International Conference on Hydroinformatics

2014

A Web Based Water Resources Analysis Portal For Occoquan Watershed

Saurav Kumar

Adil Godrej

Thomas Grizzard

Harold Post

Justin Bartlett

11th International Conference on Hydroinformatics
HIC 2014, New York City, USA

A WEB BASED WATER RESOURCES ANALYSIS PORTAL FOR THE OCCOQUAN WATERSHED

SAURAV KUMAR (1), ADIL GODREJ (1), THOMAS GRIZZARD (1), HAROLD POST (1), JUSTIN BARTLETT (1)

(1): Occoquan Watershed Monitoring Laboratory, CEE, Virginia Tech, 9408 Prince William St, Manassas, VA 20110, USA

In most of the developed world, water resources data, including both water quality and quantity, have been collected for decades. Such datasets are often used to calibrate water resources models that in turn are capable of producing large amounts of output useful for activities such as forecasting and risk analysis. Like many other domains (e.g., finance, marketing, etc.), significant education, extension, and policy making benefits may be obtained by building data analytics tools to derive knowledge from these datasets. This knowledge may translate into intelligent decisions by policy makers and may also be used to enhance environmental education among an area's stakeholders. Modern information technology infrastructure and development in GIS have eliminated several hindrances that have plagued such endeavors in the past. Using web-based methods, we have developed the OccViz system that may be used to visualize and derive knowledge from the dataset collected for the Occoquan Reservoir and its tributary watershed in northern Virginia, USA. The OccViz user portal is interactive and may be useful for local stakeholders to retrieve near-real-time data from several sampling stations in the region, and the portal's design also facilitates environmental education for lay users. In addition, the OccViz system enables data curation, integrates data from several sources into a single database, and provides tools to analyze data irregularities in massive real-time datasets, which are essential functionalities for data managers.

INTRODUCTION

Water quality and quantity data collection has been and continues to be an important aspect of water resources management. Such data is usually collected to establish a baseline, to understand the natural water system, and to support decision making. Technologies used to collect and store data have changed enormously in recent times, making it easier and cheaper to collect such data. For example, information about stream flow and water quality data, such as dissolved oxygen, nitrate concentrations, and conductivity, may be collected in near real-time at a very high frequency (e.g., a 15-minute increment) using online sensors and retrieved using cell phone networks. These automatically collected, near-real-time data are often augmented by less frequent manual sampling and analysis for a variety of compounds from nutrients to trace organics. Further adding to the complexity of water resources data collection and analysis, data may be collected at several water column depths, from different analysis matrixes (e.g., soil, water, and biological matrixes), and may be censored due to concentrations below the detection limit or other reasons. This diversity in the nature of water resources data, which often have to be analyzed together, creates unique issues for data curation, analysis, and visualization.

The environmental community has been working on these problems, and considerable advances have been made in the creation of standards for acquiring, storing, and sharing data, exemplified by the development of Observation Data Model (ODM) [1], OGC® WaterML, and CUAHSI Hydrologic Information System. Expert domain tools to discover and analyze data, such as HydroDesktop [2] and the EGRET package in R, have also been developed. However, several of these technologies have not been designed for lay stakeholders and non-expert decision makers and have a significant learning curve. In this paper we describe a web-based system, "OccViz," designed with modern, freely available, word-wide-web (web) technologies that, once set-up, may be used by lay stakeholders and experts for water resources data curation, analysis, and visualization.

Advances in web technologies made OccViz possible. The web has transitioned from the classic web with text, images, and simple navigation based on hyperlinks to an application environment with asynchronous network communications and desktop-style web apps. This transition is enabled by technologies, such as Asynchronous JavaScript and XML (AJAX), which allows modification of the Document Object Model for web pages and download new data from the Web server without undergoing a full page reload, and the HTML5 standard, which has reduced dependencies on plugins with offline support have enabled. Several new, free web-mapping application programming interfaces suitable for environmental data visualization, and spatial databases are also available. OccViz was designed to take advantage of these technologies and bring a modern web look and feel to water resources data visualization.

OCCVIZ DESIGN: KEY FEATURES

Figure 1 shows a schematic of the OccViz system. As illustrated in the figure, OccViz interfaces with telemetry systems for on-field equipment and other databases available for water resources data using a custom program designed for each input source. For example, to import data from stream stations using Sutron® dataloggers, a program is used that interfaces with Sutron® XConnect software to retrieve data from the field. Custom programs were adopted because equipment manufacturers use their unique, often proprietary, method and database schema to retrieve and store data. Nevertheless, a service was also created to import data formatted in standard WaterML format. All data in OccViz is first collected into a Central Water Resources Data Warehouse (CWRDW). Provisions are made in the CWRDW to allow expert users to directly connect and retrieve data using database connections mechanisms, such as JDBC and ODBC. OccViz, based on the configuration of the web portal, sub-samples the CWRDW to only collect data that will be available for analysis through the OccViz web portal in the Web Database Server (WDS). The WDS organizes the data in a database schema similar to the ODM that has been stream-lined for fast web retrieval using Java Persistence Architecture (JPA). The WDS also stores all web-portal configuration, user, and forum datasets.

All of the water resources data imported into the WDS are checked for outliers using methods based on individual parameter characteristics. Two different checks based on "hard-limit" and "soft-limit" criteria are performed on the data at the time of import. Hard-limits, which are used to screen out data that are likely invalid even under extreme conditions, are set for each parameter. Data outside the hard-limit bounds are not displayed to users using the web interface and are flagged for the administrator to validate. Soft-limits are automatically computed based on the historic dataset; the present OccViz system uses the 1st and 99th percentiles as the soft-limit bounds. Data outside the soft-limit bounds are displayed using the web interface, and the administrator is notified about such data. Note that these simple outlier detection methods are been updated to utilize information that may be available from spatial neighbors, surrogate, and auto-correlation. Several statistical parameters, such as percentiles (1–100) and multiple moving window (1-10 years) standard deviations are also computed for each constituent at every station and stored in the WDS. These statistical parameters are updated on a daily schedule.

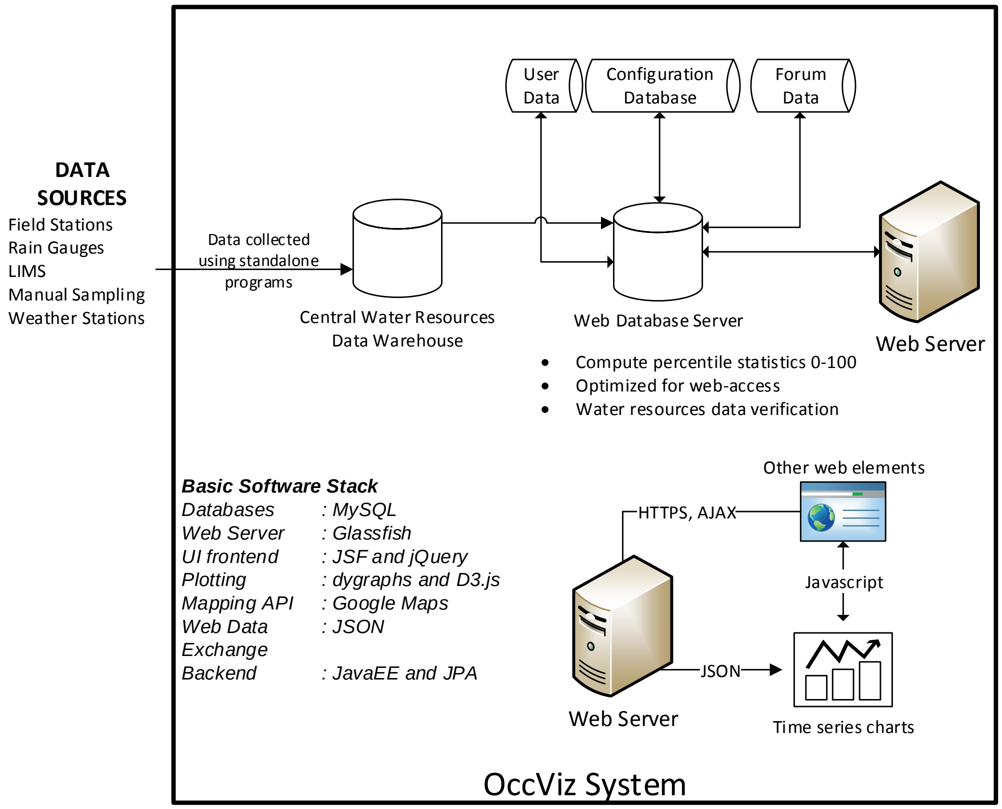


Figure 1. Schematic of OccViz System.

The web interface for the OccViz system is designed with two set of views: Admin Views and User Views (Figure 2). Admin views provide methods to manage data (update, delete, and insert), identify irregular data/outliers, manage stations (locations, descriptions, etc.), manage parameters at every station and the hard-limit for each parameter, manage users and their data access, and manage forum posts. User views are developed for user interaction and they allow users to retrieve information about stations including pictures of the station and land use of the sub-watershed, plot historic and near-real-time data, construct graphs comparing parameters from different stations, post in the forum, and change configurations to customize the look and feel of OccViz. Screenshots of some of these views from the OccViz implementation for the Occoquan Watershed (available at https://wqdata.owml.vt.edu), located in the northern Virginia suburbs of Washington, D.C., are shown in Appendices A and B.

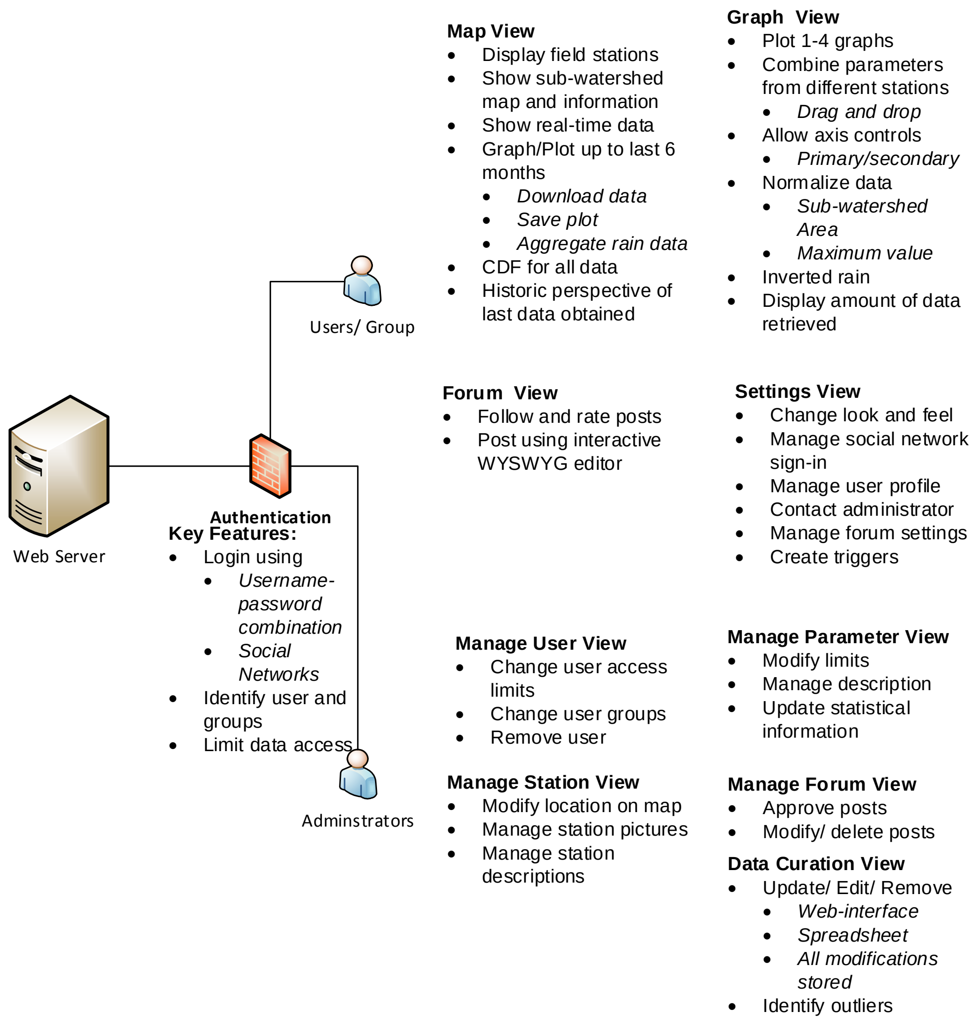


Figure 2. Key views available through OccViz web interface and their major purposes.

DISCUSSIONS AND FUTURE ENHANCEMENTS

One key purpose of the OccViz system is to update current environmental data visualization and analysis methods and bring it at par with the state-of-the-art in web designs. Modern browsers and web standards have enabled creation of desktop-style web applications that are more interactive and may be distributed widely. Moreover, as demonstrated by OccViz, these interactive visualizations enable users to analyze relations between parameters easily. There is considerable evidence in literature from various domains that visualization has a key role to play in communicating complex science to the public [3-5]. Thus, it may be expected that environmental education, in schools and for other purposes, is also likely to benefit from such endeavors.

OccViz also provides a platform which may be built-on and customized. Tremendous work has been done to visualize and statistically analyze water resources data using the programming language R. Packages in R, such as the Exploration and Graphics for RivEr Trends (EGRET) package developed by the USGS, provide unique and informative visualization methods, especially when censored data are involved. These visualization methods may be adopted for the web using JavaScripts libraries, such as D3.js, and easily integrated with OccViz. Further, the OccViz platform incorporates several data streams together that may also be utilized to analyze relationships among parameters and drive models for forecasting. For example, meteorological prediction in conjunction with rainfall-runoff models and other water quality prediction models may be used to display not just real-time but forecasted water quantity and quality information. Integrating several streams of data also aids in data curation. Integration of data makes is simpler to use surrogates and spatial neighbors to identify irregularities, especially in high frequency water resources datasets (e.g., precipitation, stream flow, etc.).

The first version of OccViz was developed for the Occoquan system and is available at https://wqdata.owml.vt.edu. Moving forward, it is important to have community partnership for continued development. We intend to release a freely available version of the OccViz system that may be used by other experts seeking to utilize a similar data collection and visualization mechanism. Also, we are trying to incorporate actively used water resources data standards such as WaterML in the OccViz system.

REFERENCES

[1] J. S. Horsburgh, D. G. Tarboton, D. R. Maidment, and I. Zaslavsky, "A relational model for environmental and water resources data," *Water Resources Research,* vol. 44, p. W05406, 2008.

[2] D. P. Ames, J. S. Horsburgh, Y. Cao, J. Kadlec, T. Whiteaker, and D. Valentine, "HydroDesktop: Web services-based software for hydrologic data discovery, download, visualization, and analysis," *Environmental Modelling & Software,* vol. 37, pp. 146-156, 11// 2012.

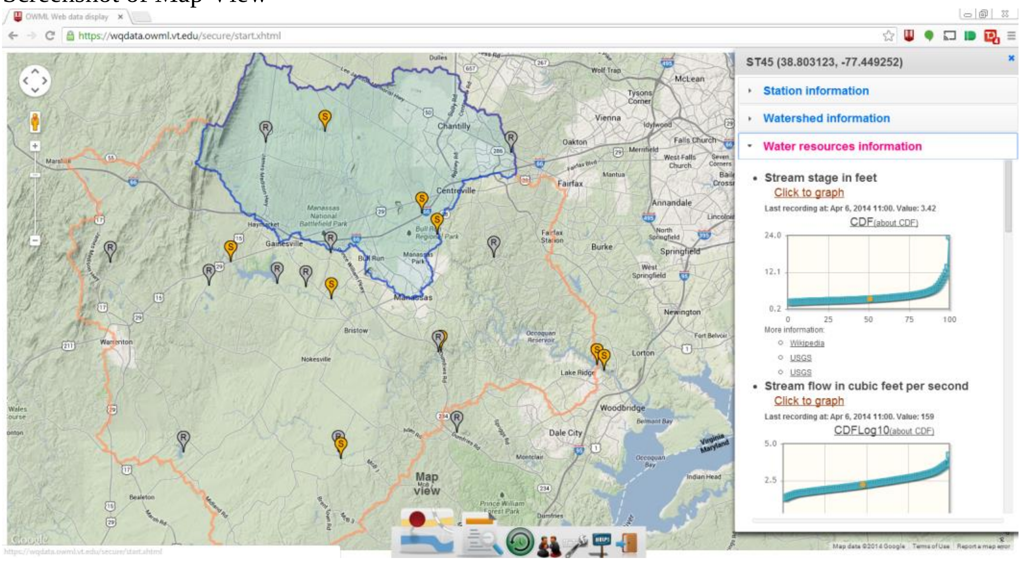
[3] M. Stieff, "Mental rotation and diagrammatic reasoning in science," *Learning and Instruction,* vol. 17, pp. 219-234, 4// 2007.

[4] J. Gilbert, "Visualization: A Metacognitive Skill in Science and Science Education," in *Visualization in Science Education*. vol. 1, J. Gilbert, Ed., ed: Springer Netherlands, 2005, pp. 9-27.

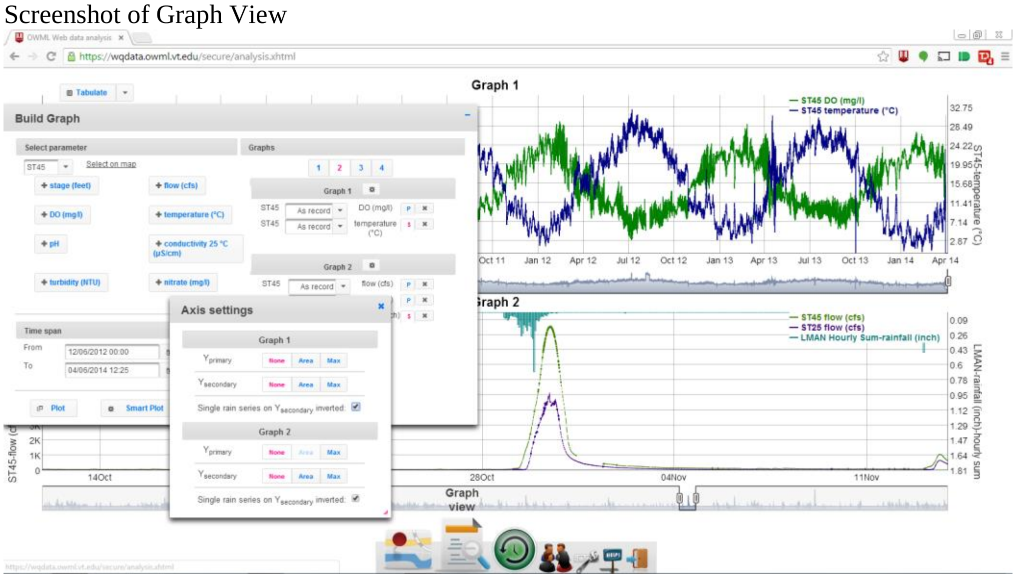
[5] M. Hegarty, "Diagrams in the Mind and in the World: Relations between Internal and External Visualizations," in *Diagrammatic Representation and Inference*. vol. 2980, A. Blackwell, K. Marriott, and A. Shimojima, Eds., ed: Springer Berlin Heidelberg, 2004, pp. 1-13.

APPENDIX A (USER VIEWS)

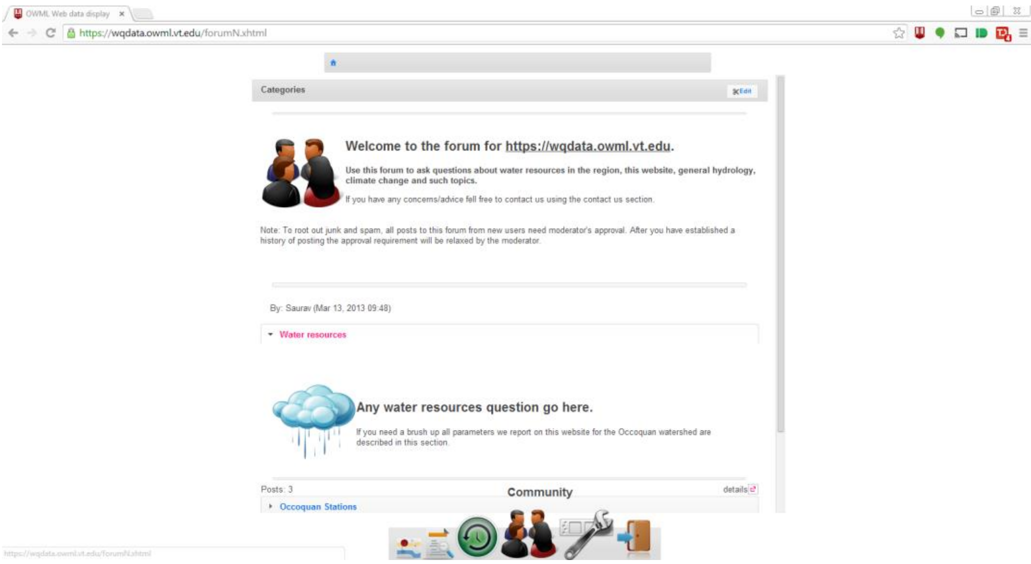
Screenshot of Map View



Screenshot of Graph View



Screenshot of the Forum View



APPENDIX B (ADMIN VIEWS)

Screenshot of Data View to edit and verify possible outliers.