Designing Geovisual Analytics Application for Exploring Hydrological Data

Mahshid Marbouti, Rahul Bhaskar, Zahra Shakeri, Craig Anslow, Leland Jackson[1], Frank Maurer
Department of Computer Science
Department of Biological Sciences[1]
University of Calgary, Canada.
E-mail: {mmarbout, rbhaskar, zshakeri, canslow, ljackson, fmaurer}@ucalgary.ca

***Abstract*—The ability to analyze and visualize large amounts of environmental and hydrological data on maps is difficult. Interaction and manipulation of data is crucial for decision making in natural disasters like floods. Hence, we present a Water Management Application (WMA), a geovisual analytics application to help analysts explore large amounts of hydrological data on web-based maps, create early flood warnings, and make strategic decisions in critical situations. With WMA we have provided monitoring and management of environmental resources which can help inform the design of future geospatial and geovisual analytics applications.**

I. Introduction

Collecting and analyzing relevant data are fundamental for describing hydrological phenomena and climate change [2] . To understand these environmental issues a significant amount of hydrological and weather data is collected hourly from different stations and instruments. For example, there are many weather stations that provide hourly data such as precipitation, snowfall, temperature, and wind. Such data are useful to model and predict events in critical situations.

In late June, 2013, there was a century frequency flood in Alberta that caused an estimated $5 billion worth of damage because of late warnings. To prevent these kinds of damages in similar situations, we built an application to help monitor water operations effectively in the Bow River basin to provide flood forecasting. The application target users are hydrological analysts.

In this paper we present Water Management Application (WMA), which is a geovisual analytics web-based application that utilizes different information visualization and data mining techniques to provide insights from hydrological and weather data. WMA provides interactive visualizations, historical charts for environmental parameters, investigation of parameters that can cause change rivers' water level and forecasting future discharge. WMA designed to support monitoring and managing hydrological and environmental resources.

There are existing flood information systems[1][4] but what makes our work different with them is that our application provide analysis and visualization techniques for a centralized hydrological and environmental dataset for Bow River Valley which can also be applied to other other regions.

II. WMA: Water Management Application

WMA provides interactive visualizations to help analysts gain and discover insight into hydrological data. As a geovisual analytics application, WMA integrates geospatial, hydrological, and weather data. It facilitates exploration of water resources by utilizing different types of visualizations.

A. Design

The following requirements were considered in the design of the WMA:

Environmental Data Exploration: In WMA, environmental data includes hydrological and weather data. Environmental data are associated with stations. Analysts need to retrieve and compare these data by selecting different stations on the map.

Map Interaction: Due to significant amounts of varied geospatial data, WMA should be able to represent data on web-based maps. This helps analysts gain a geospatial awareness of environmental data. Moreover, Analysts need to interact with geospatial data on the map and WMA allows users to interact with datasets by separating the geospatial data into different layers (e.g. stream layers, flood plains, and watersheds). WMA should also be a GIS application to represent geospatial data on web based maps.

Web Enabled: Web platform make the application accessible from different locations and devices.

To design WMA, the following issues were considered:

- Data Acquisition: How can we collect, prepare and store all necessary data for monitoring hydrological and environmental events?
- Representation: What visualization techniques are most effective to help analysts understand and discover hidden trends in data?
- Interaction: What kind of map interaction is effective to explore geospatial data?
- Prediction: What analytical and data mining methods are effective to predict future floods?

B. Data

We acquired environmental and geospatial data for the Bow River basin from different publicly available sources. Environmental data included hydrological data (gauging stations which has parameters like water level, discharge) and weather

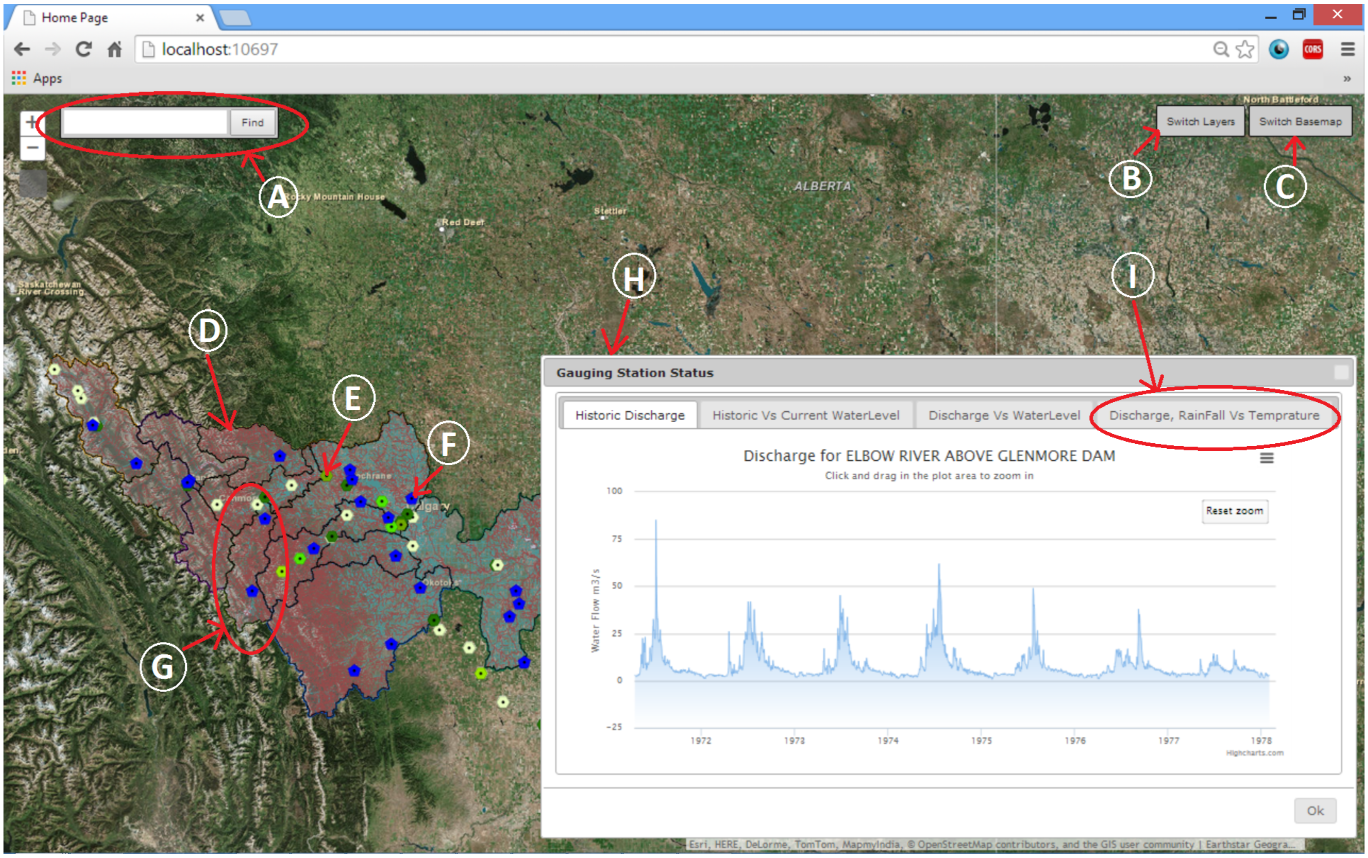


Fig. 1. User Interface of WMA - Water Management Application. (A. Search Box, B. Switch Layer Widget, C. Switch Basemap Widget D. Bow River Basin, E. Weather Station (blue icons), F. Gauging Station (green icons), G. Watershed, H. Data Visualization screen for station, I.Tabs for multiple visualizations.)

stations data (temperature, rainfall, waterfall, and snowfall). WMA uses these data, processes them and generates different visualizations.

WMA displays geospatial data of the Bow River basin geographical structure including location of different stations on the river, watersheds, streams and floodplain data. Therefore, multiple geospatial datasets were collected from different sources to provide geospatial visualizations.

C. User Interface

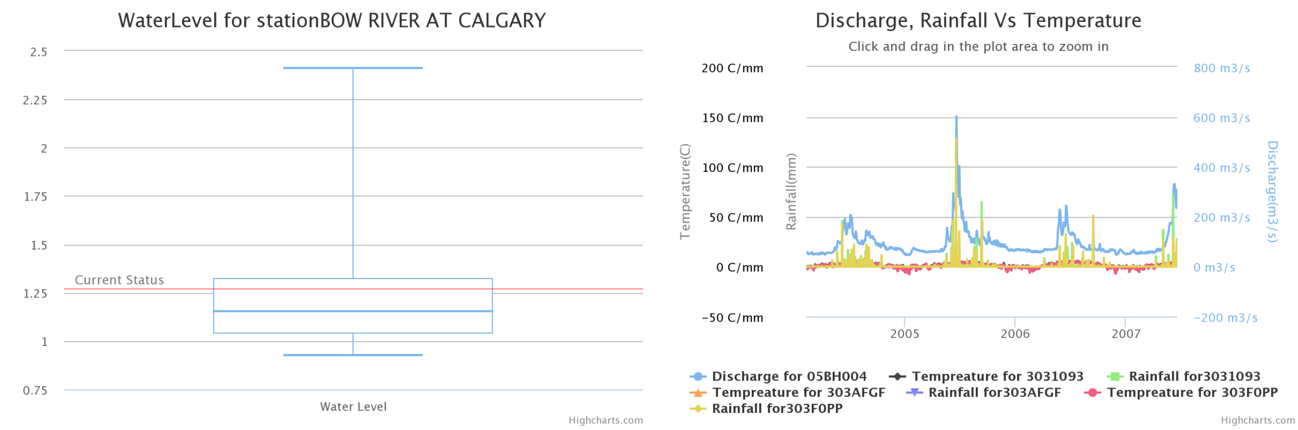
Figure 1 shows the user interface of WMA which contains features to visualize and analyze environmental data.

Geospatial Features: Initially when an analyst opens WMA (Figure 1) a map is centered on the Bow River basin (label D). The map display contains different geospatial information like the DEM (3D representation of land surface), rivers and lakes, watersheds (an area where all water drains into one place), weather stations, and gauging stations. There are multiple widget panels on the top right and top left of WMA. The panels at the top of the screen provide features to interact with the geospatial data and update according to an analyst´s requirements. The switch basemap (label C) widget provides a feature to change the base map. Default base map is the street view and user can switch between nine available basemaps. The default behavior of WMA is that it displays all layers at the launch time. To change the visibility of layers the analyst can use the switch layers widget (label B) from the top and toggle a layer's visibility on or off. Weather stations are represented by blue icons. Water gauging stations are represented by varying shades of green which help analysts see information about data ranges of a station (0-20, 20-40, 40-60 and 60+ years).

Analytical Features: WMA uses time series and multidimensional graphs for analytical features (label H). This feature can be viewed by clicking on weather station (label E) or gauging station (label F). When an analyst selects any station a window (label H) is displayed on the screen and shows different visualizations in different tabs. For weather stations there are two tabs showing different analysis results. Similarly gauging stations use four tabs to show analysis results.

D. Visualizations

Exploring information from a complex dataset is difficult. Structuring and visualizing the data could help to summarize and extract patterns from datasets [6]. The visualizations provide analysts the ability to recognize patterns, trends, and outliers in data. WMA supports different types of visualizations including maps, line graphs, dual x axis line graphs, box-plots, and infographics.



(a) Current water level in comparison with historic quartile values.

(b) Discharge, Rainfall Vs Temperature graph.

Fig. 2. WMA - Gauging Station Visualizations.

Geospatial Visualizations: Geospatial visualizations help display data like DEM (Digital elevation model), lakes, rivers, weather and gauging station locations which are hard to visualize using temporal or multi-dimensional visualizations. WMA displays world base maps with different layers. Layers display the Bow River basin geospatial data. Layers show small sub-areas (i.e. watershed), which help analysts see geographic regions that might have influence on hydrological phenomena. These watershed areas contain multiple markers that represent stations in each watershed area. This geospatial reference provides analysts the ability to navigate from one station to another and check the status of each station quickly.

Visualizations for Gauge Station: WMA contains multiple visualizations for gauging stations. WMA uses line graphs for displaying historic trends of water flow for a particular station (Label H in Figure 1). Box plot is used to show the current status of water level in comparison to the historical high and low water level. The Box plot can be used for generating alarms when water level is higher than a threshold (Figure 2(a)). WMA has multiple y-axis line graphs called 'discharge, rainfall vs temperature' to show comparative analysis between discharge and different environmental factors like rainfall and temperature (Figure 2(b)).

Visualizations for Weather Station: Weather stations contains current as well as the future weather forecast for each station and they use temporal charts for rainfall and snowfall historical data, which helps analysts understand climate trends at each station location.

E. Hydrological data exploration scenario

Suppose an analyst is seeking insight about a gauging station within the Bow River basin. The analyst wants to find past trends and determine which months extreme events like sudden increase and decrease in water level occurred.

The analyst begins by navigating to WMA from a web browser. Once WMA is loaded the analyst will locate a station. Searching for stations in WMA can be accomplished by visual or manual search. A visual search is done by locating a station on the map by looking at the map. A manual search is performed by using the search box feature in which the analyst can type the station name or number. If the station exists, WMA will highlight the station on the map by showing a red square over the station on the map.

Once the analyst has found the location of a station on the map and clicked on gauging station, a pop up is displayed that shows the station number and name. To retrieve the analysis result, the analyst will then click on the detail button to see different tabs. The first tab on the visualization labelled historic discharge (label H in Figure 1) that can be used to see the historic trend of water levels that mostly occur between late June and July.

An analyst may want to check that the current water level is within some defined limits, to determine if there is an alert situation. For this task the analyst switches to the second tab 'current vs historic' (Figure 2(a)) where they can see a comparative analysis between the current water level and the historic maximum and minimum values.

To check which parameter has significant impact on the flow of rivers. The analyst can navigate to the fourth tab (Figure 2(b)) of the gauging station to observe combinations of different parameters from multiple weather stations like snowfall, rainfall, temperature imposed over each gauge station discharge value in the same visualization. This helps to find the parameter that has significant correlation with the discharge. According to the graph (Figure 2(b)) it is apparent that temperature has a major impact on discharge values for the selected station as the change in the temperature (i.e. yellow line) also changes discharge, but it is not the same with the rainfall. Moreover, this graph contains data of all weather stations within the watershed for the selected gauging station.

In order to find the weather stations which has impact on the particular gauge station,the analyst can check the watershed (label G in (Figure 1)) to find weather station on the same watershed where the gauging is located.

F. Architecture and Implementation

The architecture of WMA has three layers. First a data layer to store different kinds of geospatial and environmental data in a relational DB. Second a server layer to process data queries and deliver geospatial maps. Third a user interface layer to display the geospatial and environmental visualizations. As data were available in different formats, collecting the data and parsing was accomplished manually and loaded into the database. The server layer focuses on the preparation and filtering the data. Therefore database provides input for an ArcGIS server for geospatial processing and analytics API for analytical processing. The user interface layer helps visualize the processed data on the web interface using web APIs (i.e. Highchart JS, ArcGIS JavaScript API).

III. Discussion

Performance: One of the technical challenges that we faced was performance. The challenge was regarding the delay in rendering visualizations of long term temporal data. In the UI, each station contains different charts about environmental temporal data and each chart was related to different parameters that were stored in different database tables. Hence, to display a temporal chart, the following steps were necessary: querying database and load data, filtering fetched data and transferring it from the server layer to the user interface layer and processing and then displaying the transferred data on the screen. Running these steps for huge amounts of data (i.e. 100 years of daily temporal data) was slow. The future plan to overcome this issue is display the data at multiple levels and let users interactively select and compare different time intervals .

Data Acquisition: We had a large variety of data distributed across many resources and not all the data we needed was publicly available so we needed to contact and explore different resources to gather the data. We are confident that we collected a valuable dataset of hydrological, weather, and geospatial data for the Bow River basin.

Evaluation: We focused on our domain expert for evaluation. We conducted preliminary user studies with the domain expert to evaluate our application. We had bi-weekly meetings with the domain expert and at each meeting we requested him to perform user scenarios and confirm if features aligned with his requirements. This iterative feedback and hints helped us to develop an application based on real requirements. After developing the applications we conducted a demo in which we gathered feedback from a group of students and business people. The feedback can be categorized as follows:

UI Enhancements: How does the design of the web interface can be more effective in understanding different features?

Representation Power: How can representation of hydrological elements on a map can be more expressive? For example after visualizing different gauging stations on the map, the domain expert asked us to colour code gauges (from dark green to light green) to show the range of data that is available for each gauge.

Hydrological Knowledge: The bi-weekly meetings with our customer helped gain hydrological knowledge we needed for data gathering and analysis.

Predictive Analytics: One of the design considerations for developing WMA was to provide easy access to geospatial and hydrological data for analysts. Our further consideration was to provide early flood warnings. To predict what might happen in critical situations is important. To move forward in this direction, we focused on predictive analytics. We needed to predict water level and discharge to help estimate what areas are going to be covered during a flood. As a first step, we chose Neural Networks (NN) – a verified Artificial Intelligence(AI) algorithm-which is one of the most suitable methods to predict discharge according to literature[5][3]. We used Matlab for the implementation of the NN method. The next step is to integrate the Matlab code with WMA to provide prediction regarding endangered areas in flood time.

Summary Exploring large amounts of flood related data on a map is challenging. In July 2013 there was a major flood in Alberta that caused huge damage. The late warnings during the flood demonstrated the need for an effective application to help hydrologists predict such disasters. To help hydrological analysts, we developed WMA, a domain specific geovisual analytics web-based application that helps analysts to explore and monitor environmental data with respect to climate change. Next, we plan to extend WMA by integrating our predictive analytics code and conduct usability studies with hydrological analyst to evaluate the effectiveness of the application.

References

[1] I. Demir and W. F. Krajewski. Towards an integrated flood information system: Centralized data access, analysis, and visualization. *Environmental Modelling & Software*, 50:77–84, 2013.

[2] J. Guzman, D. Moriasi, M. Chu, P. Starks, J. Steiner, and P. Gowda. A tool for mapping and spatio-temporal analysis of hydrological data. *Environmental Modelling & Software*, 48:163–170, 2013.

[3] C. Imrie, S. Durucan, and A. Korre. River flow prediction using artificial neural networks: generalisation beyond the calibration range. *Journal of Hydrology*, 233(1):138–153, 2000.

[4] A. Merchant, M. Kumar, P. Ravindra, P. Vyas, and U. Manohar. Analytics driven water management system for bangalore city. *Procedia Engineering*, 70:1137–1146, 2014.

[5] E. Toth, A. Brath, and A. Montanari. Comparison of short-term rainfall prediction models for real-time flood forecasting. *Journal of Hydrology*, 239(1):132–147, 2000.

[6] C. Ware. *Information Visualization: Perception for Design*. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 2004.