogram plots (Figure 4.37) for the selected variables

Click the 'Print Report' button to print the report as a PDF file or on paper from any printer

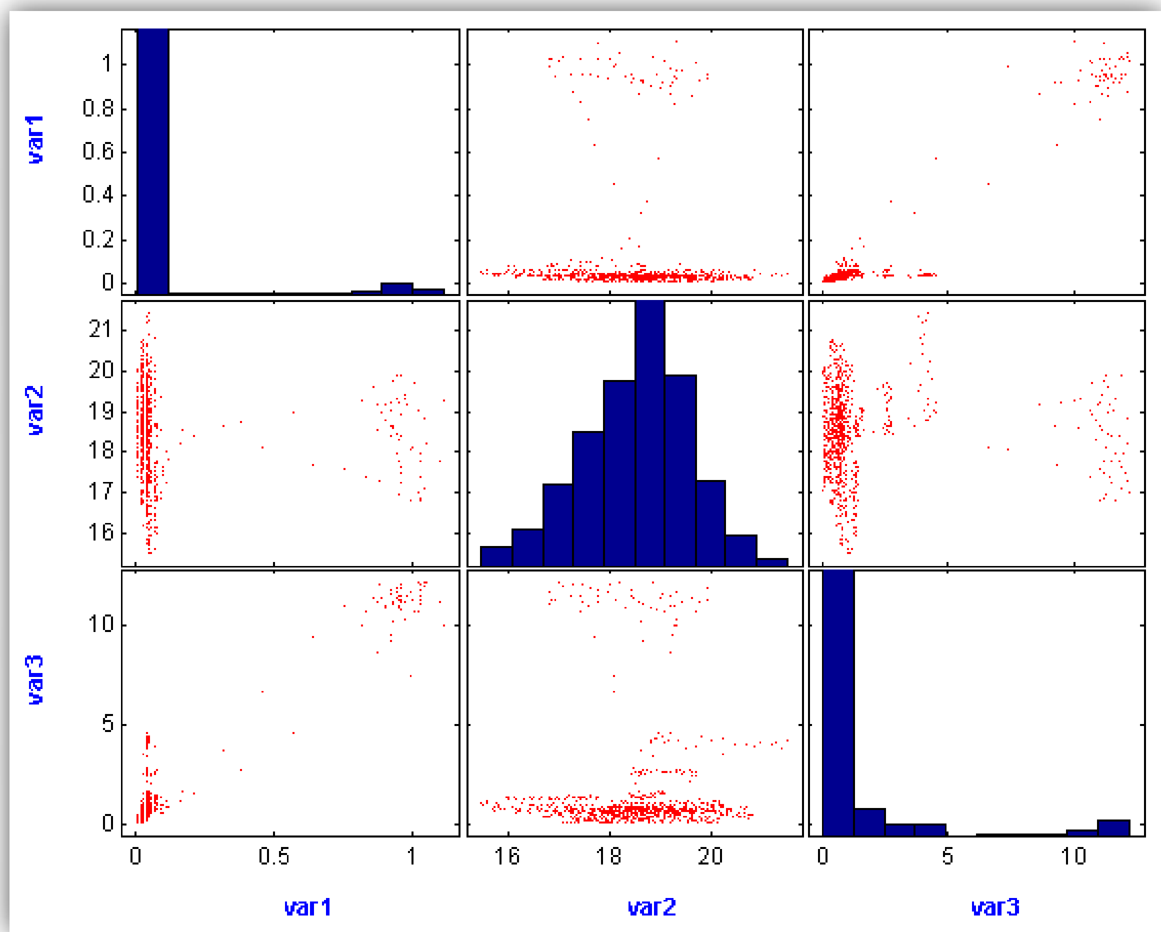


Figure 4.37. Scatter and histogram plots in a 3x3 matrix

4.3.4.3. Interactive Charts

Scenario: User wants to compare the time series of the dissolved oxygen variable for upstream and downstream locations at an observation site (Soque River) uploaded to GWIS

Instructions: Go to 'Visuals' component and then 'Interactive Charts' section

Click 'Start' button

Select the observation site 'Soque River Site #1 - Upstream' from the selection box (Figure 4.38)

Select the variable 'Dissolved Oxygen' from the selection box

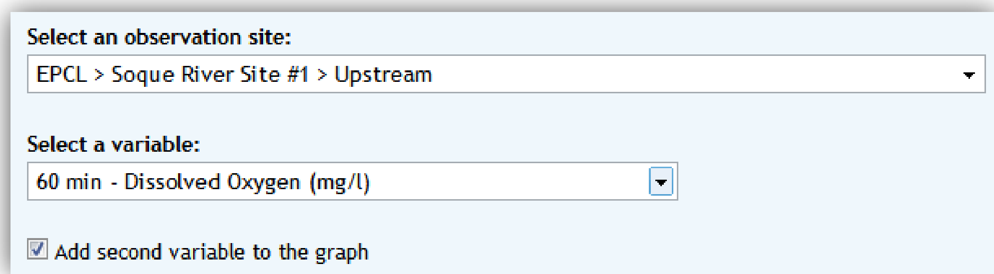


Figure 4.38. Selection of upstream site and dissolved oxygen in the interface

Check 'Add second variable to the graph' (Figure 4.38)

Select the observation site 'Soque River Site #2 - Downstream' from the selection box (Figure 4.39)

Select the variable 'Dissolved Oxygen' from the selection box

Click 'Visualize' button (Figure 4.39)

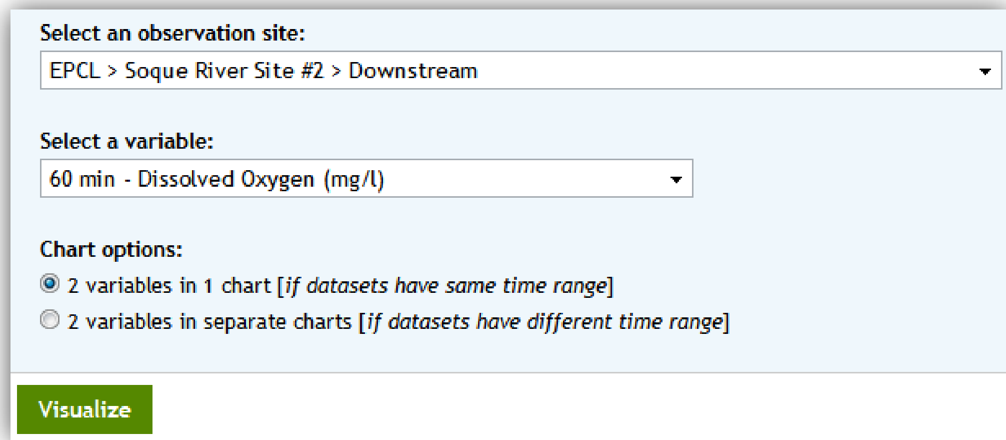


Figure 4.39. Selection of downstream site and dissolved oxygen on the interface

Results: The next screen provides basic information about the observation sites and an interactive plot of 'Dissolved Oxygen' for 'Soque River'. The plot overlays meta-data on the graph (Figure 4.40)

Right click on the plot and select 'Export GWIS Chart as image' to save the image to the computer

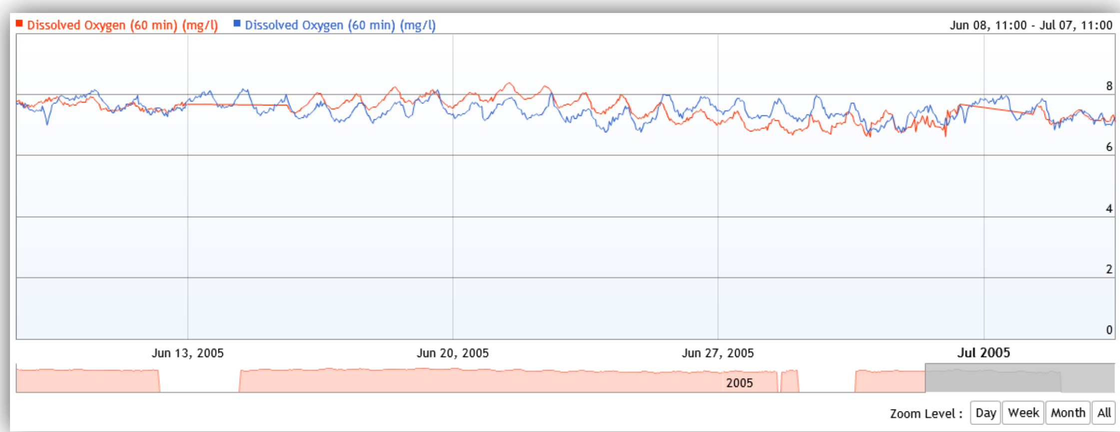


Figure 4.40. Comparison of upstream and downstream dissolved oxygen

4.3.4.4. Static Charts

Scenario: User wants to generate scatter plot of variables for the observation site (Soque River #1 - Upstream) uploaded to GWIS

Instructions: Go to 'Visuals' component and then 'Static Charts' section

Click 'Start' button

Select the observation site 'Soque River Site #1 - Upstream' from the selection box (Figure 4.41)

Select any number of variables from the selection box

Click 'Visualize' button

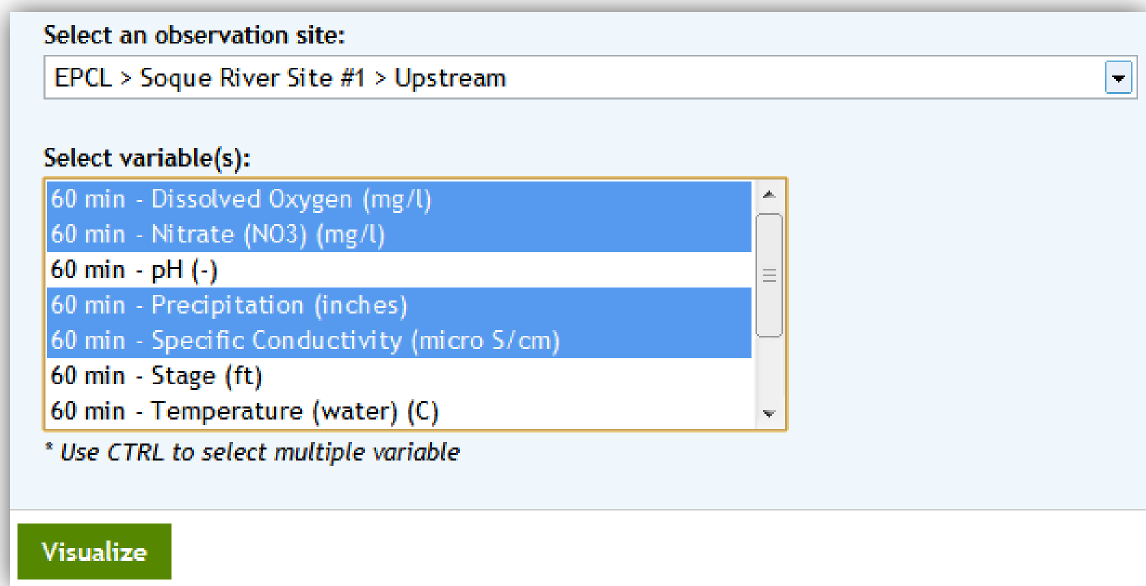


Figure 4.41. Selection of Soque River and variables for static chart

Results: The next screen provides basic information about the observation sites and a static plot of selected variables (Figure 4.42) for the 'Soque River'

Click the 'Print Report' button to print the report as a PDF file or on paper from any printer

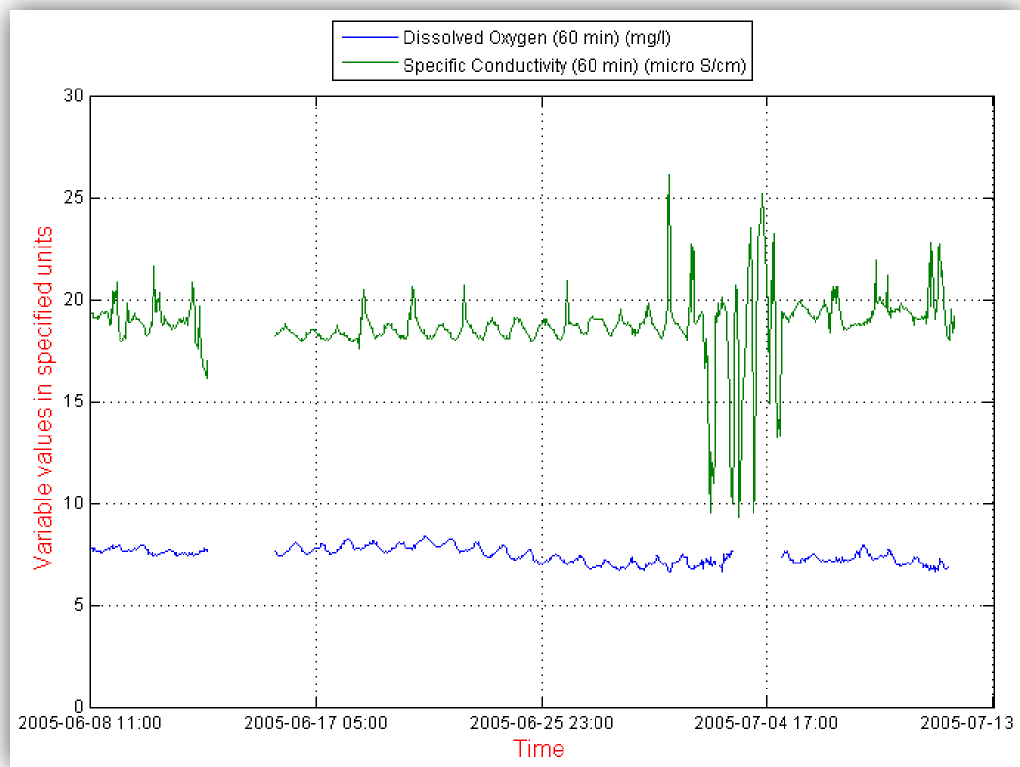


Figure 4.42. Static chart of dissolved oxygen and specific conductivity

4.3.5. Mapping Component

4.3.5.1. USGS Interactive Information Interface

Scenario: User wants to find observation sites from USGS data sources near the Soque River and quickly examine site information and data for discharge

Instructions: Go to 'Maps' component and then 'USGS Interactive Information Interface' section

Click 'Start' button

Navigate to Soque River area (Figure 4.43) on the interactive map by zooming in/out. Click on the blue cluster centers (e.g. '2' at lower left corner) to see the individual sites

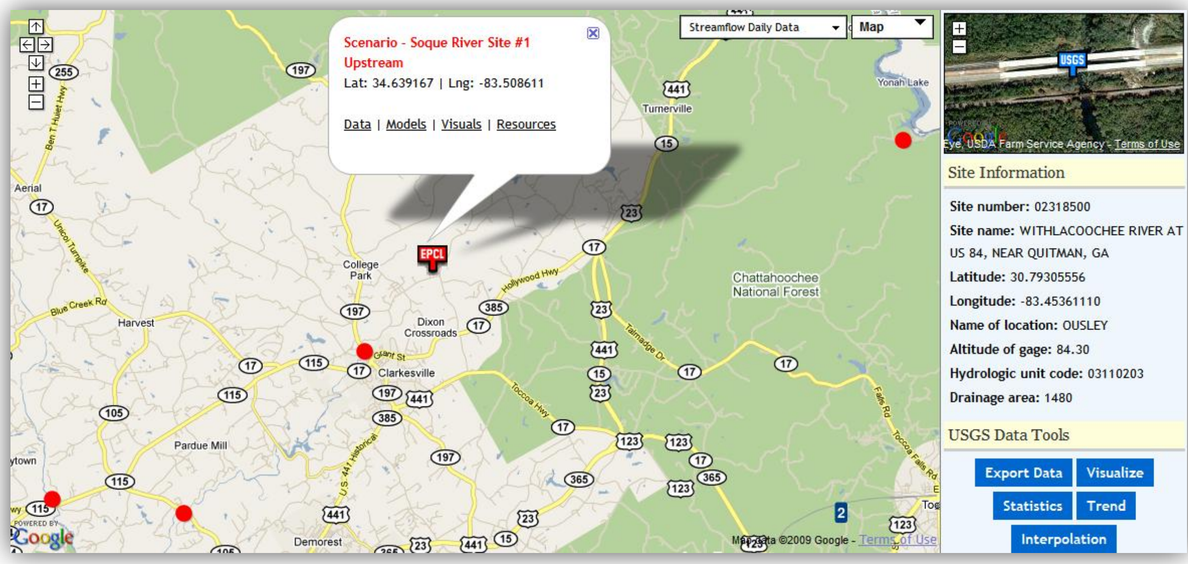


Figure 4.43. USGS sites near user's study area (Soque River) and information panel

Results: Click red circle to load information about selected observation site to the right panel (Figure 4.43). Discharge data will be transferred to the GWIS server automatically to make the data available for rapid download, or for carrying out analysis and visualization

4.3.5.2. GWIS Mapping Services

Scenario: User wants to navigate observation sites from GWIS server near his study area, Weracoba Creek, Columbus, and access related data and tools

Instructions: Go to 'Maps' component and then 'Maps powered by Google Maps API' section

Click 'Start' button

User can either select an observation site from the top selection site, or navigate to his study area (Figure 4.44) on the interactive map by zooming in/out

Results: Click markers to access information about selected observation site. Click links on the information box to access data and tools on GWIS server

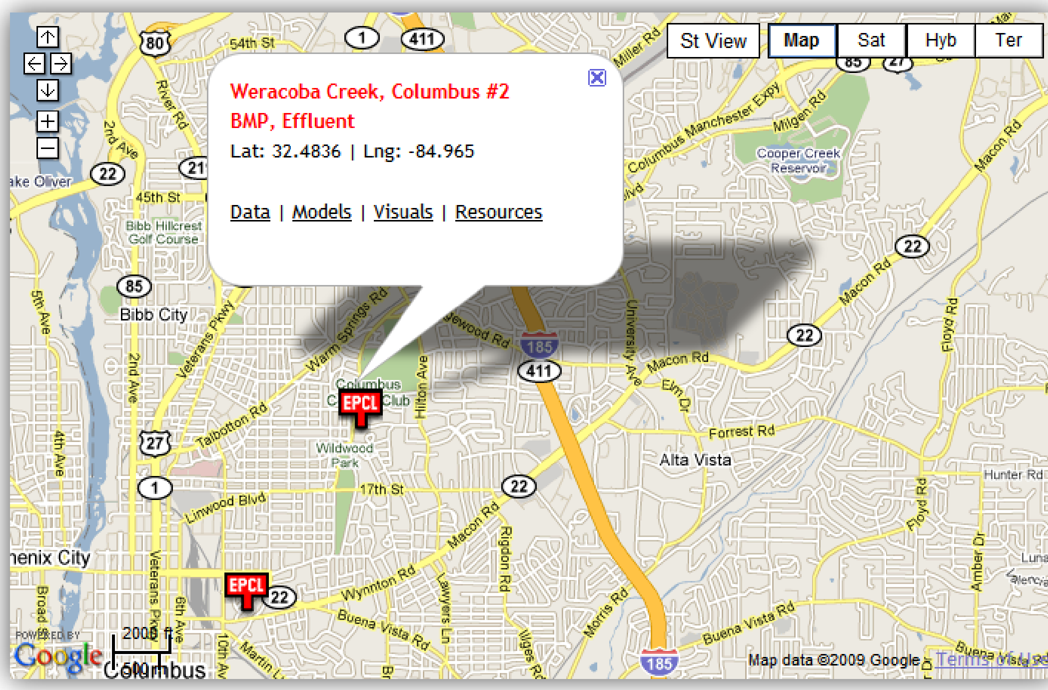


Figure 4.44. Near-by observation sites for Weracoba Creek, Columbus

4.3.5.3. KML Support on GWIS Maps

Scenario: User wants to access USGS real-time stream-gage values for North Oconee River, Athens, GA (site number: 02217770) on interactive maps

Instructions: Go to ‘Maps’ component and then ‘KML on Google Maps API’ section

Click ‘Start’ button

Enter KML url as ‘http://water.usgs.gov/waterwatch/kmls/real.kmz’ to input box

Click the ‘Set’ button

Results: KML content will be overlaid on the map (Figure 4.45). There are many freely available KML files on the overlay rich geographic content on the maps.

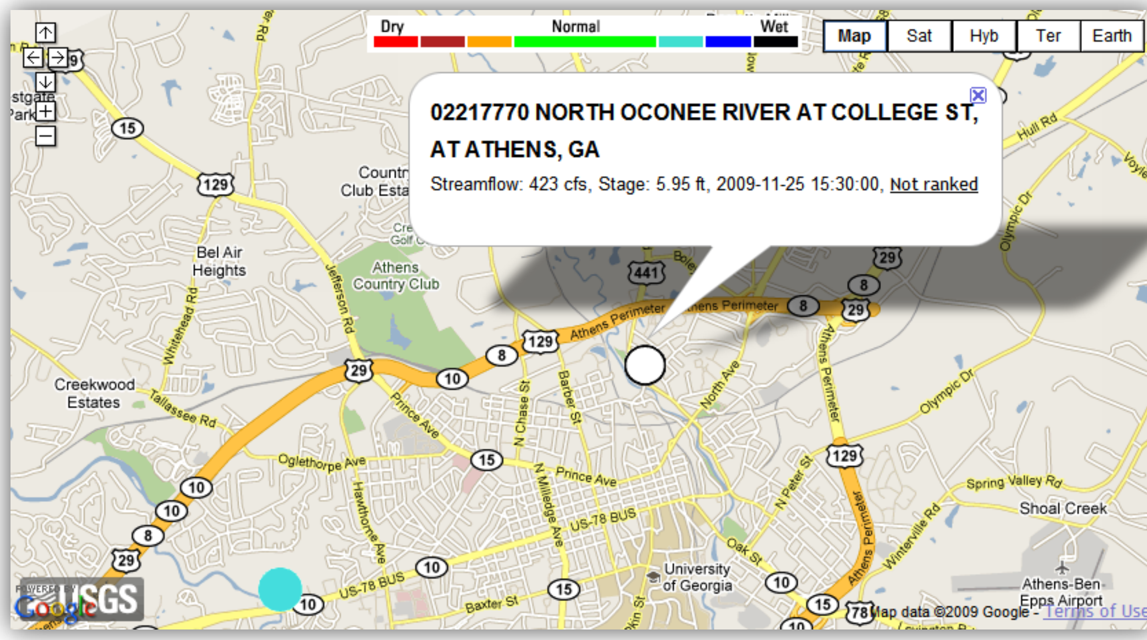


Figure 4.45. USGS real-time stream-gage overlay by KML for North Oconee River

4.3.6. Digital Resources Component

4.3.6.1. GWIS Digital Library

Scenario: User wants to access digital resources uploaded by GWIS users for an observation site (Athens Wastewater Treatment Plant No 2)

Instructions: Go to 'Library' component and then 'GWIS Digital Library' section

Click 'Start' button

Select the observation site 'Athens WWTP No.2 #1'

Click 'Show Resources' button (Figure 4.46)

Results: The next screen provides basic information about the observation site and a list of available digital resources (Figure 4.47)

Click to the link on the resources to download to your computer

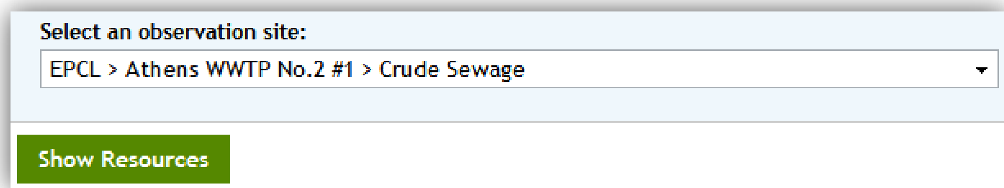
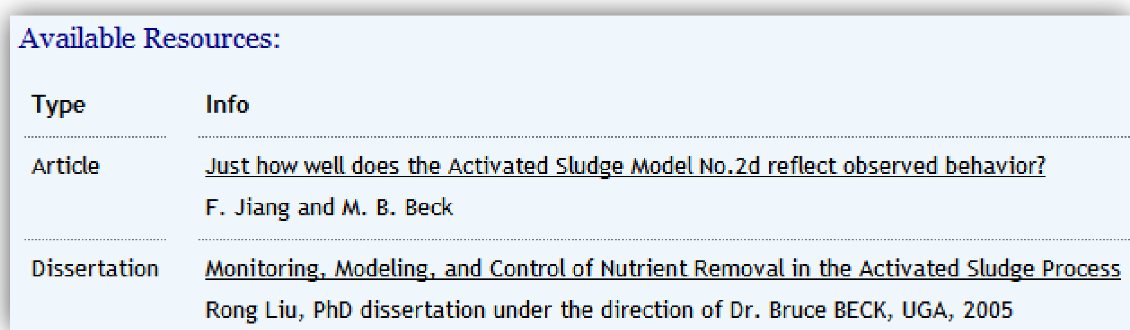


Figure 4.46. Selecting an observation site from GWIS Digital Library


Available Resources:

Type	Info
Article	Just how well does the Activated Sludge Model No.2d reflect observed behavior? F. Jiang and M. B. Beck
Dissertation	Monitoring, Modeling, and Control of Nutrient Removal in the Activated Sludge Process Rong Liu, PhD dissertation under the direction of Dr. Bruce BECK, UGA, 2005

Figure 4.47. Available digital resources are listed for selected observation site

4.3.6.2. GWIS Site Description

Scenario: User wants to access site description defined by GWIS users for an observation site (Soque River Site #1 - Upstream)

Instructions: Go to 'Library' component and then 'GWIS Site Description' section

Click 'Start' button

Select the observation site 'Soque River Site #1 - Upstream'

Click 'Site Info and Outside Sources' button

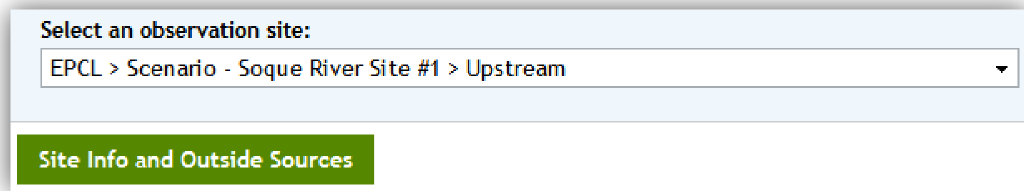


Figure 4.48. Selecting an observation site for site description

Results: The next screen provides site description about the observation site and links of related sources, such as EPA subbasin information and USGS data sources (Figure 4.49)

Click to the link of the sources to access more information

4.3.6.3. USGS Site Description

Scenario: User wants to access site description from USGS site for an observation site (Withlacoochee River, Quitman, GA) with site number '02318500'

Instructions: Go to 'Library' component and then 'Site Description by USGS Site Number'

Click 'Start' button

Enter '02318500' as USGS site number for Missouri River at Fort Benton

Click 'USGS Site Description' button (Figure 4.50)

Results: The next screen (Figure 4.51) provides site description about the observation site from USGS

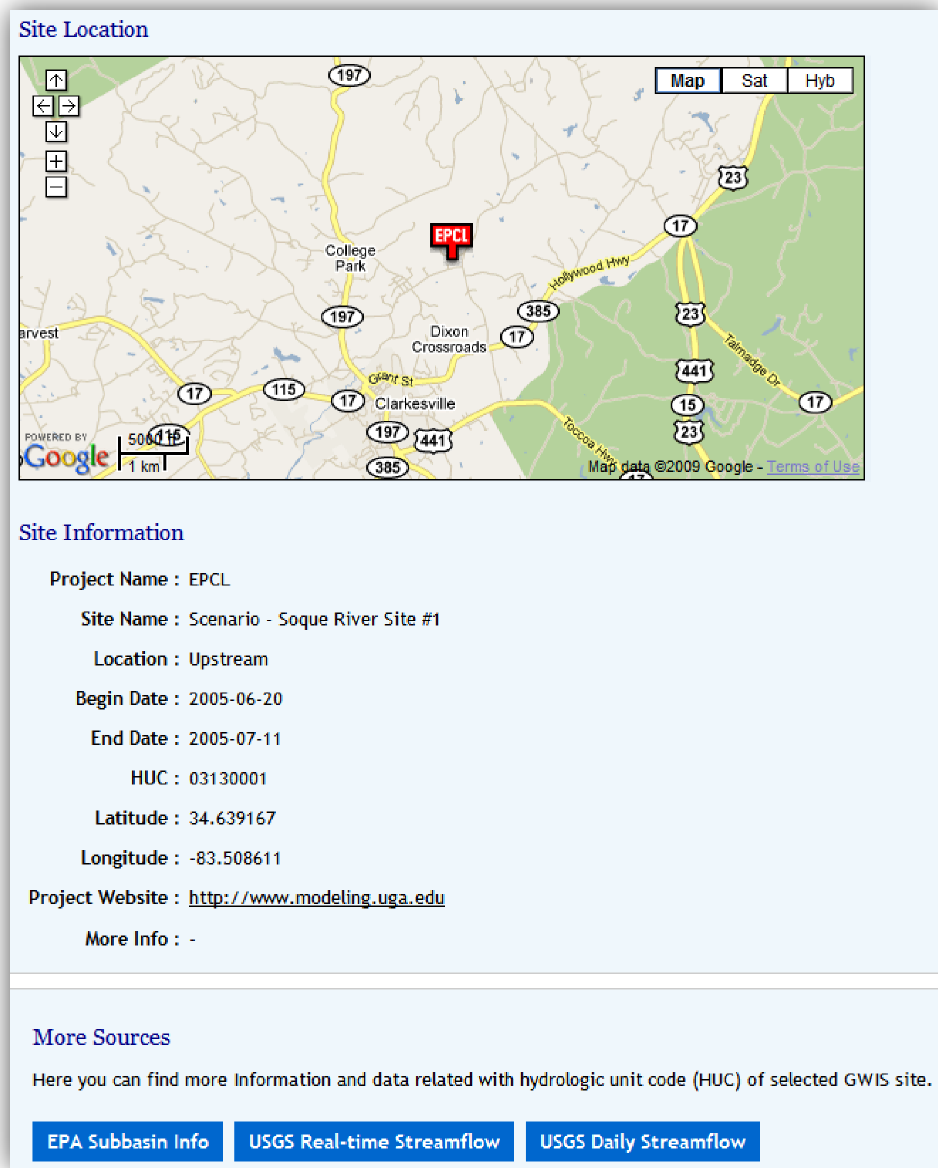


Figure 4.49. GWIS site description with links to related sources

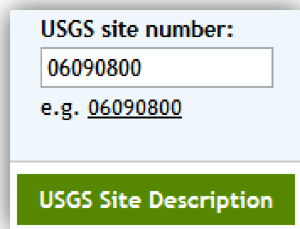


Figure 4.50. Site description by USGS site number tool

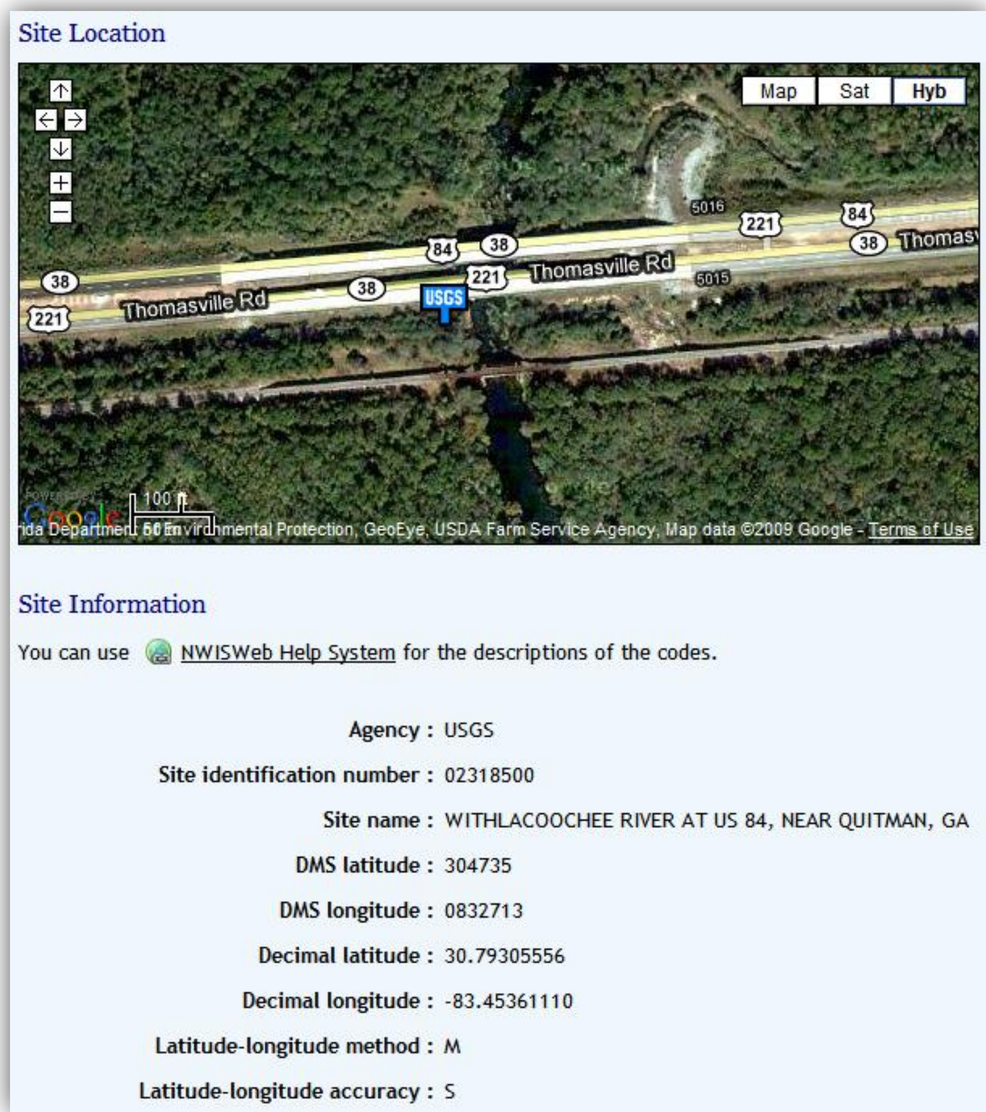


Figure 4.51. Site description from USGS by using site number

4.3.6.4. Site Description Web Service

Scenario #1: User wants to generate a web service URL to access site description of the observation site 'Soque River Site #1 - Upstream' to use in his software

Instructions: Go to the 'Library' component and then the 'GWIS Site Description Web Service' section

Click 'Start' button

Select the observation site 'Soque River Site #1 - Upstream' from the selection box (Figure 4.52)

Results: Copy the automatically generated URL from the yellow box and it paste into your application (Figure 4.52)

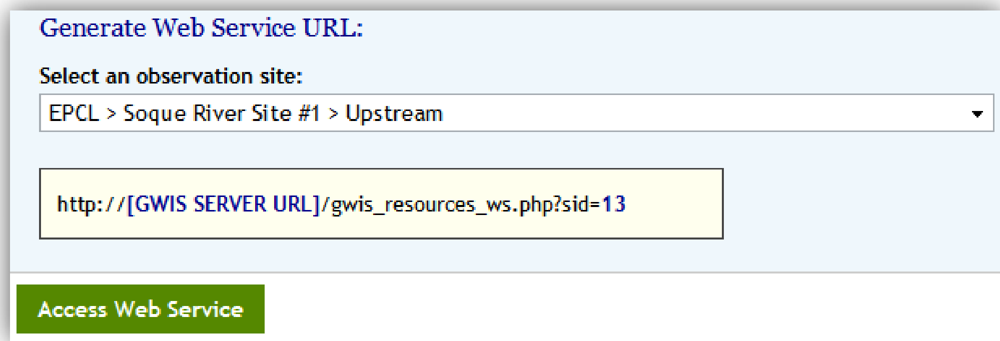


Figure 4.52. GWIS site description web service URL generation

Scenario #2: User wants to download the site description file of the observation site 'Soque River Site #1 - Upstream' to use on his computer by using GWIS web service

Instructions: Go to ‘Library’ component and then ‘GWIS Site Description Web Service’ section

Click ‘Start’ button

Select an observation site from the list

Click ‘Access Web Service’ button

Results: Save the file (Figure 4.53) to your computer


```xml
<gwis_ws>
  <site>
    <project_name>EPCL</project_name>
    <project_website>http://www.modeling.uga.edu</project_website>
    <site_id>12</site_id>
    <site_name>Weracoba Creek, Columbus #1</site_name>
    <site_location>BMP, Influent</site_location>
    <site_huc>03130003</site_huc>
    <site_info>-</site_info>
    <dec_lat>32.4837</dec_lat>
    <dec_lng>-84.96499</dec_lng>
  </site>
</gwis_ws>
```


Figure 4.53. Structure and contents of an XML file for site description

4.3.6.5. Wikipedia Retriever Tool

Scenario: User wants to access information about ‘Great Lakes Commission’

Instructions: Go to ‘Library’ component and then ‘GWIS Wiki Fork’ section

Click ‘Start’ button

Enter the term ‘Great Lakes Commission’ in the input box (Figure 4.54)

Click ‘GO’ button

Results: The next screen (Figure 4.55) transfers information for the selected term from Wikipedia directly to GWIS servers, including links, images and text.

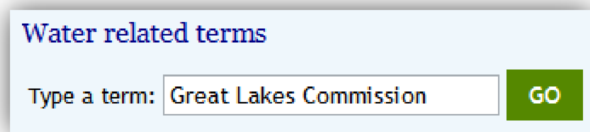


Figure 4.54. Input term in the input box and click GO

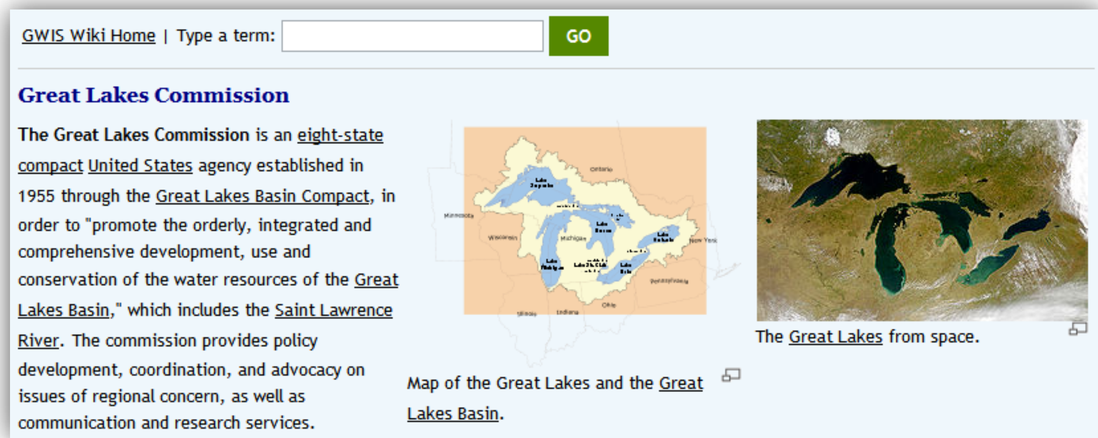


Figure 4.55. Information transferred directly from Wikipedia to GWIS server

4.3.6.6. Meta-data Generator Tool

Scenario #1: User wants to create a meta-data file for the observation site 'Athens WWTP No.2' monitoring campaign

Instructions: Go to 'Library' component and then 'Meta-data Generator' section

Click 'Start' button

Select the observation site 'Soque River Site #1 - Upstream' from the selection box (Figure 4.56)

Click 'Start Generator' button

Results: The next screen provides an input form to create a new meta-data file. The input form for the meta-data file has input boxes for date, time and description. Click the 'Add More' button to add new events for the meta-data file

Click the 'Create Meta-data' button to save the XML file (Figure 4.57)

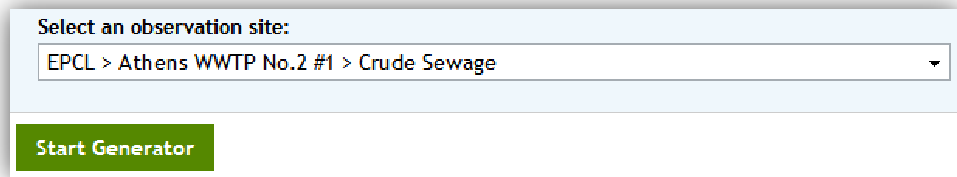


Figure 4.56. Site selection list for meta-data generator

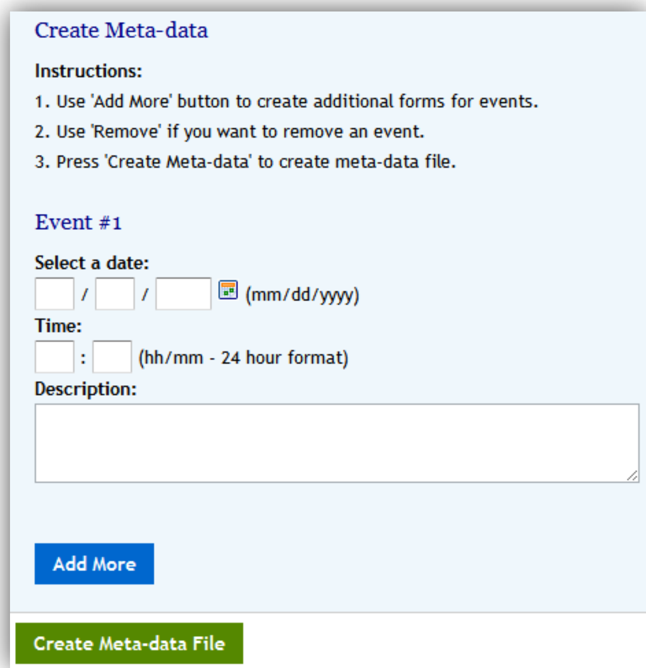


Figure 4.57. Meta-data generation screen with data, time and description input fields

Scenario #2: User wants to download or view the meta-data file for the observation site 'Soque River Site #1 - Upstream'

Instructions: Go to the 'Library' component and then the 'Meta-data Generator' section

Click 'Start' button

Select the observation site 'Soque River Site #1 - Upstream' from the selection box (Figure 4.56)

Click the 'Start Generator' button

Results: The next screen (Figure 4.58) provides information about the observation site. If the site does not have a meta-data file, an input form will be provided to create a new meta-data file.

Click the 'View Meta-data' or 'Export Meta-data' button to view (Figure 4.59) or download the XML file to your computer.

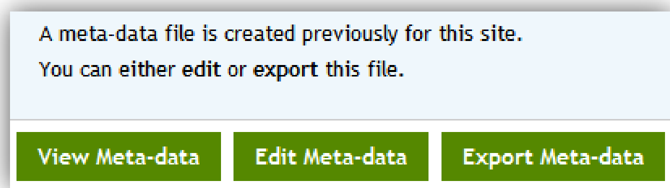
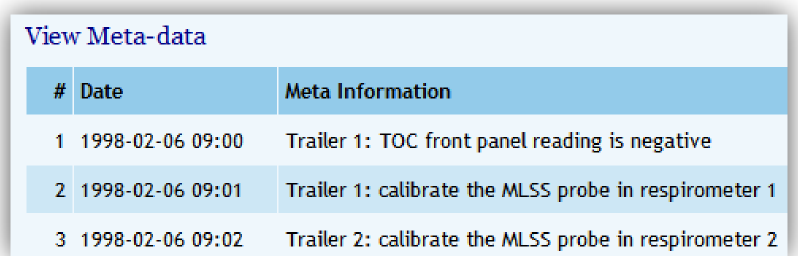


Figure 4.58. View, edit and export buttons for meta-data



View Meta-data

#	Date	Meta Information
1	1998-02-06 09:00	Trailer 1: TOC front panel reading is negative
2	1998-02-06 09:01	Trailer 1: calibrate the MLSS probe in respirometer 1
3	1998-02-06 09:02	Trailer 2: calibrate the MLSS probe in respirometer 2

Figure 4.59. Date and information fields of the meta-data file

4.4. User and Content Management

This section describes user management and content submission features of GWIS. User management includes user registration, login and editing profile. The content submission section describes starting a project, adding observation site, variables, and digital resources.

4.4.1. User Management

User registration is required for submitting content to GWIS. The registration form (Figure 4.60) asks only basic user profile and login information. Users create a username and password to login the GWIS. If a user forgets the password, GWIS automatically sends the password to the user by asking for his/her email address. Users can change their profile information from the management panel.

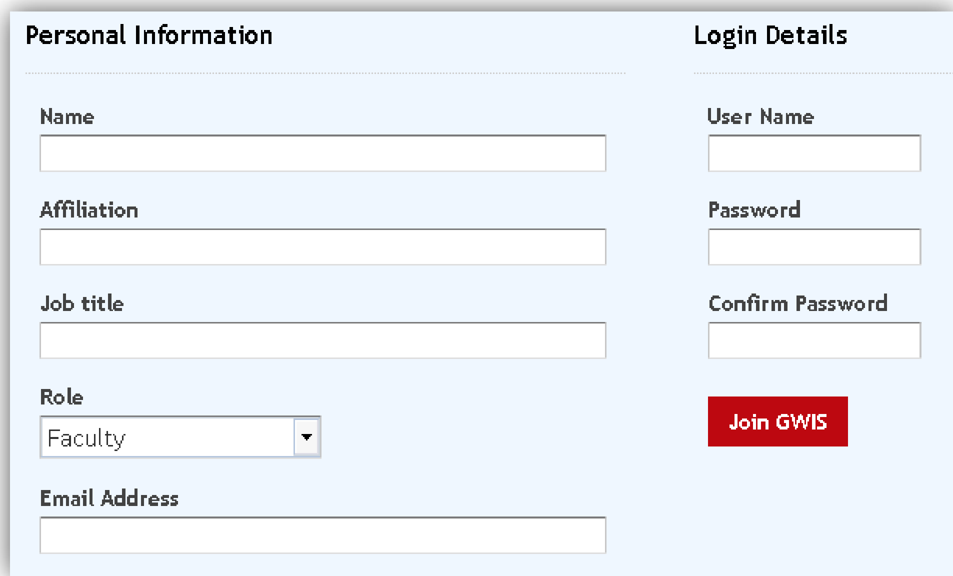


Figure 4.60. User registration form

4.4.2. Content Submission

Users are required to login to GWIS to submit content. The management panel (Figure 4.61) is accessible by on the site functions menu with 'USERS' link. GWIS provides interactive forms to submit or upload content. Several content types are supported by the system, especially for digital resources management.

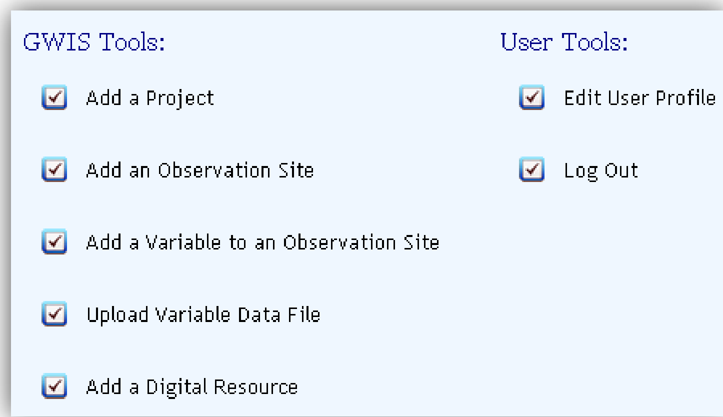


Figure 4.61. GWIS management panel

4.4.2.1. Defining a Project

Users can define a project from the management panel. It suffices to provide only a name and a link (Figure 4.62) to define a project. The project will be defined after clicking the button 'Add Project'.

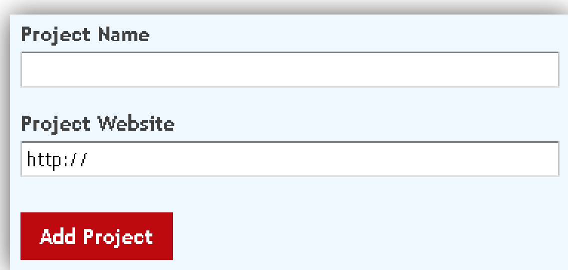


Figure 4.62. Adding a project to GWIS

4.4.2.2. Adding an Observation Site

Users can add an observation site to their projects from the management panel. Users need to select a project for the observation site (Figure 4.63), and provide the name, location, HUC number, latitude, longitude, campaign start and end dates, and campaign description for the observation site. The observation site will be added after clicking the 'Add Site' button.

Figure 4.63. Adding an observation site to GWIS

4.4.2.3. Adding a Variable

Users can also add a variable to an observation site from the management panel. Users need to select a project and observation site (Figure 4.63), and provide name, unit and sampling frequency for the variable. The variable will be added after clicking the button 'Add Variable'.

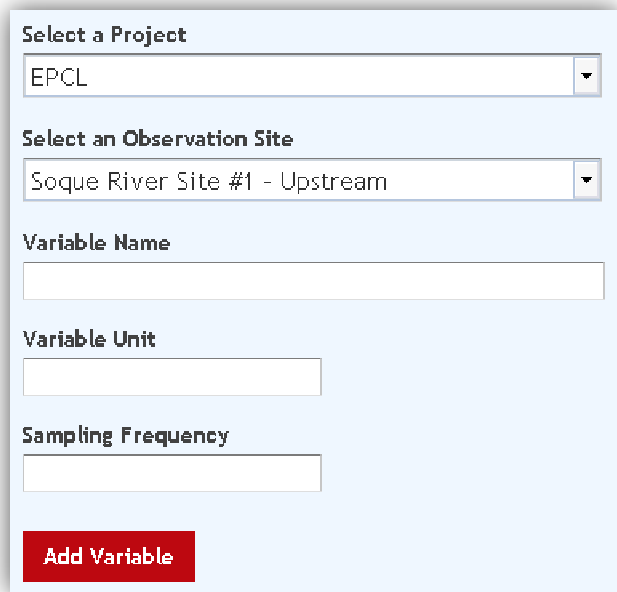


Figure 4.64. Adding a variable to an observation site

4.4.2.4. Uploading Variable Data

Users can upload variable data for a variable from the management panel. Users need to select a project, an observation site and a variable (Figure 4.65), and provide a data file for the variable. An acceptable file format is Comma Separated Value (CSV) with time series defining each sample point in a line. Each line should have a date and concentration value in specified units. The data file will be added after clicking the button 'Upload Variable'.

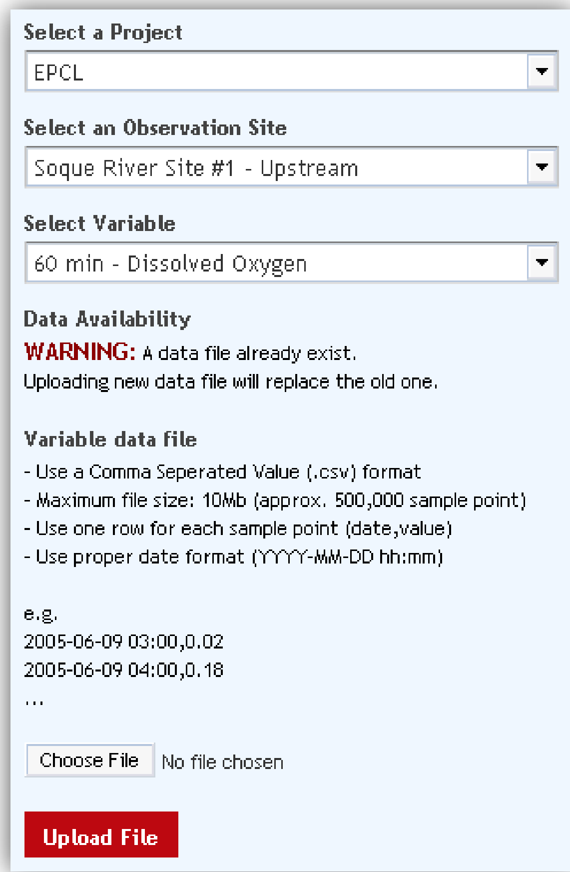


Figure 4.65. Uploading time series data for selected variable

4.4.2.5. Adding Digital Resources

Users can submit digital resources from the management panel. Users need to select a project, an observation site and resource type (e.g. dissertation, article, report, meta-data, meeting minutes, report, campaign notes, and etc), and provide title, source link and additional information for the digital resource (Figure 4.66). GWIS accepts various file formats for digital resources like documents, presentations, images, and other common file formats. The digital resource will be added after clicking the button 'Add Resource'.

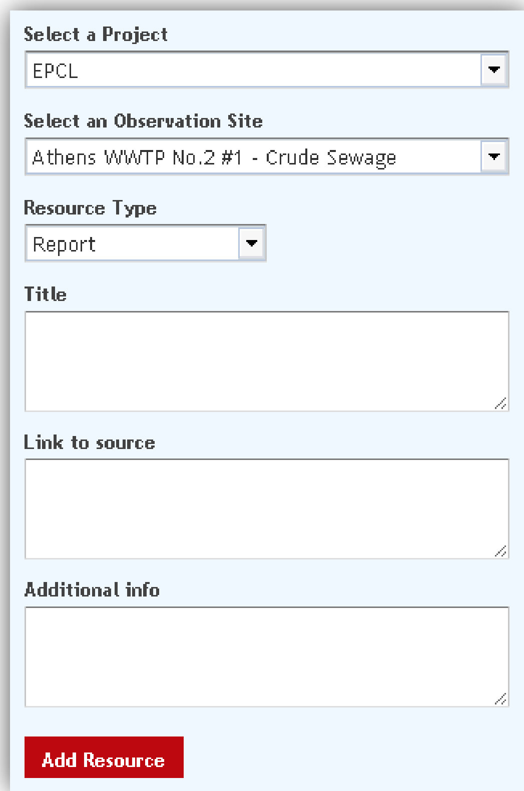


Figure 4.66. Adding digital resources to GWIS

4.4.3. New Feature Submission

As an open development platform, GWIS provides a submission form so that users can suggest or provide new tools and features to expand GWIS. Users can develop new analysis, modeling or scientific visualization tools and make them available to the public by using GWIS. The system provides programming codes for sample tools and features and provides instructions to build new interfaces. Users can also suggest new data and mapping services to be integrated to the system. Users can submit a new feature by suggesting a name, selecting a component and feature type and instructions or codes for the tools by using the web interface (Figure 4.67).

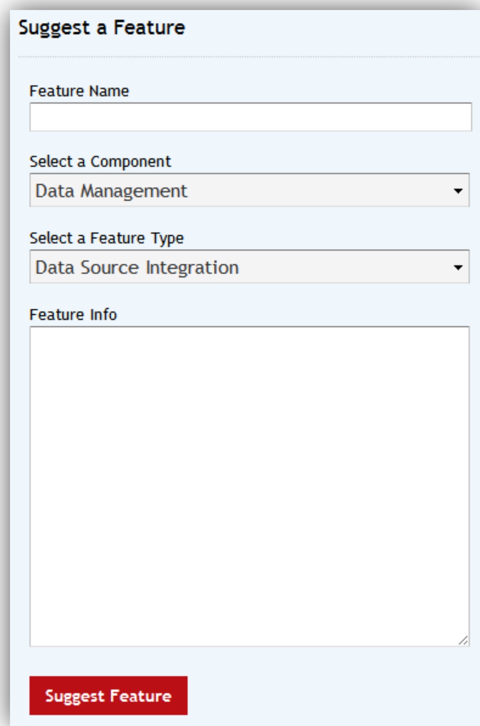


Figure 4.67. Suggesting a new feature or tool to GWIS

4.5. Interactive Help Tool

GWIS provides a novel interactive help tool, called GWIZARD (Figure 4.68). It is available on the left panel of the website on every page. It provides a step-by-step guidance to users so that users can utilize the system to its full extent. At each step, GWIZARD provides description of the current page or tool. Users can go backward and forward with the tool, or skip any steps. GWIZARD remembers the last step the user visited on the system, so that users can continue their navigation from the point they left. GWIZARD follows a specific order as in a management project, starting from understanding the characteristics of the area, examining the observation site and data, data visualization, analysis, modeling and digital resources.

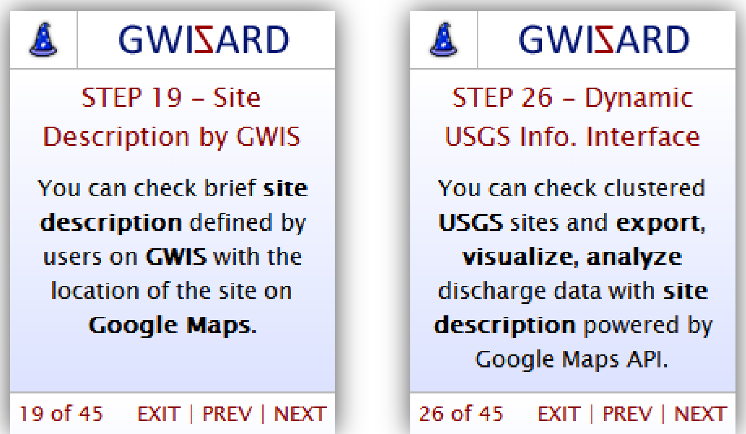


Figure 4.68. Sample interfaces of interactive help tool (GWIZARD)

4.6. Conclusion

GWIS tools and features are explained extensively with references to related methods and literature in Chapter 3. Mini case studies with real water quality data for each tool of the main components have been explained in this chapter. The mini examples explain the novel and advanced tools of GWIS step by step with instructions and screenshots to help users fully understand and utilize interfaces. The next chapter presents a large-scale case study, a step-by-step comparison of GWIS with traditional methods, and how GWIS can make watershed research and management easy, user-friendly, and more efficient with its novel and distinct features supported by the latest web technologies.

CHAPTER 5

CASE STUDIES

5.1. Introduction

This chapter illustrates novel features and approaches of GWIS as an Information System with reference to case studies in Environmental Research and Watershed Management with respect to the use of data, digital resources and the tools of GWIS. A specific case study, the Chattahoochee Watershed project of Shi (2008), is reviewed to demonstrate and compare (a) the way this project was carried out in the absence of GWIS, and (b) the way it could have been implemented using GWIS. During the comparison of these two cases, it is assumed that GWIS is fully developed with substantial content, including high-volume high-quality water quality data, digital resources and analysis tools.

5.2. GWIS as an Information System

Environmental studies typically involve many types of information that need to be managed and analyzed by using various tools. Each step of the management and analysis during a case study may require choosing and fully utilizing different tools to manage this information. Each tool or software with different data input and output formats with differing platform and computing requirements, license and management costs, and user interfaces, makes watershed research harder. While most of the software packages require a one-time purchase and annual maintenance/update costs, some of them are based on annual service subscriptions with no initial cost.

While each software system has advantages, as a web-based system, GWIS comes with many benefits, including easy access to information and tools, allowing users to share data and digital resources, integrating data sources and models, analyzing data and allowing access to digital resources in a way that is customized to fit the specific needs of given users and projects. GWIS increases collaboration in projects through on-line access by project members to share up-to-date project data, documents, and information.

As a web-based information system, GWIS has several advantages over its more traditional software alternatives. The key features applicable to environmental projects in general are listed in below (Table 5.1).

Table 5.1. Key features and descriptions of GWIS as a web-based information system

Features	Description
Ease of Use	As a web-based information system, GWIS has a simple and user-friendly interface, and ensures a consistent look-and-feel across all of its tools; on the other hand traditional systems generally have complex interfaces. Their interfaces also vary for each task, since separate software packages are required for each specific purpose.
Cross-platform compatibility	GWIS is accessible through any operating system and web browser on a PC, Mac or most of the mobile platforms, while traditional systems have specific operating system and platform requirements.
Updating	GWIS can be easily updated and modified without requiring the user to take any action, and without needing to interfere with his/her work, whereas traditional systems need to be updated on the user-side by downloading and installing the new software releases.

Table 5.1. continues

Accessibility	The user can access GWIS online and be ready to work, no matter what their setup or hardware is, while traditional systems have a limited accessibility.
Ease of experimentation	Users can try GWIS tools and features easily and effectively before deciding to transfer their work fully to GWIS, while traditional systems allow limited try-out with additional download and installation work.
Less hardware requirements	GWIS applications impose more reasonable demands on end-user hardware sources than thereof locally installed programs, while distributing the load to online servers and distributed sources.
Fewer Problems	GWIS tools should be less prone to technical problems due to software or hardware conflicts with other existing software. All problems can be fixed on the server side as soon as they are discovered, while traditional systems have to be updated by each user for every problem and change.
Flexible Development	Since GWIS is a collection of tools rather than a single program, each component could be written in a different programming language, while traditional systems are bound to use the same language most of the time.
Highly Deployable	Deploying GWIS to the end user is far easier. The user simply needs a website address to log in to access the latest version of GWIS. On the other hand, traditional systems require a computer system with specific requirements and software to be installed by an administrator to start using the systems.

Table 5.1. continues

Expandability	Data sources, mapping services and analysis tools are not limited by what GWIS provides in the system. GWIS demonstrates examples for integration of such tools and sources. GWIS provides a flexible structure that makes it easy to integrate various web services. Most of the advanced modeling environments provide non-graphical user-interface access to the system. It is easy to develop an interface between a web server and the modeling environment to integrate new tools to GWIS. For example, GWIS features two different programming environments, MATLAB and Visual Basic, integrated to its analysis and modeling components.
Reduced Costs	GWIS can dramatically lower costs due to reduced support and maintenance, lower requirements on the end user system and its simplified architecture.

5.3. GWIS for Environmental Research

All environmental research projects need to handle one or more of the following steps: (a) information management; (b) assembling data; (c) analysis of data; (d) collecting digital resources; (e) modeling; (f) scientific visualization; and (g) accessing spatial data and maps. GWIS can provide specific advantages for most of these topics compared to traditional approaches as given below.

a) Information Management

- Centralized access to data, tools and digital resources
- Sharing of project and campaign data, publications, models

- Publicizing your study and research group by sharing your data, campaign notes, publications and other digital resources about your projects
- Collaboration with other researchers related to your study area

b) Assembling Data

- Access to high-volume high-quality data sets (e.g., the EPCL data sets)
- Integration of distributed data sources (e.g., USGS daily and real-time data sets)

c) Data Analysis

- Carry out statistical and hydrological analyses
- Calculate trends and seasonal components
- Carry out correlation and interpolation analyses
- Analyze and complete missing data sets
- No external software installation required for analyses
- Integrating other analysis tools

See Chapter 3.4.2 for detailed information on analysis tools

d) Digital Resources Collection

- Easy access to digital resources related to the study area
- Access to publicly unavailable digital resources, such as internal reports, project summaries, modeling results, presentations, meta-data about sampling campaigns, meeting minutes, etc.

e) Modeling

- Run statistical and hydrological models
- Advanced modeling capabilities (e.g., WWQS)
- No external software installation for modeling
- Modeling through a user-friendly web interface
- Sample case studies and modeling instructions
- Integrating other hydrological models

f) Scientific Visualization

- Easily visualize raw data, modeling outputs, simulation results
- Compare monitoring site variables (e.g., upstream and downstream water quality comparison)
- Access statistical and expert system visualization tools
- Carry out pattern recognition analysis by using Self Organizing Maps

See Chapter 3.4.2 for detailed information on analysis tools

g) Spatial Data and Maps

- Access rich geographical content from multiple content providers by using Web Services
- Integration with KML for dynamic spatial data import

5.4. GWIS for Watershed Management

Steps in the watershed planning and implementation process are listed in the EPA Watershed Handbook (2009), as follows: (a) build partnerships; (b) characterize the watershed; (c) finalize goals and identify solutions; (d) design implementation program; (e) implement watershed plan, and (f) measure progress and make adjustments.

GWIS can be fully utilized for some of these steps and make the process much easier for stakeholders. Each step of the planning process will be detailed below with accompanying GWIS features. Some of the features suggested below are not yet available in GWIS, but they are easier to be developed than the current GWIS features.

Step (a) Build a Partnership

During the involvement of the stakeholders, surveys can be completed, documents shared for review, and discussion conducted through forums, can all be carried out using GWIS. Documents and files can be shared in GWIS, for integrating local, state, tribal, and federal programs into an overall watershed planning effort. Educational resources can be shared with stakeholders through GWIS's digital resources library. Without GWIS all these steps would have to be carried out either by using several different websites, email communications or offline methods (like paper surveys, conducting meetings, distributing documents, etc.), which will increase the time, cost and efforts spent on each step.

Step (b) Characterize the Watershed

Watershed boundaries can be identified by using the topographic overlays of GWIS's mapping tools. GWIS data tools easily connect users to sources, such as EPA's sub-basin information and USGS's stream-flow data to understand and characterize the watershed. GWIS data and digital resources tools may contain existing data series and resources related to the study area. A watershed inventory can be created by collecting additional data and resources from distributed sources, which will help not only the current program but also future management programs. GWIS easily transfers daily and real-time data to the watershed inventory from sources, such as USGS and EPA. Without GWIS, all these steps require excessive manual work, such as accessing different websites, collecting data and digital resources from various sources, and data management (storing, managing, pre-processing).

It is very important to visualize multi-dimensional datasets to have a better understanding of the quality and characteristics of data before downloading them for further analysis (Zheng, 2005; Jeong and Liang, 2005). Most information systems do not provide proper tools to their users to visualize and analyze datasets (Zheng, 2005; Jeong and Liang, 2005). It is a difficult and time/resource-consuming process to download datasets and develop/access advanced data analysis and visualization tools that can be used easily with various data formats (Liu and Liang, 2003; Liu et al., 2003).

Data series can be analyzed through GWIS's descriptive statistics tool to identify gaps, so that additional data can be collected if needed. The GWIS statistical and hydrological analysis tools and visualization features can be utilized to analyze and understand the data to identify causes and sources of pollution that need to be controlled. Without GWIS, such data analysis and visualization will be a real problem, because the user will need several software packages to carry out the various different analyses, pre-process data in different formats for use in these software packages, and collect and manage data from other users or sources.

Steps (c), (d), (e) and (f) together

GWIS can provide tools for the remaining steps in digital resource management, conducting discussions, project management, training and educational purposes, collaborative plan development, sharing results, reviewing information, and in other processes that require information management, analysis, visualization and mapping.

An essential process of watershed management for assisting stakeholders is in understanding models, and effectively discussing policy and management options. This can be possible by using advanced scientific visualizations about the observations and modeling results. It is easy to generate graphs, scatter plots, histograms, trends and seasonal component simulations, statistical and expert system analysis plots, and comparisons with a user-friendly, easy web interface on GWIS. Spatial information, 3D mapping visualizations, street views, and rich geographical content can be accessed through the GWIS mapping component. Digital resources including publicly unavailable internal meta-data, reports, and research summaries will be available in GWIS. All these data, information, tools and resources make policy discussions and management easier.

In general, most of the steps in a watershed management process can be carried out by a web based information system like GWIS. It may reduce time, cost and human resources at each step. Key features of GWIS are its centralized approach to modeling, visualization, data and digital resources management, and integration of distributed resources. With the benefits of a web-based information system, as detailed in section 5.1., GWIS can make a significant difference and an improvement in watershed management. On the other hand, GWIS may not provide solutions for all of the steps of watershed management, since there might be steps to be handled manually or offline. However, all of the resources (time, cost and human resources) saved by the GWIS can be utilized on other steps carried out without an information system.

5.5. GWIS for the Chattahoochee Watershed Case Study

A comprehensive watershed study requires a large amount of information to be collected, analyzed and managed for further processing for an accurate understanding of the watershed. This would include natural, social and economic dimensions like meteorological conditions, land cover type, residence distribution, point sources, current and historical water quality conditions, existing pollution control plans, future and developing trends in the surrounding areas (Shi, 2008). This section compares how a comprehensive watershed study has been carried out with traditional computational methods and how it could be done with the help of GWIS.

Shi (2008) developed a dynamic watershed simulation system for the Chattahoochee watershed to explore the overall watershed hydrological and pollutant behavior. Some of the steps listed below and how these steps were handled for the Chattahoochee watershed case study are not clearly mentioned one by one in the dissertation (Shi, 2008). If the detail is not given for a specific step in the case study, it is assumed to have been carried out by traditional method and the comparison made accordingly. Also, GWIS is assumed to be fully populated with the necessary or planned data, tools and digital resources for the comparison.

a) Chattahoochee Watershed Characteristics

GWIS: The digital resources component can be used for gathering information, images and maps for watershed characteristics (Figure 5.1a).

Traditional: Searching multiple websites like USGS, EPA and online resources to collect this information is required (Figure 5.1b).

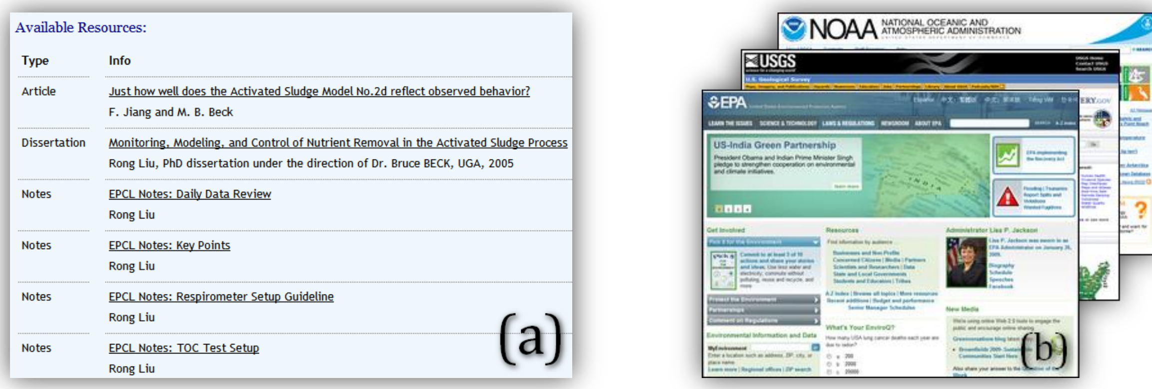


Figure 5.1. (a) GWIS digital resources component, (b) Multiple websites

b) Pictures and Maps of the Study Area

GWIS: The mapping and digital resources component can be used for pictures and maps of the study area (Figure 5.2a).

Traditional: Maps can be found through searching multiple web sites (Figure 5.1b) and using complex GIS software (Figure 5.2b).

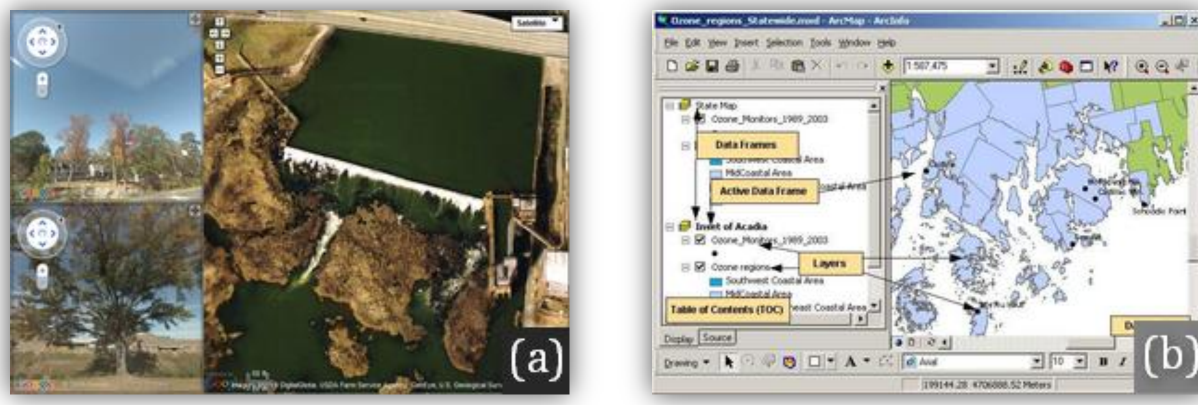


Figure 5.2. (a) Panoramic visuals and satellite maps, (b) GIS software

c) Land Use / Land Cover Spatial Data and Maps

GWIS: The mapping component can be used to transfer spatial data and maps from USGS land cover web services (Figure 5.3a). KML layers can provide land use and cover overlays on GWIS maps.

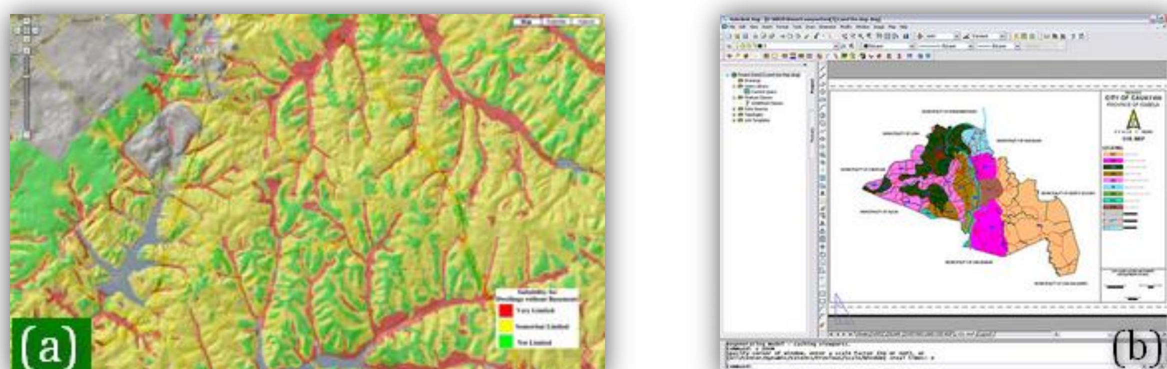


Figure 5.3. (a) Land use overlay on GWIS maps, (b) GIS software with land use map

Traditional: Downloading and using complex GIS software to access USGS land cover database is required (Figure 5.3b).

d) Meteorological Data Collection

GWIS: GWIS data tools can be used to transfer meteorological data from USGS weather station web services to GWIS with a simple interface (Figure 5.4a). Transferred data can be used without any pre-processing on GWIS.

Traditional: Data are collected from USGS weather stations manually by using complex user interfaces and preprocessed to visualize and analyze the raw data (Figure 5.4b).

Figure 5.4. (a) Simple interface to transfer meteorological data, (b) USGS interface

e) Environmental Management Issues of Study Area

GWIS: The digital resources component is available for related journal articles, reports, dissertation and other documents about the study area. Access to unpublished resources (not available on the web), such as internal reports and meta-data, project plans and summary, and others is available by GWIS

(Figure 5.5a). Spatial features on GWIS helps users filter related content easily for their study area.

Traditional: Search from scientific databases (Figure 5.5b) for articles, using search engines (e.g. Google) for reports and other related documents, dissertation databases and document related sources. Spatial filtering is not available for most of the scientific databases.

Figure 5.5. (a) GWIS digital resources interface, (b) Scientific database interface

f) Annual Inflow and Average Concentration of Phosphorus Calculation

GWIS: User-uploaded data can be used from GWIS directly or easily transferred from USGS (Figure 5.4a). No data preprocessing is required, user needs to run selected models, visualize outputs (Figure 5.6a) and export results.

Traditional: It is required to search and download data from USGS or other sources (Figure 5.4b), find out and learn a calculation method, prepare data for calculation, calculate manually or by a software package (Figure 5.6b).

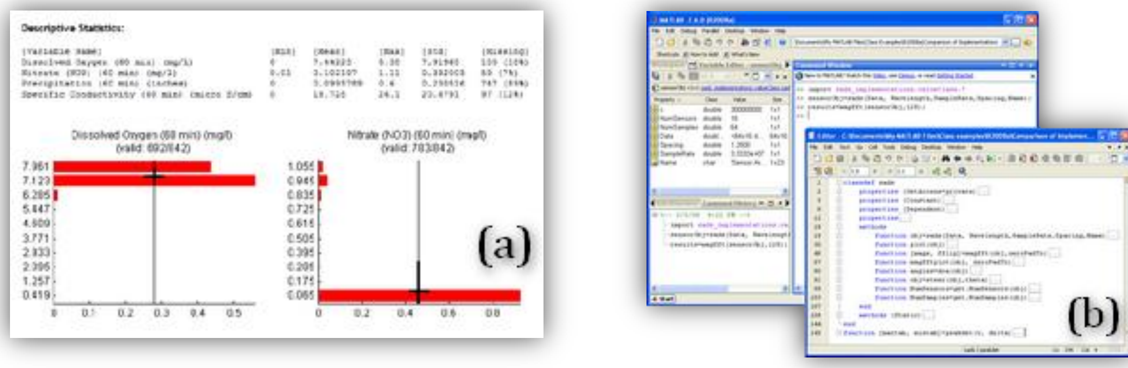


Figure 5.6. Statistical Analysis (a) with GWIS, (b) complex Matlab interface

g) Data Collection - Modeling Framework

In the case study, the Modeling Framework steps explained here are repeated for point source assessments, non-point source assessments, and in-stream simulations. This comparison applies to all of the repeated modeling processes.

GWIS: One of the most distinctive advantages of GWIS is having high-volume high-quality data set (e.g. EPCL data sets), which may not be available for most of the other information systems. Users can also transfer data from USGS (Figure 5.4a) directly to GWIS servers for further analysis.

Traditional: Data is collected from USGS manually (Figure 5.4b) and saved to computer for further analysis in the following steps.

h) Data Pre-processing - Modeling Framework

GWIS: No format change for pre-processing is necessary for the data available in GWIS or transferred to GWIS. The descriptive statistics tool (Figure 5.6a) gives a quick overview of the data. If required, missing data can be completed by using dynamic harmonic regression or Self Organizing Maps.

Traditional: After searching and downloading data from USGS (Figure 5.4b), data should be pre-processed to be used for various calculations and visualizations. Missing data needs to be completed with software (Figure 5.6b). All of these steps might require different data formats and preparation depending on the software to be used (Figure 5.7b).

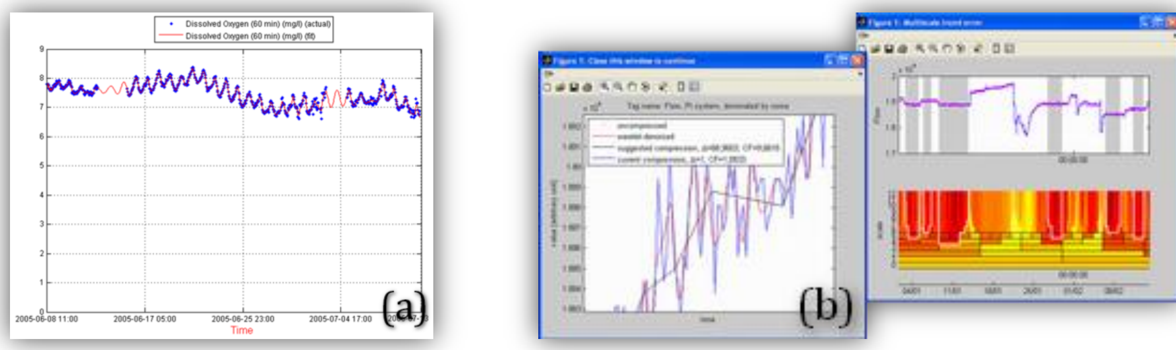


Figure 5.7. (a) Completing missing data with GWIS, (b) data preprocessing software

i) Data Examination - Modeling Framework

GWIS: Data can be visualized easily with GWIS's visualization tools (Figure 5.8a). Trend and seasonal component analysis will help extracting additional details from the data. Statistical and expert system visualization tools might give a better understanding of the data and the study area.

Traditional: Data can be visualized with Excel (Figure 5.8b) or other visualization software. Trend and seasonal component analysis might be required. Every additional analysis might require additional processing of data and knowledge of related software packages. In the case study a box-plot is used to identify outliers.

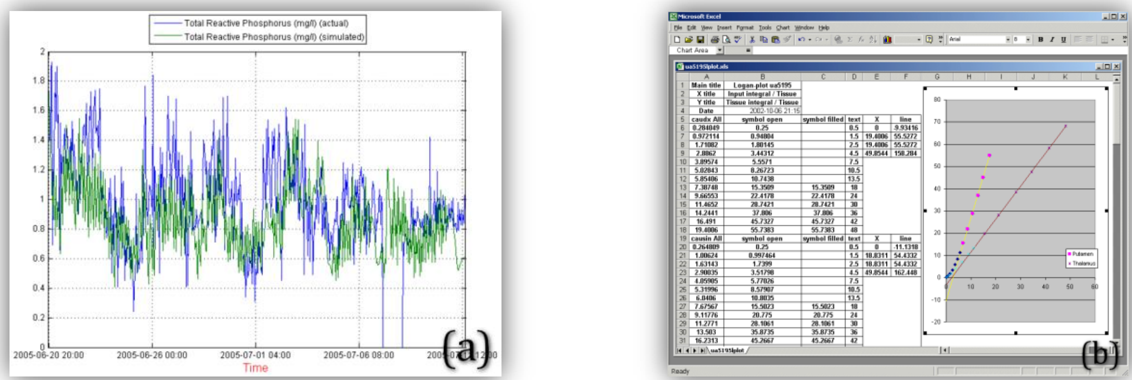


Figure 5.8. (a) GWIS visualization example, (b) MS Excel for visualization

j) Running Models - Modeling Framework

GWIS: Assuming the availability of the required modeling frameworks or interfaces through GWIS, it is very easy to run models within a web interface. No data discovery, downloading, processing, and installation of models are required. Users can run a model (Figure 5.9a) with a couple clicks with user-friendly instructions. Simulation results can be exported easily to be used for further analysis.

Traditional: There are a lot of challenges for a regular user in this step, such as finding the appropriate model, downloading it, providing the platform requirements for the model, installation of the model, consulting with an expert on model, preparing data for the model (Figure 5.9b), exporting results.

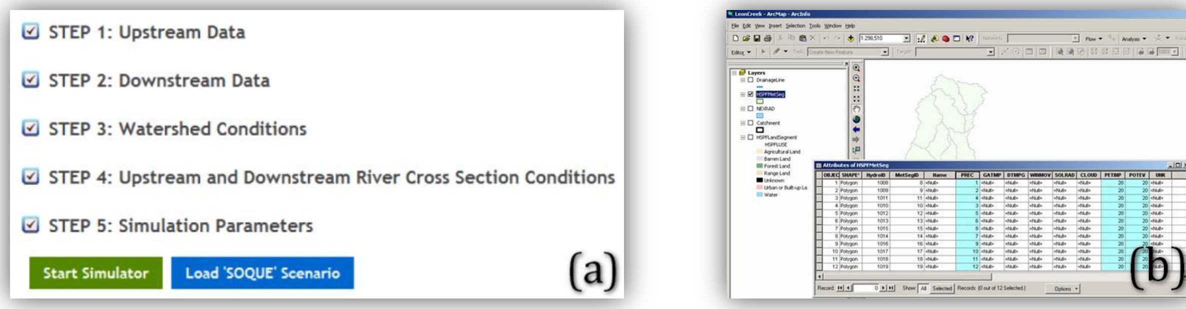


Figure 5.9. (a) GWIS watershed water quality simulator, (b) ArcGIS HSPF interface

k) Generating Reports - Modeling Framework

GWIS: Simulation results from models can be visualized (Figure 5.8a) and exported easily, and reports (Figure 5.10a) are ready to print as a PDF file or on printer.

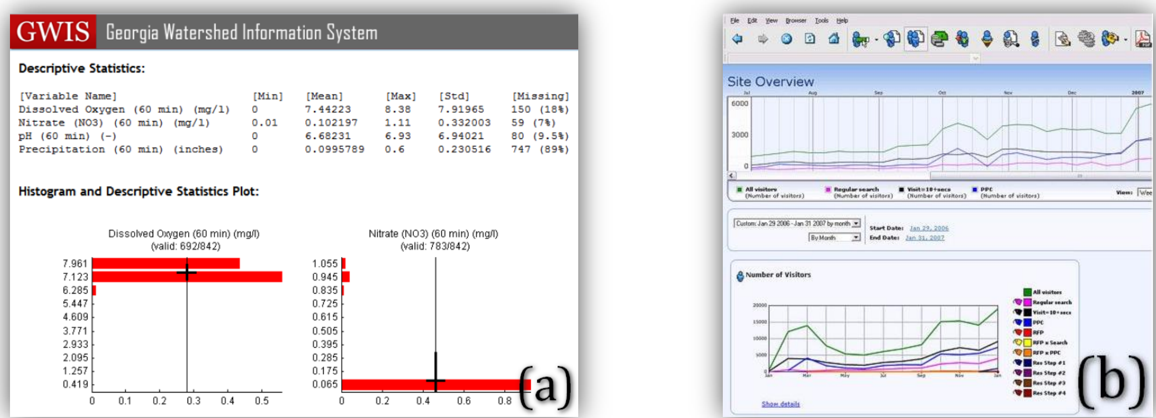


Figure 5.10. (a) GWIS analysis report, (b) report generation by other software

Traditional: After gathering results from the model, outputs should be visualized (Figure 5.8b) and prepared for printing reports by additional software (Figure 5.10b) which might require additional output processing.

5.6. Conclusion

This chapter has presented the advantages of GWIS as an information system in general, and specifically for environmental research and watershed management cases, with a step by step comparison of a watershed study with and without the presence of GWIS. Some of the steps for the case require accessing multiple websites to gather information, and the use of several software packages to visualize and analyze the data, which will increase the time and resources required in the process. When compared to traditional approaches, GWIS provides advantages in respect of ease of use, saving time and resources through centralized access, and minimizing expert requirement. However, as a prototype, GWIS does not provide all available or necessary tools or features for environmental research; further development is required to make the required tools available in the system.

One of the most important advantages of GWIS is combining data collection (download and transfer), pre-processing (changing format and structure), and examining data (visualization and analysis) in one interface, which is otherwise a multistep process requires various information system tools, expertise, and time. As a conclusion, GWIS can make watershed research and management easy, user-friendly, and more efficient with its novel and distinct features powered by the latest web technologies. Since internet technologies are evolving rapidly, it is important to automatically adapt an information system to the new developments. The next chapter discusses adapting and operating GWIS in the light of new technology.

CHAPTER 6

ADAPTING AND OPERATING GWIS IN THE LIGHT OF NEW TECHNOLOGY

6.1. Introduction

The development of web-based information systems is a dynamic process of continuous improvement. There are a number of driving forces for the development of web-based systems, such as distributed information, developments in internet technologies and the need for application portability (application virtualization or application as a service). However, the adaptation and operation of information systems in the light of new technologies is not adequately discussed (Suchman, 1987; Ciborra, 1994; Kyng and Mathiassen, 1997; WHO, 2000; Lenz and Kuhn, 2004).

This chapter presents issues and methods for allowing GWIS to be systematically adapted and operated (over the years) in order to accommodate and benefit from the furious pace of change in the technologies supporting and defining GWIS. The adaptation process is discussed with references to related components and layers in the GWIS architecture (Figure 6.1) for the following aspects; presentation, scalability, reusability and maintainability. A special emphasis is given to Web Services (WS) and Semantic Web (SW) technologies. Figure 6.1 presents GWIS elements corresponding to related adaptation issues in the architecture and the speed of change in technology is given in a colored background frame. While dark blue color represents a faster change in technology, lighter colors represent slower changes.

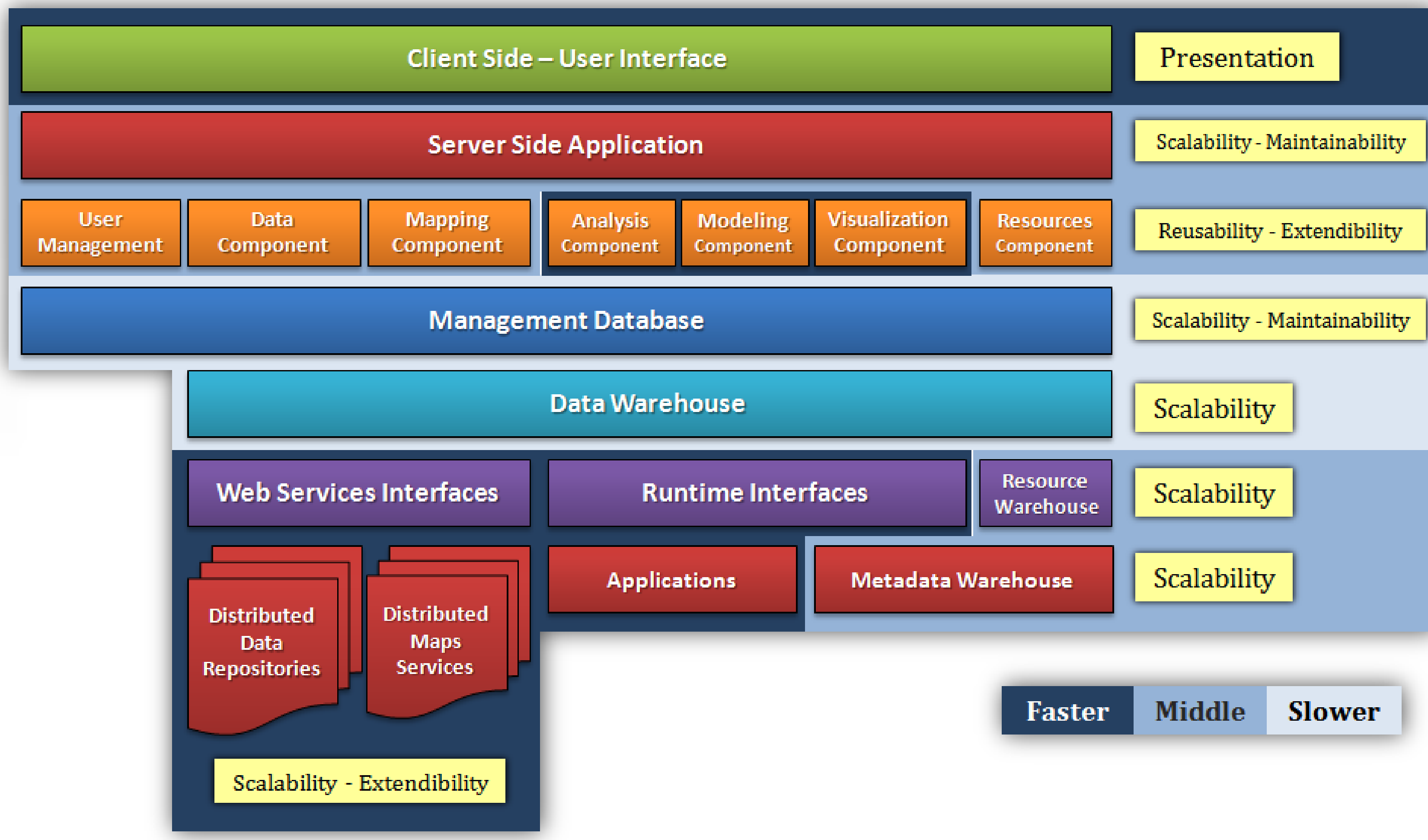


Figure 6.1. Architecture of GWIS with adaptation issues and change in technology

6.2. Presentation

GWIS uses a common approach based on the latest web development trends to separate the content and presentation (navigation and layout) layers in the development process. Content layer refers to data and information gathered from system users and distributed sources. Navigation layer defines the application logic of the tools in the system or whole system. A layout layer deals with the design related issues in the system.

Having a three layer design process helps GWIS in handling adaptation separately for each layer. Adaptation of the content layer will be discussed in the scalability section below. Currently GWIS provides various navigation options to users, such as web forms, interactive mapping interfaces, and hybrid approaches. The design of GWIS uses a central definition of style format, header and footer files. Clear separation of the content and a centralized design approach provides flexibility to change and adapt the layout and navigation structures.

The presentation layer is subject to faster changes in technology (dark blue frame in Figure 6.1). The size and capabilities of devices that are accessing internet are changing day by day. While the size of devices is changing from a desktop or notebook to netbook or mobile, the functionality is increasing from basic mobile operations to full desktop capabilities. The characteristics of these devices are varying over such a wide range that they require changes in not only presentational aspects but also navigational and layout aspects in order to deliver information and services from web-based systems.

It is easy to add new navigation structures and features to GWIS by simply defining the structure and making the necessary database connections with content. The design of GWIS can be adapted to properly function on devices with different screen sizes by changing the central style definition, header and footer files. In traditional information systems, components usually have specific requirements to run on desktop systems. They should be re-designed specifically for the selected system according to new requirements and capabilities. It is usually a manual process requiring a lot of resources.

6.3. Scalability

Scalability is one of the most common issues in information system design. It is subject to mid-level or long-run changes in the technology (Figure 6.1). It can be related to an increase in the number of users and content. An increase in the number of users can be handled by upgrading the hardware that manages the required network load. Also the design of the system can help in decreasing the load by automating some of the processes and providing mechanisms like caching and distributed data access. An increase in the content might require more hardware resources for storage.

As a prototype information system, GWIS has not been designed to support large scalability issues in its current structure. However, GWIS uses PHP scripting language and MySQL database system on the server, which are the two common technologies that are in use for large and scalable systems handling millions of users and large content resources.

GWIS utilizes web services to provide users with an alternative method of accessing content, which decreases the network load by sending only structured content to the users without any heavy design and content layout. Second, the use of web services in GWIS allows access to distributed data and content from various sources. This is a way of increasing the content and data of the system, while distributing the network and storage load to other systems.

Web services can be used in other ways to help scalability. They can be used for sharing structured content, manually integrating distributed sources into the system, and automating integration of data and tools with an API. Other aspects of content are meta-data and vocabularies, which can utilize semantic web for technological adaptations. The following sections discuss web services and semantic web technologies with respect to scalability and adaptation issues.

6.3.1. Web Services

GWIS utilizes Web Services in its data management and mapping components. The data management component has two integration interfaces to access and transfer data from USGS real-time and daily data sources to GWIS, so that users can either import, analyze and visualize USGS data directly within GWIS. However, GWIS can benefit more from Web Services technologies in various ways. Some of these approaches can be as follows:

a. ***Integration with additional Web Services:*** GWIS has been integrated with two USGS data sources. Additional data sources from USGS, Environmental Protection Agency

(EPA), National Oceanic and Atmospheric Administration (NOAA), National Oceanographic Data Center (NODC), and others can be integrated into GWIS. Web Services (WS) studies for modeling and visualization are very limited in environmental science. There is a total of 135 articles only from 2000 to 2009 in ScienceDirect, one of the largest scientific databases. The next component for WS integration is online mapping. GWIS has been interfaced with four powerful WS (Google, Microsoft and Yahoo Maps APIs) for mapping features. It should be scalable for integration with other mapping services.

One problem with distributed sources like USGS is changes in their access methods or specifications, which will break the connection between other information systems. Actually, this problem occurred two times during the development of GWIS. The USGS changed or removed some of the parameters from the connection URL. The ideal way of making changes to a web service is providing a backward compatibility to older access methods. For example, GWIS uses Google Maps Service, which has been updated a couple of times during the development process. But it never caused a problem because Google still supports the older access methods to their web services. This issue should also be considered carefully, and connection procedures for sources such as USGS should be checked periodically.

b. ***Interface for web services integration:*** Web Services integration can be automated by building an interface, so that manual integration of available data sources can be done by users. The GWIS mapping component can benefit also from such interfaces to increase the number of available map services to users.

c. ***Web Services for GWIS data and tools:*** Since GWIS has been populated with high-volume high-quality water quality data acquired during field monitoring campaigns between 1998 and 2008 with the Environmental Process Control Laboratory (EPCL), having a web service to make data available for other information systems makes GWIS data more accessible to the public. GWIS currently provides two web services to share water quality data and site description with users. Two of the novel features of GWIS are modeling and visualization capabilities. Users can carry out visualizations and run models for the data uploaded to GWIS or the ones transferred from distributed data sources. These modeling and visualization tools can be made fully available for other information systems through a web service from GWIS. Sharing available data and tools over a web service can attract more users and developers to GWIS, which will extend its use and keep it live.

Even making some of these approaches partially available in GWIS makes it an effective and widely acceptable open collaboration and community platform.

6.3.2. Semantic Web

Semantic Web (SW) is defined by the Semantic Web Agreement Group (SWAG, 2001) as, "a web that includes documents, or portions of documents, describing explicit relationships between things and containing semantic information intended for automated processing by our machines". Ontology is a formal specification of a shared conceptualization (Gruber, 1991). Ontologies provide an explicit and declarative description of a domain. Often environmental databases make datasets available with little known information, such as file names, certain identification and distribution details. Resource Description Framework (RDF), a standard model for data interchange on the web, can be used to describe metadata. Most of the time, this information is not enough to understand the usefulness of a dataset for a specific analysis before actually downloading and looking at the data. Semantic information and relations are needed to help users make decisions about the data available for their research.

Semantic web technology can be used to provide a platform to allow users to submit formal descriptions of concepts, terms, and relationships within watershed management, which will increase the understanding of the concepts and tools in GWIS and keep the system up-to-date within the selected domain. Semantic web technologies have been applied widely in many disciplines outside information technology, including linguistics, philosophy, library science and life science.

There are examples of studies in environmental science including ontology-based discovery of geographic information services (Klein, 2005), semantic approaches to environmental modeling (Villa, 2008), and an integrated system for publishing environmental observations data (Horsburgh, 2009). However there are very few studies on watershed data, such as a streaming data ontology to handle temporal proximity concepts (Liu, et al., 2008), the NASA semantic web ontologies for earth and environmental terminology (Raskin, 2010), and mapping hydrological variables to ontology concepts (CUAHSI HIS, 2008).

In particular, a connection from data to a description of their meaning (e.g. meta-data) is extremely important for data interoperability. Sharing a common description of the meaning of the data (a common ontology) and having a common protocol for the interconnection are key issues for correctly sharing the data. Currently, semantic features are not available in GWIS. Instructions are available for all GWIS features to build a standard format for data and meta-data. Most of the inputs asked of users for data entry are in widely acceptable standard formats like the USGS site number, HUC number, latitude, longitude, etc. An auto-complete feature is available for fields like variable name, unit, and frequency, to provide consistency. GWIS can benefit from semantic web technologies in various ways for adapting it to new technologies. Some of these approaches can be as follows:

a. ***Publish data in RDF:*** RDF has features to merge data even if the underlying schemas differ, and it supports the modifications on schemas over time without requiring all the data consumers to be changed. It can be used to define and publish meta-data for water quality observations. This will improve interchangeability and accessibility of GWIS data between information systems.
b. ***Semantic integration of distributed sources:*** This will increase the value of otherwise isolated environmental information as in data, meta-data, and digital resources.

c. ***Semantically defined interfaces:*** This will enable a flexible, future-oriented extendible integration framework for environmental research and collaboration.

However, there are several research challenges for establishing semantic web technologies such as: (a) high cost of existing technologies for semantic integration; (b) low availability of best practices in environmental science domains; (c) the need for efforts to setup a semantic framework (ontology creation process); and (d) the challenge of establishing an easy-to-use access to a complex semantic framework.

6.4. Reusability and Extendibility

Modular design of a system increases the extent to which components represent a separation of functionality, and improve maintainability by providing boundaries between components. GWIS consists of six main components and many tools under each component.

Reusability can be discussed for the system as a whole or for the components in the system. GWIS has been developed as a prototype for the management of watershed related data and resources for the state of Georgia. However none of the features and structure elements limits its use for other states or geographic regions. The system can be modified and rebranded for other watershed projects easily by changing only the presentation layout (central style definition, related text and files).

The current structure of GWIS does not limit its use for purposes other than watershed-related data management. It can accept any time series data from other environmental or ecological areas. Most of the analysis and visualization tools are for general use. The system can be modified and rebranded easily for other projects or fields by changing only the presentation layout and removing watershed-related analyses and models. New tools and features can be added to extend the use of GWIS for different environmental research areas.

Individual components are designed in a modular structure. GWIS features hybrid applications of visualization and mapping features combining separate components. Each component or tool can be reused for future adaptations or hybrid applications in a project but also between projects.

6.5. Maintainability

It is usually difficult to maintain and expand web-based information systems than the first development process (Ginige, 2001). The increase in the number of users and content, scalability issues, moderation of the system, and new requirements increase the load on routine maintenance. The main reason for this is not having internal IT staff to develop the system and hosting the system on an outside provider (Marmaridis, 2004). It is important that an information system is developed to become economical in its use of resources for its maintenance, operation, and use. Maintainability of the system should be defined clearly in the developmental process.

Most of the tools and services provided in GWIS are automatic and require minimal maintenance. Web Services integration to distributed sources provides not only data and content but also decreases the load on the network, storage and maintenance for integrated sources. User uploaded data and content require reviews and checking of resources from time to time. If the system is in use by a limited or trusted user group, the maintenance will decrease for content reviews. GWIS is not designed for a large user base, so expansion of the functions or scale might require modification of the system.

6.6. Conclusion

There are many factors that can affect the failure and success of an information system, such as the planning process, adaptability, scalability, reusability, extendibility and maintainability. This chapter has discussed issues and methods for allowing GWIS to be systematically adapted and operated in order to keep it live and current in the light of new technology. Development of information systems is a dynamic process of continuous improvement that needs to be planned during the development process to accommodate issues of scalability, reusability and maintainability.

The problem with contemporary information systems is their rapid development process and lack of interdisciplinary collaboration on planning and development. The rapid development process depends on minimal planning in favor of rapid prototyping. The planning process is usually interchanged with developing the system itself in rapid development. The lack of extensive pre-planning generally allows the system to be developed much faster, but it does not provide the flexibility to handle issues of adaptability for future needs and requirements. A comprehensive information system requires a team of people from various disciplines to accommodate the need of understanding of the issues that will adapt the system in the light of new technological developments.

While some of the design approaches have already been implemented in GWIS to overcome the adaptation problems of contemporary information systems, some of these approaches are suggested for future research. Web Services and the Semantic Web are two important technologies that will help GWIS in this dynamic adaptation process. The next chapter presents conclusions of the research as a whole and recommendations for future research in the area of information systems.

CHAPTER 7

CONCLUSIONS AND RECOMMENDATIONS

7.1. Discussion and Conclusions

Information technologies are experiencing rapid progress. There is a significant opportunity for information systems to contribute to watershed management. Web-based information systems can help in understanding watershed systems, with the integration and analysis of local and distributed data and digital resources. However this requires advanced data management, visualization and analysis capabilities. New technologies and techniques even appear during the long development process of an information system. They may provide a better, faster, or cost-effective alternative to technologies for the planned system, which requires designing the system so that it can be easily adapted in the light of new technology.

A review of the literature on information systems for watershed management shows that they are either implemented on local infrastructures for offline use, or focused on a specific problem of watershed management, and do not utilize the latest web technologies. In the larger context, the EU WFD and US TMDL programs give us an idea about the missing points, needs and trends in Information Systems for environmental studies. No research was found that provides a comprehensive framework for Watershed Information Systems with high volume data management, interactive scientific visualization, statistical analysis, expert systems modeling, dynamic mapping, and digital resources components.

This dissertation presents a prototype of and a web-based information system, called the Georgia Watershed Information System (GWIS), to support the integration, dissemination, and exploration of environmental data for Georgia watersheds. GWIS has been initialized with several tools and data collected with the Environmental Process Control Laboratory (EPCL). Once fully populated with increasingly varied content and capabilities, GWIS promises to be extremely beneficial to research and management in Georgia watersheds, and the members of the general public. Significant contributions made in this dissertation are summarized as follows:

- ✓ GWIS provides a platform to help watershed research by decreasing time and resources on data management, collection, pre-processing and examination by web-based tools and interfaces.
- ✓ GWIS extends current watershed information systems by providing online analysis, scientific visualization, modeling and mapping capabilities.
- ✓ GWIS provides an alternative information interface for USGS monitoring sites. The interface has an advanced navigation method that allows users to locate an observation site on an interactive map, access site related information and data, and carry out visualizations and analyses directly on the interface.
- ✓ GWIS integrates mapping, visualization and analysis capabilities to provide interactive scientific visualization interfaces to increase public awareness on environmental issues.
- ✓ While implementing the latest web technologies for watershed information systems, GWIS is an example of how an information system can be designed to be adaptable to continuous improvement in web technologies.

- ✓ GWIS features a smart interface designs that provide novel and interactive navigation structures, and easy user input methods and advanced user guidance.
- ✓ GWIS provides user-friendly interfaces to integrate data from distributed data sources by using web services.

7.2. Recommendations on Future Research

Environmental information systems, and web-based systems in particular, have been discussed and a prototype demonstrated in this dissertation, which is important and beneficial for supporting watershed research and management. This dissertation has a focus on a number of main issues on this context. Further research is recommended in this area along the following lines.

- ✓ Developing a prototype information system is the first step for web-based watershed management and research. Further research is required to launch the system with the public, publicize it to potential users, operate, maintain, and adapt it in the light of user feedback, and update it accordingly. It is important to have proper instructions, help, contact information and a feedback mechanism in the system. When the system is up and running, there are several ways to spread the word. Some of the examples are listed below:
    - Announcing to the research team, former students, colleagues, departments, academic groups and faculty contacts in different institutions.
    - Sending a press release to departments, universities and local newspapers and magazines.

- Sending a technical note to related journals and scientific databases.
- Presenting the system at various conferences, seminars, and workshops.
- Reaching federal, state, policy, management stakeholders and public.
- Sending posts to university, department, scientific, and local blogs.
- Providing tools on the system that makes it easy for users to spread the word about the system (e.g. email forms, social network tools).
- Improving rankings of the system on major search engines (e.g. Google, Bing) to appear on top for specific keywords.
- Opening pages on social networks to spread the word and creating a network of users for the system.
- Increasing communication with users by creating a blog for the system, and using social networking tools (e.g. Twitter, Facebook).

✓ The performance of GWIS when compared with traditional information system approaches to data management, visualization, analysis and modeling can be evaluated quantitatively by using international standards. This can be used for defining and evaluating the usability of GWIS as an information platform and interactive system. GWIS and alternative systems/tools can be tested with a sample of users performing a set of pre-determined tasks to yield the most reliable and valid estimate of the systems' usability via surveys. There are also other methods for further research to evaluate information systems such as expert-based (Human Computer Interaction expert examines the system and estimates its likely usability for a given user population) and model-based methods.

- ✓ While GWIS is primarily developed with a research focus, it can serve a management-oriented function for watersheds with further research. Information systems (IS) have had a great effect on traditional management functions and participation. Management of systems relies less on manual processes to perform everyday actions. IS have automated many of the key management activities. For example, communication changed a lot with email, web-based conferencing, desktop sharing and video calls. Online systems allow easy access to data, information and analysis tools. Because of the online and easy access to data, managers are feeling the pressure to better manage critical information.

- ✓ While providing the design of a comprehensive information system, this dissertation does not focus much on the issues of scalability. Methods and design strategies for handling the increasing number of data, integrated sources and users, is a target for future research.

- ✓ GWIS is designed in a flexible structure to support adaptation to new developments in internet technologies. However, strategies and issues outlined in Chapter 6 for procedures to adapt an information system in the light of new technologies can be further studied extensively for each component or issue in particular.

- ✓ Data quality is a multi-dimensional concept and has many dependents such as individuals involved with the data and measurements based on the data set. GWIS does not provide methods necessarily to handle data quality issues to limit the focus of dissertation. Data quality assurance and data metrics are obviously an area for future direction of research. Currently, assuring the quality of the data in GWIS is left to users who uploaded to the system in the first place and who want to use them in their studies.

- ✓ This dissertation does not focus extensively on discussing and developing data and meta-standards. There are ongoing studies in these areas; however we are far from having a universal or widely accepted standard for water quality data and meta-data. This would be an important direction for future research.

- ✓ The structure of GWIS is open to implement more integrated sources. However, automation of this process by providing methodologies and smart interfaces with the help of web services technologies and APIs can be further studied. This approach can be used to provide an integrated analysis and modeling environment to third parties who want to share their analysis tools on the system.

- ✓ Scientific visualization and analysis tools are limited with the examples presented in the prototype. New visualization and analysis tools for specific purposes in watershed management, and an open integration system to provide mechanisms to connect these tools to web-based information systems can be a future direction on this area.

- ✓ GWIS integrates web-based mapping systems from Google and third-party systems. This approach can be extended by studying web-based GIS systems that can provide advanced mapping and spatial analysis capabilities.

- ✓ The semantic web and systems are promising areas in information systems. They are not widely studied in environmental science. Semantic information and relations have a vital role to help users to access related data and information, and make decisions about the data available for their research. A connection from data to a description of their meaning is important for data interoperability. Meta-data and vocabularies can find use of semantic web for technological adaptations. This would be an important direction for future research.

REFERENCES

Ahlers, D., Finke, J., Stadler M., and Sonnenschein, M., 2005. A web-based reviewing process guidance system for an ecological database of plant traits. In: Hrebicek, J., Racek, J. (Eds.), *Proceedings of the 19th International Conference Informatics for Environmental Protection - Enviroinfo Brn*. Masaryk University, Brno, 80-84.

Al-Ghadban, A. N., 1997. Towards a GIS-Based environmental information system for Kuwait, *GIS Conference 1997*, Qatar.

Alhoniemi, E., 2002. Unsupervised Pattern Recognition Methods for Exploratory Analysis of Industrial Process Data, *PhD Dissertation*, Helsinki University of Technology.

Alonso, G., Casati, F., Kuno, H., and Machiraju, V., 2004. Web Services. Concepts, Architectures and Applications. *Springer*, Berlin/Heidelberg.

Alves, C. D. A., and Loucks, D. P., 2007. Development and application of a Watershed Information System (WIS) for water quality analyses, *Water Quality and Sediment Behaviour of the Future: Predictions for the 21st Century* (Proceedings of Symposium HS2005 at IUGG 2007, Perugia, July 2007). IAHS Publ. 314, 2007, 251-264.

Anderson, G., and Moreno-Sanchez, R., 2002. Building web-based spatial information solutions around open specifications and open source software. *Transactions in GIS,* **7,** 4, 447-466.

Argent, R.M., and Grayson, R.B., 2001. Design of information systems for environmental managers: an example using interface prototyping. *Environmental Modelling & Software,* **16**, 433–438

Arkansas Watershed IS. Available online http://watersheds.cast.uark.edu. Last accessed September 2008.

Aronoff, S., 1989. Geographic Information Systems: A Management Perspective. *WDL Publications*, Ottawa, Canada.

Aydiner, C., Demir, I., Keskinler, B., and Ince, O., 2009. Joint Analysis of Transient Flux Behaviors via Membrane Fouling in Hybrid PAC/MF Processes using Neural Network. *Desalination*, **250**, 1.

Aydıner, C., Demir, I., and Yıldız, E., 2005. Modeling of Flux Decline in Crossflow Microfiltration Using Neural Networks: The Case of Phosphate Removal. *Journal of Membrane Science*, **248**, 1-2, 53-62.

Band, L.E., 1989. A terrain based watershed information system. *Hydrological Processes*, 3, 151-162.

Bayar S., Demir, I., and Onkal-Engin, G., 2009. Modeling Leaching Behavior of Solidified Wastes Using Back-Propagation Neural Networks. *Ecotoxicology and Environmental Safety*, **72**, 3, 2009.

Beck, M. B., Jiang, F., Feng, S., Walker, R.V., Osidele, O. O., Lin, Z., Demir, I., and Hall, J., 2010. Re-Engineering Cities as Forces for Good in the Environment. *Proceedings of the ICE - Engineering Sustainability (in press)*.

Belfiore, S., 2002. Introduction to the special issue on capacity building, From the 1992 Earth Summit to the 2002 World Summit on Sustainable Development: continuing challenges and new opportunities for capacity building in ocean and coastal management. *Ocean & Coastal Management,* **45**, 541-548.

Bianconi, R., Galmarini, S., and Bellasio, R., 2004. Web-based system for decision support in case of emergency: ensemble modelling of long-range atmospheric dispersion of radionuclides. *Environmental Modelling & Software,* **19,** 4, 401-411.

Blomqvist, A., 2004. How can stakeholder participation improve European watershed management: the Water Framework Directive, watercourse groups and Swedish contributions to Baltic Sea eutrophication. *Water Policy,* **6,** 39–52.

Boner, M., Gurney, P., Bowman, J., Birkhead, B., Jurgensen, M., Beck, M. B., Villarroel Walker, R., Shi, F., Gore, J., and Banning, J., 2008. Demonstration of Disinfection and Pollutant Load Reduction BMP in Weracoba Creek Watershed, *Technical Report*.

Briggs, D.J., 1995. Environmental statistics for environmental policy: genealogy and data quality. *Journal of Environmental Management*, **44**, 39-54.

Brunt, J.W., 2000. Data management principles, implementation and administration. In: Michener, W.K., Brunt, J.W. (Eds.), *Ecological Data: Design, Management and Processing*. Blackwell Science, Oxford, 25-47.

Burrough, P.A., and McDonnell, R.A., 1998. Principles of Geographical Information Systems. *Oxford University Press*, Oxford.

Busse, K. M., 1998. Tillamook Bay National Estuary Project Watershed Information System, Research Report, *Oregon State University*.

Carleton C. J., Randy, A. D., and Kenneth, W. T., 2005. A relational database for the monitoring and analysis of watershed hydrologic functions: I. Database design and pertinent queries. *Computers & Geosciences,* **31,** 393–402.

Chen, C. W., Herr, J., and Weintraub, L., 2004. Decision support system for stakeholder involvement. *Journal of Environmental Engineering*, **130,** 6, 714-721.

Ciborra, C., 1994, The grassroots of IT and strategy. In Strategic information systems—A European perspective, eds. C. Ciborra and T. Jelassi, 3–24. Chichester: Wiley.

COMPRIS, 2005. Final Conference October 24-27, 2005. COMPRIS Internet Site – Products - Deliverables. Available online http://www.euro-compris.org/content/ascii/. Last accessed September 2008.

Crompvoets, J., Bregt, A., Rajabifard, A., and Williamson, I., 2004. Assessing the worldwide developments of national spatial data clearinghouses. *International Journal of Geographical Information Science,* **18,** 7, 665-689.

CUAHSI-HIS. Available online http://his.cuahsi.org. Last accessed September 2008.

Culshaw, M.G., Nathanail, C.P., Leeks, G.J.L., Alker, S., Bridge, D., Duffy, T., Fowler, D., Packman, J.C., Swetnam, R., Wadsworth, R., and Wyatt, B., 2006. The role of web-based environmental information in urban planning—the environmental information system for planners. *Science of the Total Environment.* **360,** 233– 245.

Demir, I., 2003. Fuzzy Logic and Neural Network Applications in Environmental Science, *M.S. Thesis*, Gebze Institute of Technology, Kocaeli, TURKEY.

Demir, I., and Beck, M. B., 2009a. Dynamic Spatial Information Interface for USGS Monitoring Sites, *Southern Forestry & Natural Resource Management GIS Conference*, December 7-9, Athens, GA, USA.

Demir, I., and Beck, M. B., 2009b. GWIS: A Prototype Information System for Georgia Watersheds, *Georgia Water Resources Conference: Regional Water Management Opportunities*, April 27-29, UGA, Athens, GA, USA.

Demir, I., Jiang, F., Villarroel Walker, R., Parker, A.K., and Beck, M.B., 2009c. Information systems and social legitimacy. *Proceedings of Annual IEEE Conference on Systems, Man and Cybernetics*, 1093-1098, October 11-14, 2009, San Antonio, TX, USA.

Denisov, N., and Christoffersen, L., 2001. Impact of environmental information on decisionmaking processes and the environment. *UNEP/GRID-Arendal Occasional paper,* **1**. 2001, 50.

Denzer, R., 2005. Generic integration of environmental decision support systems - state-of-the-art. *Environmental Modelling & Software*, **20**, 10, 1217-1223.

Denzer, R., Riparbelli, C., Villa, M., and Güttler, R., 2005. GIMMI: Geographic Information and Mathematical Models Inter-operability, *Environmental Modelling & Software*, **20**, 12, 1478-1485.

Denzer, R., Swayne, D.A., and Schimak, G., 1997. Environmental Software Systems, **2**. *Chapman and Hall*, London.

Díez, E., and McIntosh, B. S., 2009. A review of the factors which influence the use and usefulness of information systems. *Environmental Modelling & Software*, **24**, 5, May 2009, 588-602.

Dublin, 2004, Common Implementation Strategy for the Water Framework Directive Principles and Communication of Results of the First Analysis under the Water Framework Directive. Policy Summary and Main Document. 22/23 June.

Dupmeier, C., and Geiger, W., 2006. Theme Park Environment as an example of environmental information systems for the public. *Environmental Modelling & Software,* **21,** 11, 1528-1535.

Eimer, U., and Krause, S., 2005. European waste sector assistant - EUWAS - implementation of a web-based knowledge management system for the European waste sector. *Proceedings of the 19th International Conference Informatics for Environmental Protection*, September 7-9. Masaryk University in Brno, Czech Republic, 406-410.

EPA, California, Total Maximum Daily Loads Questions and Answers. Available online http://www.swrcb.ca.gov/tmdl/docs/tmdl_factsheet.pdf. Last accessed December 2007.

EPA, Environmental Protection Agency, Watershed Handbook. Available online http://www.epa.gov/nps/watershed_handbook/. Last accessed November, 2009.

EPA, STORET. Available online http://www.epa.gov/storet/dbtop.html. Last accessed May 2009.

European Commission, 2000. Directive of the European Parliament and of the Council 2000/60/EC establishing a Framework for Community Action in the Field of Water Policy. 1997/0067(COD) C5-0347/2000 LEX 224.

European Commission, 2005a. River information services, as policy implementation flows from research. *EXTR@Web Consortium for DG Energy and Transport*. European Commission, Belgium, 19.

European Commission, 2005b. WISE Newsletter, The bulletin of the Water Information System for Europe, **1**, 1, EC, DG Environment, Unit D2, 1-2.

Evans, S., 1999. A prototype environmental information system for London (London Environment Online or "LEO"). Working paper. CASA Working Papers (16). Centre for Advanced Spatial Analysis (UCL), London, UK.

Feng, S., 2008. Watershed Management and Water Quality Trading: A Systematic Framework for Assessing Uncertainty, *PhD Dissertation*, University of Georgia.

Feng, S., and Beck, M. B., 2008. An Integrated Study of Watershed Modeling under Uncertainty. *Proceedings of the 4th IWA Young Water Professionals International Conference*, The International Water Association. July, 2008. Berkeley, U.S.

Fielding, R. T., 2000. Architectural Styles and the Design of Network-based Software Architec-tures, *PhD Dissertation*.

Fisher, M. J., and Marshall, A.P., Understanding descriptive statistics, *Australian Critical Care*, **22**, 2, May 2009, Pages 93-97.

Foote, K.E., and Kirvan, A.P., 2000. WebGIS. Unit 133. NCGIA core curriculum in geographic information science. Available online http://ncgia.ucsb.edu/giscc/units/u133 /u133_f.html. Last accessed May 2009.

Freedman, P. L., Nemura, A. D., and Dilks, D. W., 2004. Viewing total maximum daily loads as a process, not a singular value: adaptive watershed management. *Journal of Environmental Engineering*, **130**, 6, 695-702.

Frehner M., and Brandli, M., 2006. Virtual database: Spatial analysis in a Web-based data management system for distributed ecological data. *Environmental Modelling & Software,* **21,** 1544-1554.

GAO, 2004. US Government Accountability Office Report, Watershed Management: Better Coordination of Data Collection Efforts Needed to Support Key Decisions, GAO-04-382, June 7, 2004.

Ghaemi, P., Swift, J.N., Goldberg, D.W., and Wilson, J.P., 2006. Prototype geographic search and query tools for a library digital archive, *ArcUse,* **9,** 2, 30–32.

Ginige, A., Murugesan, S., Khandelwal, V., Pollard, T., Costadopoulos, N., and Kazanis, P., 2001. Information Technology in Western Sydney-Status and Potential-Survey. Sydney, Australia, University of Western Sydney.

Goodall, J. L., Horsburgh, J. S., Whiteaker, T. L., Maidment, D.R., and Zaslavsky, I., 2008. A first approach to web services for the National Water Information System. *Environmental Modelling & Software,* **23,** 404-411.

Groot, R., and McLaughlin, J., 2000. Geospatial Data Infrastructure Concepts, Cases and Good Practice. *Oxford University Press*, New York.

Gruber, T. R., 1991. The role of common ontology in achieving sharable, reusable knowledge bases. In Fikes, R. and Sandewall, E., editors, *Proceedings of Knowledge Representation and Reasoning* (KR&R-91). Morgan Kaufmann Publishers: San Mateo, CA.

Gschwind, B., Menard, L., Albuisson, M., and Wald, L., 2006. Converting a successful research project into a sustainable service: The case of the SoDa Web service. *Environmental Modelling & Software,* **21,** 2006, 1555-1561.

Guariso, G., and Werthner, H., 1989. Environmental Decision Support Systems. *Ellis Horwood Limited*, Chichester.

Guttler R., Denzer, R., and Houy, P., 2001. User interfaces for environmental information systems interactive maps or catalog structures? Or both? *Advances in Environmental Research*, **5**, 345-350.

GWP, 2000. Integrated Water Resources Management. TAC Background Papers No 4, *Global Water Partnership*, Stockholm, 67.

Haklay, M. E., 2003. Public access to environmental information: past, present and future, Computers, *Environment and Urban Systems*, **27,** 163–180.

He, C., 2003. Integration of geographic information systems and simulation model for watershed management, Environmental Modelling & Software, **18,** 809–813.

Henry, T., Shen, J., Lahlou, M., Shoemaker, L., Parker, A., Ouyang, J., Yang, H., and Ludwig, J., 2002. Mining Data Analysis System (MDAS), *Proceedings of the Water Environment Federation*, Watershed, **15**, 674-688.

Himberg, J., 2004, From Insights to Innovations: Data Mining, Visualization, and User Interfaces, *PhD Dissertation*, Helsinki University of Technology.

Histogram, Wikipedia. Available online http://en.wikipedia.org/wiki/Histogram. Last accessed November, 2009.

Horsburgh, J. S., 2004. Bear River Watershed Information System.

Horsburgh, J. S., Tarboton, D. G., Piasecki, M., Maidment, D. R., Zaslavsky, I., Valentine, D., and Whitenack, T., 2009. An integrated system for publishing environmental observations data, *Environmental Modelling & Software*, **24**, 8, 879-888.

Huang, G. H., and Chang, N. B., 2003. Perspectives of Environmental Informatics and Systems Analysis, *Journal of Environmental Informatics,* **1,** 1, 1-6.

Huang, G. H., Liu, L., Chakma, A., Wang, X. H. and Yin, Y.Y., 1999. A hybrid GIS-supported watershed modeling system. *Hydrol. Sci. J.*, **44**, 597-610.

ISESS, International Symposium on Environmental Software Systems. Available online http://www.isess.org. Last accessed October, 2004.

ISO 8879, 1986, Information Processing -- Text and Office Systems -- Standard Generalized Markup Language (SGML).

Jeong, S., and Liang, X., 2005. Survey of CUAHSI HIS at UC Berkeley. *CUAHSI Hydrologic Information System Symposium*, University of Texas at Austin, TX, March 7–9.

Jeyaraj, A., Rottman, J.W., and Lacity, M.C., 2006. A review of the predictors, linkages, and biases in IT innovation adoption research, *Journal of Information Technology,* **21**, 1, 1–23.

Jiang, F., 2007. Towards a Sustainable Framework – Modeling, Optimization and Evaluation of Nutrient Removal and Recovery Strategies, *PhD Dissertation*. The University of Georgia. Athens, Georgia.

Kacsuk, P., Kranzlmuller, D., Nemeth, Z. and Volkert, J., 2002. Distributed and Parallel Systems-Cluster and Grid Computing, *Kluwer Academic Publishers*, Boston, MA, USA.

Kaski, S., 1997, Data exploration using self-organizing maps. Acta Polytechnica Scandinavica, *Mathematics, Computing and Management in Engineering,* **82**, 57.

Khan, M. A., Gupta, V. P., and Moharana, P. C., 2001. Watershed prioritization using remote sensing and geographical information system: a case study from Guhiya, India. *Journal of Arid Environments*, **49**, 3, 465-475.

Kingston R., Carver, S., Evans, A., and Turton, I., 2000. Web-based public participation geographical information systems: an aid to local environmental decision-making. Computers, *Environment and Urban Systems*, **24**, 109-125.

Klien, E., Lutz, M., and Kuhn, W., 2006. Ontology-based discovery of geographic information services — An application in disaster management, *Computers, Environment and Urban Systems*, **30**, 1, 102-123.

Klingseisen, B., Heiss, G., Jung, M., Schönerklee, M., Dünnebeil, G., Havlik, D., Humer, H., and Gigler, H., 2006. Watermark – A Multidisciplinary Research Initiative for The EU Water Framework Directive. *7th International Conference On Hydroinformatics HIC*, Nice, France.

Kohonen, T., 1995, Self-Organizing Maps, **30**, *Springer Series in Information Sciences*. Springer, Berlin, Heidelberg.

Kraines, S. B., and Wallace, D. R., 2003. Urban sustainability technology evaluation in a distributed object-based modeling environment, *Computers, Environment and Urban Systems*, **27**, 2, 143-161.

Kyng, M., and Mathiassen, L., 1997, Computers and design in context, *MIT Press*.

Laine, S., 2003. Using Visualization, Variable Selection and Feature Extraction to Learn from Industrial Data, *PhD Dissertation*, Helsinki University of Technology.

Lam, D.C.L., Mayfield, C.I., Swayne, D.A., and Hopkins, K., 1994. A prototype information system for watershed management and planning. *J. Biological Sys*. **2,** 4, 499–517.

Lancaster University, MSc Program website in Environmental Informatics. Available online http://geography.lancs.ac.uk/gres/ei/ei.htm. Last accessed December, 2009.

Langaas, S., Ahlenius, H., Hannerz, F., and Nilsson, S., 2004. Towards GIS and Internet based information systems for trans-boundary river basins, Chapter 11 in: *Environmental information in European trans-boundary river basin management*. (Eds. Timmerman, J.G. and Langaas, S.), IWA publishing, London, U.K.

Lauber, T.B., and Knuth, B.A., 2000. Citizen participation in natural resource management: a synthesis of HDRU research. *HDRU Series,* No 00–7. Department of Natural

Resources, Human Dimensions Research Unit, Cornell University, Ithaca, New York, USA.

LeVine, D., Wirt, M., and Whitebook, B., 2003. Practical Grid Computing for Massively Multiplayer Games, *Charles River Media*, MA, USA.

Lee, B., and Scholz, M., 2006. Application of the self-organizing map (SOM) to assess the heavy metal removal performance in experimental constructed wetlands, *Water Research*, **40**, 18, 3367-3374.

Lee, C., Jang, J., Jeong, K., Kim, D., and Joo, G., 2009. Patterning habitat preference of avifaunal assemblage on the Nakdong River estuary (South Korea) using self-organizing map, *Ecological Informatics*, **5**, 2, 89-96.

Lenz, R., and Kuhn, K. A., 2004. Towards a continuous evolution and adaptation of information systems in healthcare, *International Journal of Medical Informatics*, **73**, 1, 75-89.

Liu, R., 2000. Monitoring, Modeling, and Control of Nutrient Removal in the Activated Sludge Process, *PhD Dissertation*. The University of Georgia. Athens, Georgia.

Lin, Z., 2003. Modeling environmental systems under uncertainty: towards a synthesis of data-based and theory-based models. *PhD dissertation*, Warnell School of Forest Resources, University of Georgia, USA.

Lin, Z., and Beck, M.B., 2006. Understanding complex environmental systems: a dual approach, *Environmetrics*, **18,** 1, 11–26.

Liu, Y., Hill, D. J., Rodriguez, A., Marini, L., Kooper, R., Futrelle, J., Minsker, B., and Myers, J. D., 2008. Near-real-time precipitation virtual sensor using NEXRAD data, *Proceedings* *of the 16th ACM SIGSPATIAL international conference on Advances in geographic information systems*, November 05-07, Irvine, California.

Liu, Z., and Liang, Y., 2003. Design of metadata in a hydrological integrated scientific data management system. *The Seventh Joint Conference on Information Sciences* (JCIS), Cary, North Carolina, 418–421.

Liu, Z., Liang, Y., and Liang, X., 2003. Integrated data management, retrieval and visualization system for earth science datasets. *The 17th Conference on Hydrology*, Long Beach, CA.

Lorenz M. H., Arnfalk, P., Erdmann, L., Goodman, J., Lehmann, M., and Wager, P. A., 2006. The relevance of information and communication technologies for environmental sustainability, A prospective simulation study. *Environmental Modelling & Software,* **21,** 1618-1629.

Loucks, D.P., and da Costa, J.R., 1991. Decision Support Systems. *Water Resources Planning*. NATO ASI, Series G. Springer-Verlag, Heidelberg.

Lovejoy, S.B., 1997. Watershed management for water quality protection: Are GIS and simulation models the answer?. *J. Soil Water Conserv*., **52**, 103-110.

MacDougall, D., and Crummett, W. B., 1980. Guidelines for Data Acquisition and Data Quality Evaluation in Environmental Chemistry, *Anal. Chem.,* **52**, 2242–2249.

Maguire, D.J., and Longley, P.A., 2005. The emergence of geoportals and their role in spatil data infrastructures. *Computers, Environment and Urban Systems*, **29**, 1, 3-14.

Maguire, L.A., 2003a. Interplay of science and stakeholder values in Neuse River total maximum daily load process. *ASCE Journal of Water Resources Planning and Management*, **129**, 4, 261–270.

Maguire, L.A., 2003b. Interplay of science and stakeholder values in Neuse River National Research Council, 2001. Assessing the TMDL Approach to Water Quality Management, by the Committee to Assess the Scientific Basis of the Total Maximum Daily Load Approach to Water Pollution Reduction. *National Academy of Sciences*, Washington, D.C.

Mailhot, A., Rousseau, A.N., Massicotte, S., and Dupont, J., 1997. A watershed-based system for the integrated manage-ment of surface water quality: the GIBSI system. *Water Sci. Technol.*, **36**, 381-387.

Manolakos, E., Virani, H., and Novotny, V., 2007. Extracting knowledge on the links between the water body stressors and biotic integrity, *Water Research*, **41**, 18, 4041-4050.

Marmaridis, I., Ginige, J.A., Ginige, A., and Arunatilaka, S., 2004. Architecture for evolving and maintainable web information systems. In: *Proceedings of the international resource management association conference*, IRMA 2004.

Mateescu, G. D., 2008. Polynomial interpolation and applications to autoregressive models, *Romanian Journal of Economic Forecasting*, **5**, 1, 119-129.

Mayfield, C., Joliat, M., and Cowand, D., 2001. The roles of community networks in environmental monitoring and environmental informatics. *Advances in Environmental Research*, **5**, 385-393

Michener, W.K., 2000. Research design: translating ideas to data. In: Michener, W.K., Brunt, J.W. (Eds.), Ecological Data: Design, Management and Processing. *Blackwell Science*, Oxford, 1-24.

Millward A. A., and Mersey, J. E., 2001. Conservation strategies for effective land management of protected areas using an erosion prediction information system (EPIS), *Journal of Environmental Management*, **61**, 4, 329-343.

Mustajoki, J., Hamalainen, R.P., and Marttunen, M., 2004. Participatory multicriteria decision analysis with Web-HIPRE: a case of lake regulation policy. *Environmental Modelling & Software,* **19**, 6, 537-547.

NERC, Science topics, list of definitions – Environmental Informatics. Available online http://www.nerc.ac.uk/funding/policy/sciencetopics.shtml#envinf. Last accessed March 3, 2008.

Olsson, P., and Folke, C., 2001. Local ecological knowledge and institutional dynamics for ecosystem management: A study of crayfish management in the Lake Racken watershed, Sweden. *Ecosystems,* **4**, 85-104.

Onkal-Engin, G., Demir, I., and Engin, S. N., 2005a, e-NOSE Response Classification of Sewage Odors by Neural Networks and Fuzzy Clustering, *Advances in Natural Computation, Pt 2, Proceedings Lecture Notes In Computer Science*, 3611, 648-651.

Onkal-Engin, G., Demir, I., and Engin, S. N., 2005b, Determination of the Relationship between Sewage Odor and BOD by Neural Networks, *Environmental Modeling & Software*, **20**, 7, 843-850.

Pahl-Wostl, C., 2002. Participative and stakeholder-based policy design, evaluation and modeling processes. *Integrated Assessment*, **3**, 1, 3–14.

Parker, A.K., 2004. The role of iron-rich sediment in the biogeochemical cycling of phosphorus in Georgia Piedmont impoundments. *PhD Dissertation*. University of Georgia, Warnell School of Forest Resources, Athens GA.

Pecar-Ilic J., and Ruzic, I., 2006. Application of GIS and Web technologies for Danube waterway data management in Croatia. *Environmental Modelling & Software,* **21,** 1562-1571.

Peng, Z. R., and Tsou, M. H., 2003. Distributed Geographic Information Services for the Internet and Wireless Networks. *John Wiley & Sons*, New Jersey, 336.

Piasecki, M., HYDROSEEK. Available online http://his.cuahsi.org. Last accessed September, 2008.

Pillmann, W., Geiger, W., and Voigt, K., 2006. Survey of environmental informatics in Europe. *Environmental Modelling & Software,* **21,** 1519-1527.

Plewe, B., 1997. GIS online, information retrieval, mapping, and the internet, Santa Fe NM, *OnWord Press,* 336.

Poch, M., Comas, J., Rodrıguez-Roda, I., Sanchez-Marre, M., and Cortes, U., 2004. Designing and building real environmental decision support systems. *Environmental Modelling & Software*, **19**, 9, 857-873.

Pokorny. J., 2006. Database architectures: Current trends and their relationships to environmental data management. *Environmental Modelling & Software*, **21**, 1579-1586.

Power, M.J., Strome, M., and Daniel, T.J., 1995. Editors, *Proceedings Decision Support 2001 Vol. 2, American Society for Photogrammetry and Remote Sensing,* Bethesda, Maryland.

Preston, M., Clayton, P., and Wells, G., 2003. Dynamic run-time application development using CORBA objects and XML in the field of distributed GIS. *International Journal of Geographical Information Science*, **17**, 4, 321-341.

Purvis M., Cranefield, S., Ward, R., Nowostawski, M., Carter, D., and Bush, G., 2003. A multi-agent system for the integration of distributed environmental information. *Environmental Modelling & Software*, **18**, 565–572.

Rajabifard, A., Feeney, M. E. F., Williamson, I., and Masser, I., 2003. National SDI initiatives. In: Williamson, I., Rajabifard, A., Feeney, M.F. (Eds.), *Developing Spatial Data Infrastructures From Concept to Reality*. Taylor and Francis, London, 95-109.

Raskin, R., Pan, M., and Mattmann, C., 2010. Enabling Semantic Interoperability for Earth Science Data, Final Report to NASA Earth Science Technology Office (ESTO). Available online http://citeseerx.ist.psu.edu/viewdoc/download ?doi =10.1.1.2.3585&rep=rep1&type=pdf. Last accessed January 2010.

Richards, R. P., 2004. Improving Total Maximum Daily Loads with Lessons Learned from Long-Term Detailed Monitoring. *Journal of Environmental Engineering*, **130**, 6, 657-663.

Romanowicz, R., Young, P., Brown, P., and Diggle, P., 2006. A recursive estimation approach to the spatio-temporal analysis and modelling of air quality data, *Environmental Modelling & Software*, **21**, 6, 759-769.

Schimak, G., 2003. Environmental data management and monitoring system UWEDAT. *Environmental Modeling & Software*, **18**, 573–580.

Schumann, A. H., Funke, R., and Schultz, G. A., 2000. Application of a geographic information system for conceptual rainfall–runoff modeling, *J. of Hydrology*, **240**, 1-2, 45-61.

Sharma, M., Maheshwari, M., Sengupta, B., and Shukla, B.P., 2003. Design of a website for dissemination of air quality index in India. *Environmental Modeling & Software*, **18**, 405–411.

Shiffer, M.J., 2001. Spatial multimedia for planning support. In: Brail, R.K. and Klosterman, R.E., Editors, 2001. *Planning support systems: integrating geographic information systems, models, and visualization tools*, ESRI, Redlands, CA, 361–385.

Smiatek, G., 2005. SOAP-based web services in GIS / RDBMS environment. *Environmental Modelling and Software*, **20**, 775-782.

Spooner, J., and Line, D.E., 1993. Effective monitoring strategies for demonstrating water quality changes from nonpoint source controls on a watershed scale. *Water Science and Technology*, **28**, 3-5, 143-148.

Sprague, R.A.J., and Carlson, E.D., 1982. Building Effective Decision Support Systems. *Prentice-Hall*, Englewood Cliffs, NJ.

Stadler, M., Ahlers, D., Bekker, R. M., Finke, J., Kunzmann, D., and Sonnenschein, M., 2006. Web-based tools for data analysis and quality assurance on a life-history trait database of plants of Northwest Europe. *Environmental Modelling & Software*, **21**, 1536-1543.

Stadler, M., Bekker, R.M., Finke, J., Kunzmann, D., and Sonnenschein, M., 2005. Using data mining techniques for exploring the key features of plant dynamics upon a newly built plant trait database. In: Hrebicek, J., Racek, J. (Eds.), *Proceedings of the 19th International Conference Informatics for Environmental Protection - Enviroinfo Brno* 2005. Masaryk University, Brno, 90-94.

Stadler, M., Kunzmann, D., Schlegelmilch, J., and Sonnenschein, M., 2004. Data quality, abstraction and aggregation in the LEDA traitbase. In: *"Sharing" e-Proceedings of the 18th International Conference Informatics for Environmental Protection*. Part 1, CERN, Geneve, Switzerland. editions du Tricorne, 515-525.

Suchman, L. A., 1987. Plans and situated actions—The problem of human machine communication. Cambridge: *Cambridge University Press*.

Sunlu, F.S., Demir I., Onkal-Engin, G., Büyükışık, B., Sunlu, U., Koray, T., and Kükrer, S., 2009. Nano-plankton Population Dynamics and Dissolved Oxygen Change across The Bay of Izmir by Neural Networks, *Journal of Environmental Monitoring*, **11**, 1165-1168.

Sunlu, F.S., Demir, I, Onkal-Engin, G., Sunlu, U., Buyukisik, B., Koray, T., Sever, T., Kukrer, S., and Aydin, A., 2007. Prediction of Nano-plankton Change in Izmir Bay (Aegean Sea - TURKEY) with Time by Neural Networks, *38th CIESM Congress*, April, 9-13, Istanbul, TURKEY.

SWAG, Semantic Web Agreement Group. Available online http://swag.webns.net. Last accessed September 2008.

Swift, J. N., Stepp, J. C., Bobbitt, J., Futrelle, J., Tiwana, S., and Ali, M., 2004. Information technology issues in the development of the pilot COSMOS/PEER-LL (Pacific Earthquake Engineering Research Institute (Berkeley) Lifelines Program). Geotechnical Virtual Data Center. *Geo-Trans 2004*, ASCE, UCLA, July 27–31.

Tait, M.G., 2005. Implementing geoportals: applications of distributed GIS. *Computers, Environment and Urban Systems*, **29**, 1, 33-47.

Tate, K.W., Dahlgren, R.A., Singer, M.J., Allen-Diaz, B., and Atwill, E.R., 1999. On California rangeland watershed: timing, frequency of sampling affect accuracy of water quality monitoring. *California Agriculture,* **53,** 6, 44–48.

Taylor, C. J., Pedregal, D. J., Young, P. C., and Tych, W., 2007. Environmental time series analysis and forecasting with the Captain toolbox, *Environmental Modelling & Software*, **22**, 6, 797-814.

Thomsen, A.G., 1980. Watershed Information System, *Dissertation Abstracts International Part B: Science and Engineering* [DISS. ABST. INT. PT. B - SCI. & ENG.]. **41**, 4, 176.

Timmerman, S., 2002. Langaas (Eds.), Environmental information in European trans-boundary river basin management. *IWA Publishing*, London, U.K.

Tomasic, A., and Simon, E., 1997. Improving access to environmental data using context information. *ACM SIGMOD Record*, **26,** 1, 11-15.

Triantafyllou A. G., Evagelopoulos, V., and Zoras, S., 2006. Design of a web-based information system for ambient environmental data, *Journal of Environmental Management*, **80**, 230–236.

Tsou, M. H., 2004a. Present realities and the future of internet GIS. The Global Geospatial Magazine. Available online http://gisdevelopment.net/magazine/years/2004 /july/29.asp.

Tsou, M. H., 2004b. Integrating web-based GIS and image processing tools for environmental monitoring and natural resource management. *Journal of Geographical Systems*, **6**, 155-174.

Tsou, M. H., and Buttenfield, B. P., 2002. A Dynamic architecture for distributing geographic information services. *Transactions in GIS*, **6**, 4, 355-381.

Usländer, T., 2003. Motivation for an IT Framework Architecture for the Implementation of the European Water Framework Directive. In: Schimak, G., et al. (Eds.), *Proceedings of the ISESS 2003 Symposium*. Semmering, 27-30. Mai 2003, IFIP

Conference Series, International Federation for Information Processing, ISBN. 3-901882-16-2.

Usländer, T., 2005. Trend of environmental information systems in the context of the European Water Framework Directive. *Environmental Modelling & Software*, **20**, 12, Special Issue "Environmental Knowledge and Information Systems", 1532-1542.

van Bemmelen, J., Fusco, L., and Guidetti, V., 2005. Access to distributed earth science data supported by emerging technologies. In: *Proceedings of the 19th International Conference Informatics for Environmental Protection*, September 7-9. Masaryk University in Brno, Czech Republic, 181-185.

Vesanto, J., Himberg, J., Alhoniemi, E., and Parhankangas, J., 1999. Self-organizing map in matlab: the som toolbox. In *Proceedings of the Matlab DSP Conference*, 35-40, Espoo, Finland, November, 1999.

Via, D., 2002. Data warehouse for the Danube Waterway (D4D). Interreg IIIB CADSES Project (2000-2006). *Danube Transport Development Agency Ltd*., Vienna.

Villa, F., Athanasiadis, I. N., and Rizzoli, A. N., 2009. Modelling with knowledge: A review of emerging semantic approaches to environmental modeling, *Environmental Modelling & Software*, **24**, 5, 577-587.

W3C, World Wide Web Consortium. Available online http://www.w3.org. Last accessed November, 2009.

Wang, X., 2005. Integrating GIS, simulation models, and visualization in traffic impact analysis, *Computers, Environment and Urban Systems*, **29**, 4, 471-496.

Wang, X., Homer, M., Dyer, S. D., White-Hull, C., and Du, C., 2005. A river water quality model integrated with a web-based geographic information system. *Journal of Environmental Management*, **75,** 219–228.

Ward, K. T., and Johnson, G. R., 2007. Geospatial methods provide timely and comprehensive urban forest information, *Urban Forestry & Urban Greening*, **6**, 1, 15-22.

Ware, C., 2000. Information visualization: perception for design. *Morgan Kaufmann Publishers*, San Francisco, CA.

WebH2O, Inc., Web-based environmental information management systems. Available online http://info.webh2o.net Last accessed September 2008.

WFD, Final Summary of Proceedings, 2001. Public Participation, NGOs And The Water Framework Directive In Central And Eastern Europe. WWF/GWP - *Pan-CEE NGO Conference on Public Participation and the Water Framework Directive*. Budapest, March, 8-10.

WHO, 2000. The World Health Report 2000. Health Systems: Improving performance, Geneva: *World Health Organization*.

Wikipedia, The Free Encyclopedia, Confidence Interval. Available online http://en.wikipedia.org/wiki/Confidence_interval. Last accessed January 2010.

Williamson, I., Rajabifard, A., and Feeney, M.F., 2003. Developing Spatial Data Infrastructures - From Concept to Reality. *Taylor and Francis*, London.

WISE, Watershed Information System. Available online http://www.watershed concepts.com. Last accessed September 2008.

Yang, C., Wong, D.W., Yang, R., Kafatos, M., and Li, Q., 2005. Performance improving techniques in web-based GIS. *International Journal of Geographical Information Science*, **19**, 3, 319-342.

Young, P. C., Taylor, C. J., Tych, W., Pedregal, D. J., and McKenna, P.G., 2004. The Captain Toolbox, Centre for Research on Environmental Systems and Statistics, *Lancaster University*. Available online http://www.es.lancs.ac.uk/cres/captain.

Zeng, W., 2000. A Model for Understanding and Managing the Impacts of Sediment Behavior on River Water Quality, *PhD Dissertation*. The University of Georgia. Athens, Georgia.

Zhang, L., Scholz, M., Mustafa, A., and Harrington, R., 2008. Assessment of the nutrient removal performance in integrated constructed wetlands with the self-organizing map, *Water Research*, **42**, 13, 3519-3527.

Zhang, L., Scholz, M., Mustafa, A., and Harrington, R., 2009. Application of the self-organizing map as a prediction tool for an integrated constructed wetland agro-ecosystem treating agricultural runoff, *Bio-resource Technology*, **100**, 2, 559-565.

Zhang, L. J., Chung, J. Y. and Zhou, Q., 2003. Developing grid computing applications, part 1-introduction of a grid architecture and toolkit for building grid solutions. Available online http://www.106.ibm.com/developerworks/webservices/library/ws-grid1/resources.

Zheng, C., 2005. Summary of User Needs, *CUAHSI Hydrologic Information System Symposium*, University of Texas at Austin, TX, March 7–9.

Zhu, X., and Dale, A.P., 2001. JavaAHP: a web-based decision analysis tool for natural resource and environmental management. *Environmental Modeling & Software*, **16**, 251–262.

Zimmermann, R., Bardet, J. P., Ku, W. S., Hu, J., and Swift, J., 2003. Design of a geotechnical information architecture using Web Services. *In Proceedings of the seventh world multi-conference on systemics, cybernetics and informatics*, Orlando, Florida, July 27–30.

APPENDICES

APPENDIX A

ABBREVIATIONS

2D	Two dimensional
AJAX	Asynchronous JavaScript and XML
API	Application Programming Interface
AR	Auto-Regressive
CGI	Common Gateway Interface
CSS	Cascading Style Sheets
CSV	Comma Separated Value
D4D	Data warehouse for Danube waterway
DBMS	Database Management Systems
DHR	Dynamic Harmonic Regression
DSS	Decision Support System
EI	Environmental Informatics
EIS	Environmental Information System
EPA	Environmental Protection Agency
EPCL	Environmental Process Control Laboratory
ESA	European Space Agency

EU	European Union
GAO	Government Accountability Office
GIS	Geographical Information Systems
GWIS	Georgia Watershed Information System
HTTP	Hypertext Transfer Protocol
HVHQ	High Volume High Quality
IS	Information System
ISESS	International Symposium on Environmental Software Systems
ISO	International Organization for Standardization
IT	Information Technology
JCIS	Joint Conference on Information Sciences
KML	Keyhole Markup Language
NERC	Natural Environment Research Council
NOAA	National Oceanic and Atmospheric Administration
NODC	National Oceanographic Data Center
NSF	National Science Foundation
NWIS	National Water Information System
RDF	Resource Description Framework
RIS	River Information Services
SDI	Spatial Data Infrastructure
SGML	Standard Generalized Markup Language

SOM	Self Organizing Maps
SW	Semantic Web
SWAG	Semantic Web Agreement Group
TMDL	Total Maximum Daily Load
UI	User Interface
USGS	United States Geological Survey
W3C	World Wide Web Consortium
WFD	Water Framework Directive
WISE	Water Information System for Europe
WS	Web Services
WWQS	Watershed Water Quality Simulator
WWTP	Wastewater Treatment Plant
XML	Extensible Markup Language

APPENDIX B

GLOSSARY

Annealing: Information annealing or knowledge annealing, is a network-based information system in which all users of the system are permitted to change the system at will [wikipedia.com].

Bias: Bias is a term used to describe a tendency or preference towards a particular perspective, ideology or result, when the tendency interferes with the ability to be impartial, unprejudiced, or objective. In statistical hypothesis testing, a test is said to be unbiased when the probability of rejecting the null hypothesis exceeds the significance level when the alternative is true, and is less than or equal to the significance level when the null hypothesis is true [wikipedia.com].

Confidence interval: It gives an estimated range of values which is likely to include an unknown population parameter [cnx.org].

Correlation: It is a single number that describes the degree of relationship between two variables.

Data integrity: Data integrity is data that has a complete or whole structure. All characteristics of the data including business rules, rules for how pieces of data relate, dates, definitions and lineage must be correct for data to be complete [wikipedia.com].

Data inventory: A data inventory is a library in which collections are stored in digital formats and accessible by computers. The digital content may be stored locally, or accessed remotely via computer networks [wapedia.com].

Data redundancy: It sometime refers to in computer data storage that all or part of the data stored in the array can be recovered in the case of disk failure. Redundancy is attained when the same data values are stored more than once in a table, or when the same values are stored in more than one table [wikipedia.com].

Data: The term data means groups of information that represent the qualitative or quantitative attributes of a variable or set of variables. Data are typically the results of measurements [wikipedia.com].

Density estimation: It is the construction of an estimate, based on observed data, of an unobservable underlying probability density function [wikipedia.com].

Extensibility: It is a system design principle where the implementation takes into consideration future growth [wikipedia.com].

Hypothesis: It is a proposal intended to explain certain facts or observations. A suggested explanation for a group of facts or phenomena either accepted as a basis for further verification (working hypothesis) or accepted as likely to be true [wordia.com].

Information: Information is interpreted data, or knowledge derived from study, experience (by the senses), or instruction [answerbag.com].

Integration: Digital integration, in computer science, allows data from one device or software to be read or manipulated by another, resulting in ease of use [wikipedia.com].

Interoperability: Interoperability is a property referring to the ability of diverse systems and organizations to work together [wikipedia.com].

Knowledge: Knowledge is defined by the Oxford English Dictionary as (i) expertise, and skills acquired by a person through experience or education; the theoretical or practical understanding of a subject, and (ii) what is known in a particular field or in total; facts and information [wikipedia.com].

Model: A model is a pattern, plan, representation, or description designed to show the main object or workings of an object, system, or concept [wikipedia.com].

Null hypothesis: A type of hypothesis used in statistics that proposes that no statistical significance exists in a set of given observations. The null hypothesis attempts to show that no variation exists between variables, or that a single variable is no different than zero. It is presumed to be true until statistical evidence nullifies it for an alternative hypothesis [answers.com].

Ontology: Ontology is, in the context of Semantic Web, a schema that formally defines the hierarchies and relationships between different resources [altova.com].

Quality assurance: Data quality assurance is the process of profiling the data to discover inconsistencies, and other anomalies in the data and performing data cleansing activities (e.g. removing outliers, missing data interpolation) to improve the data quality [wikipedia.com].

Semantic web: It is a web of data that can be processed directly and indirectly by machines [altova.com].

Standard deviation: It is a measure of the dispersion of a set of data from its mean.

Standard error: It is the estimated standard deviation or error of a series of measurements [wikipedia.com].

Uncertainty: A state of having limited knowledge where it is impossible to exactly describe existing state or future outcome, more than one possible outcome [wikipedia.com].

Variance: Variance of a random variable or distribution is the expected value of the square of the deviation of that variable from its expected value [wikipedia.com].

Web Services: It is a standardized communication framework on web to exchange information between organizations and information systems using open standards.