A Web-based Data Visualization and Analysis System for Watershed Management

Yufeng Kou[1], Changtien Lu[1], Adil Godrej[2], Thomas Grizzard[2], Harold Post[2]

[1]Dept. of Computer Science, Virginia Polytechnic Institute and State University, 7054 Haycock Road, Falls Church, VA 22043
E-mail:{ykou, ctlu}@vt.edu
[2]Occoquan Watershed Management Laboratory, 9408 Prince William Street, Manassas, VA 20110
E-mail:{agodrej, grizzard, hpost}@vt.edu

Abstract

In recent years, the combination of databases, data analysis, and Internet technologies has greatly enriched the functionalities of GIS and facilitated the dissemination of geospatial information. In this paper, we propose a web-based data visualization and analysis system to manage the watershed of the Occoquan Basin in Northern Virginia. This system distinguishes itself from others in the following respects. First, it enables near real-time data collection and dissemination. Second, data-query tools are embedded in the system to support effective data analysis and visualization functionalities. In addition to regular queries, aggregation queries have been developed to provide summary views of historical data with various granularities. Third, the visualization and analysis tools can be accessed via the Internet. WWW technologies, such as active web pages and Java programming, provide user-friendly interfaces for browsing, plotting, comparing, and downloading information of interest, without the need of dedicated GIS software. This system can render information to support many other applications, including pollution control and biochemical detection.

I. INTRODUCTION

Traditional GIS appear as stand alone tools deployed on dedicated computers with sophisticated graphical display equipment. In recent years, however, GIS has been changing its role to become an integrated system which combines geospatial applications with databases, data mining, and Internet technology. Databases, especially spatial databases, play an important role in storing large volume of spatial and aspatial data. Dedicated GIS databases can provide powerful data processing and map management functionalities, such as efficient data accessing, flexible spatial quering, and spatial data visualization. However, they often come with expensive price tags and require high performance computers. Their maintenance also requires highly professional engineers. In addition, traditional GIS are not designed for nonprofessional users. This limits their wide adoption by the general population. Since 1990s, Internet technology has triggered a revolutionary change in data access and dissemination. It has simplified access to geographic data and now it serves as a vehicle for disseminating geographic information to a multitude of users across various hardware and operating systems. The development of web technologies, such as Common Gateway Interface (CGI), scripting languages, Java Applet, Java Server Page (JSP), and Java Servlet, has provided an efficient and effective means of delivering spatial data products using the Internet.

Data visualization is an important aspect of data analysis. During the process, the collected data are summarized and presented in visual forms to help grasp the minute details and relationships of data for more effective decision-making. Advances in visualization technology have revolutionized fields such as architecture, medicine, among others. Data visualization is essential for complex data analysis, when the structure, features, patterns, trends, anomalies and relationships of the data are not easily detectable. Visualization facilitates the extraction of hidden patterns by presenting the data in various visual forms. Visualization not only provides a qualitative overview of large and complex data sets, it can also assist in identifying regions of interest and parameters for further quantitative analysis. Along with the wide spread use of the World Wide Web, visualization on the Internet has become the trend. Users do not need purchase any specific hardware or install any dedicated graphic software. As long as they have access to internet, they can view a particular data set in either graphical or tabular formats. In this paper, we present a web-based GIS which can collect data from distributed sensors and provide a highly interactive and intuitive graphical user interface for data exploration and analysis.

The Occoquan Watershed Monitoring Laboratory (OWML), established by the Virginia Polytechnic Institute Department of Civil Engineering, began its onsite operations since 1972 to the present time on comprehensive studies of water quality and the effects of the AWT (Advanced Wastewater Treatment) discharges. In the course of its operation, OWML has developed a comprehensive database of water quality in the Occoquan Basin and has been instrumental in formulating decisions in a number of areas critical to the ongoing management of water quality. For example, OWML provided water quality data in support the decision on AWT plant expansions and rendered information on water quality effects and cost-effective control of non-point sources of pollution. In this paper, we present a web-based visualization system that can be used to collect water data in near real-time and disseminate the data to the public in an efficient way.

This hydrologic and water quality data acquisition and analysis system was designed by the Occoquan Watershed Monitoring

1082-4006/05/1101-40$5.00

Laboratory. The Occoquan Watershed is located in northern Virginia and is situated on the southwestern periphery of the Virginia suburbs of Washington, D.C. The basin encompasses six political subdivisions of the Commonwealth of Virginia, including portions of four counties, and the entire land area of two independent cities. It is bounded by the Potomac Estuary to the east and Bull Run Mountain to the west. The northern and southern boundaries lie in the Counties of Fairfax and Fauquier-Prince William, respectively.

In this paper, we propose a web-based online GIS data visualization system to monitor and analyze the surface water resource of the Occoquan Watershed in Northern Virginia. This system can render information to support many applications, including pollution control and biochemical detection. This system distinguishes itself from other systems in the following aspects. First, it enables near real-time data collection and dissemination. The readings of sensors can be collected and published on the Internet within minutes. Second, data query tools are embedded in the system to support effective data analysis functionalities. In addition to regular queries, aggregation queries have been developed to provide summary views of historical data with various granularities. Third, the new information and analysis tools can be accessed via the Internet. WWW technologies, such as active web page and Java programming, provide user-friendly interfaces with which users can browse, plot, compare, and download data of interest without the need for dedicated GIS software. Fourth, the system includes easy-to-use visualization components. The water data can be visualized with tables, JPEG figures and PDF files. Both linear scale and log scale are supported in the display of data in graphic forms. By exploring the data cube concept [1], an aggregation view of the water data can be provided with different temporal granularities.

The remainder of this paper is organized as follows. In Section II, related work was surveyed. Section III describes the hardware architecture and software components of our system. The system functionalities and visualization components are demonstrated in Section IV. Section V discusses key technical issues in distributed GIS visualization and identifies several directions for future research. Finally, we summarize our work in Section VI.

II. RELATED WORK

Web-based GIS has been rapidly evolving with the development of WWW and Internet technologies. Generally, operations of Web-based GIS can be divided into two areas: server-side and client-side [2]. Systems emphasizing server-side operations perform most processes on the server, with only a "thin" client (Web Browser) installed on the user's machine. This method maximally simplifies user's responsibility for software installation and maintenance. However, if the number of users increases, the server load will increase and the system performance may degrade. Systems emphasizing client-side operations take advantage of client computing technologies, such as Active-X controls, plug-ins, and Java Applets. The system can cache necessary data on the user machines and support a more interactive Graphic User Interface (GUI). The price is that more software components need to be installed on the client machine. The common dilemma of Web-GIS is the balance between large data sets and limited network bandwidth. Much research has been conducted to promote the efficiency of WebGIS [2,3]. In Tu et al. (2001) [4], a GIS web service is implemented to reduce the user-perceived response time and support user self-guided navigation. The core idea is to decompose GIS data sets into small blocks on the server side, and use locality-based caching and pre-fetching on the client side. Wei et al. (1999) proposed a new method of efficient spatial data transmission for Web-based GIS [5]. Their system first divides a large-size map into several parts, called "Tiles", according to appropriate granularity. A tile will be sent by server only if it overlaps with the requested region. Since data is only transmitted on demand, the bandwidth utilization is optimized. 3D techniques are increasingly used to represent GIS data in a more intuitive way, especially for urban planning. In [6], a data model was proposed to support 3D topology and 3D geometry. Coors and Flick developed a web-based 3DGIS based on Java and VRML [7]. To promote transmission and display efficiency, their system also supports multiple levels of detail for visualization.

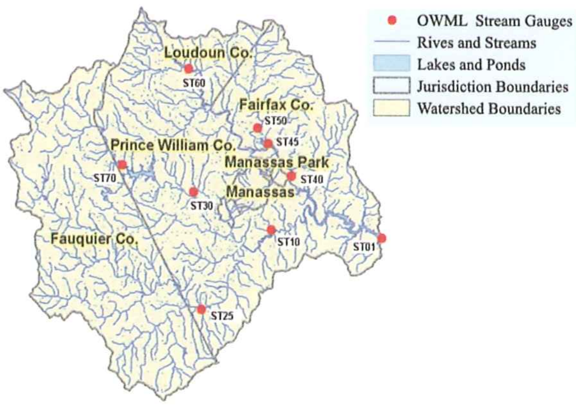
There are various Web-based GIS that have been used for regional natural resource analysis [8], mobile device positioning [9], or moving objects tracking [10]. However, these GIS are heterogeneous with different data representations and query languages. Thus, it has become an interesting problem to develop methods that will provide an integrated view of the data supplied by all sources and to devise a geographical query language for accessing and manipulating this integrated data. Tu et al, (2001) proposed a J2EE-based approach which can support the exchange of both spatial vector data and raster image data [11]. Boucelma et al. introduced a multi-tier client/server architecture which uses standard wrappers to access data and uses "derived wrappers" to capture additional query capabilities [12].

Han et al. proposed a spatial data mining protocol, DBMine, for knowledge discovery in relational databases and data warehouses [13]. They extended DBMine to GIS applications and developed a spatial data mining tool, GeoMiner [14, 15]. GeoMiner is capable of identifying various rules in spatial data, such as characteristic rules, comparison rules, and association rules.

There are several online water data query systems, for example, "Water Resources of the United States" by the U.S. Geological Survey (USGS). However, they are short of visualization and query performance. First, they cannot support aggregation views of the data. Second, they cannot provide near real-time data dissemination. Even the "real-time water data" from USGS can only display data that are two hours old. In this paper, we propose a web-based data visualization system which can support both near real-time data collection and flexible data visualization using the data cube concept, log scale plotting and JPEG/PDF/ text format displays.

III. CONCEPT AND SYSTEM ARCHITECTURE

OWML manages a dozen of monitoring stations distributed over the Occoquan Watershed. Figure 1(a) shows the locations of these stations in northern Virginia. The stations are set along rivers or runs, collecting various readings with automatic gauging devices. Figure 1(b) shows the picture of station ST01, which is on the upstream and southern side of the Occoquan Dam. Our system performs systematic collection, recording, evaluation, and dissemination of data on water quality, flow characteristics, and chemical dissolution. As shown in Figure 2, the collected data can be used for environmental quality evaluation, fish and wild life services, natural resources conservation, and weather and climate research. This system enables the local governments in northern Virginia to successfully deal with the competing uses of urban land, waste water discharges and urban runoff, and public water supply in a critical watershed-impoundment system.



(a) Location of the stations

(b) Picture of station ST01

Figure 1. The monitoring stations

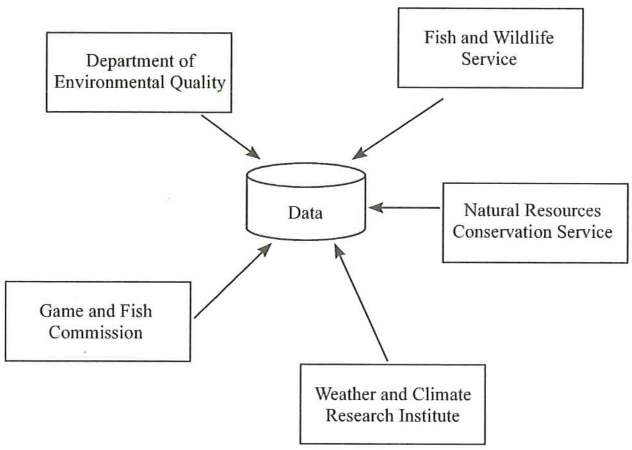


Figure 2. The applications of water monitoring data

This system provides two main functions: data collection and data visualization. Data collection controls the communication between monitoring stations and the database. It collects the readings from distributed sensors, converts them to an appropriate format, and inserts records into the database. Visualization is a web-based GUI, which accepts a user's requests and display the queried results in textual or graphical format. Data visualization can also help users perform basic data analysis, such as presenting aggregation views and statistical summaries.

The hardware architecture of our system is illustrated in Figure 3. The monitoring stations are connected to the OWML data center using phone lines, and the monitoring data are transmitted every 15 minutes. There are two databases in our system. The master database is located at OWML where it receives data from monitoring stations. The engineers can modify the master database if the data contain errors or noise. The slave database resides on the web server at the Northern Virginia Center of Virginia Tech. The slave database contains only a subset of the master database, is synchronized with the master database and only used for data visualization on the Internet. Duplicated databases are maintained for reasons of security. The data in the master database can only be operated by specific engineers at OWML, as some of the data are confidential and need to be filtered before being made available to the public. The master database and slave database are connected by a Virtual Local Area Network (VLAN), which provides secure communication. Data requests are initiated by web users, checked by the firewall, and then replied by the web server.

Figure 4 shows the software components and their relationships in the system. Our system has a standard three-tier architecture, consists of user interfaces, a web server and databases. The three-tier architecture has two major advantages. The first is a "thin client," which means that only a web browser is needed to be installed on the user's machine and all the computation and maintenance processes are

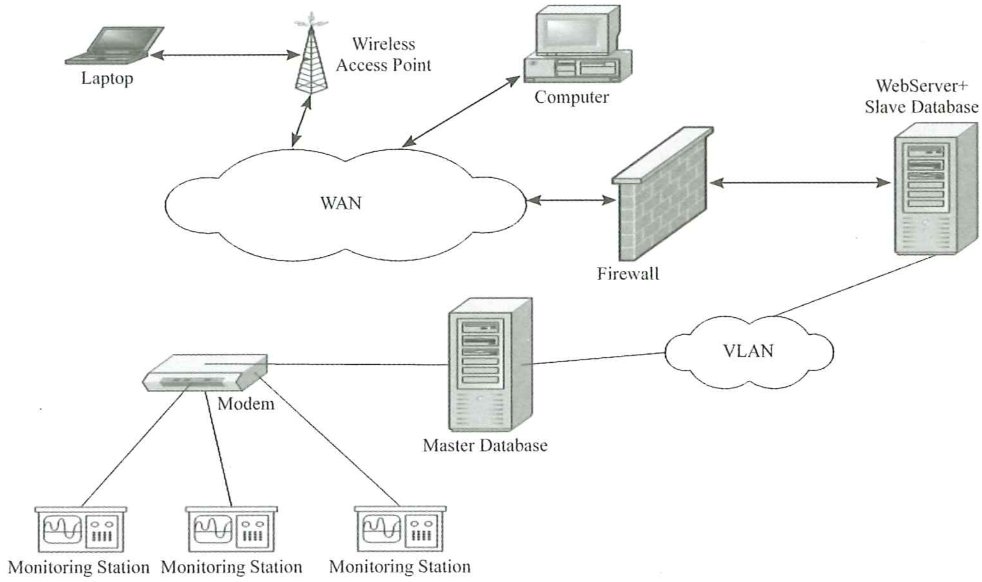


Figure 3. The hardware architecture of OWMLGIS

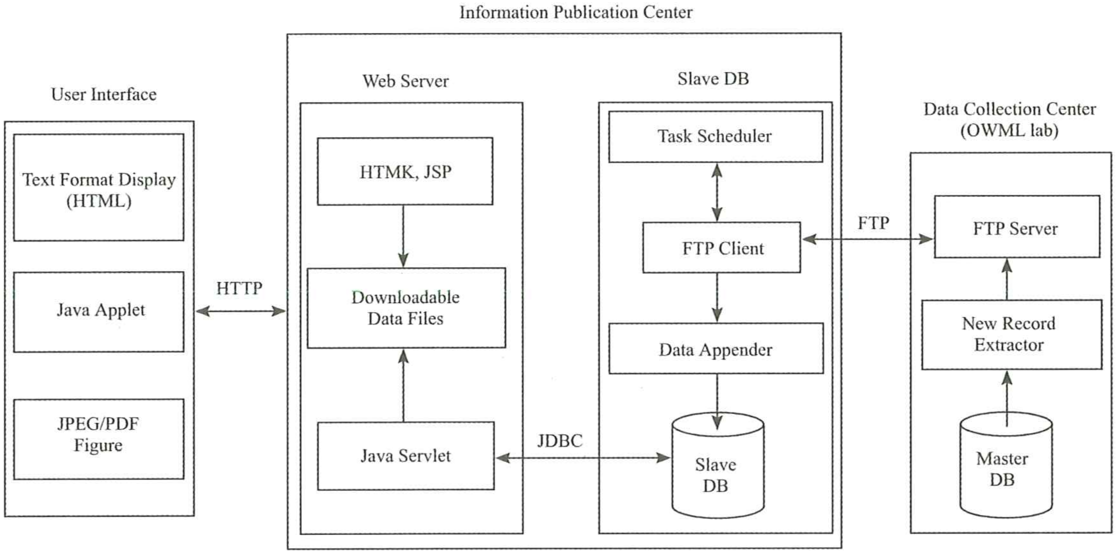
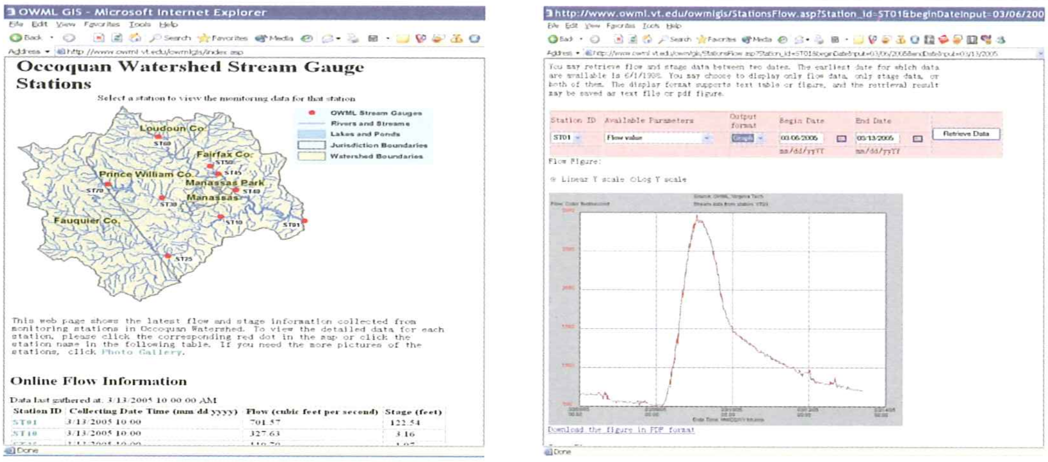


Figure 4. The software architecture of the watershed management system

performed on the web server and database. This will simplify client operations and reduce the maintenance costs. High performance is another advantage. In our system, most of the user requests are data-centric, requiring large amount of data but comparatively little CPU time. Thus, data processing should reside on the database to avoid transmitting large data packages through the network. The software components in our system can be grouped into three types according to their locations: User Interface, Information Publication Center, and Data Collection Center.

User interface: All query requests and results are displayed through a web browser. Two data display formats are supported, textual format and graphical format. In textual format, data are listed in a table, which can be downloaded as a text file. Graphical format displays changes in the data with figures in the form of curves or color maps. Users can choose to save the figure as either a JPEG or PDF file. Figure 5(a) shows the locations of the available monitoring stations and a table of their current readings. Users can click on the map or the station list to browse the data of a particular station. Figure 5(b) shows the query interface. Users can select a particular station, a time period, a set of attributes of interest, and the display format which they prefer to browse the data.

Information publication center: The information publication



(a) Station index

(b) Information for a particular station

Figure 5. User interface

center consists of a web server and the slave database. The web server is the container for static web pages, active web pages, and Java Servlets. The HTML files are used to display static information, such as the website introduction and project description. The active pages and Java Servlets are responsible for receiving user requests, converting them to SQL, communicating with the slave database, and translating the query results to various display formats. The query result can be saved as downloadable files in text, JPEG or PDF format. Java DataBase Connector (JDBC) is used as an open interface between Java Servlets and the database management system. The main task of the slave database is to synchronize with the master database. An FTP client programmed with Java is in charge of the data synchronization. Since the size of the master database is large, we employ an "incremental update" scheme to accomplish synchronization. Once the master database receives a batch of new data, it will generate a flat file containing these new data. An FTP client Java program on the web server will check frequently to see if the flat file on the master database server is new. If the file is new, it builds an FTP connection, fetches the file, and calls a "Data Appender" program to write the new data to the slave database. Since only increments of data are transmitted, this does not consume much bandwidth.

Data collection center: The data collection center is located at OWML and controls the data collection from the monitoring stations. The master database obtains readings from monitoring stations periodically, integrates and transforms them to the desired format, and builds an index for them. Finally the newly created data will be output as a flat file for synchronization with the slave database. There is always a possibility that some malfunction of the sensors or communication equipment may cause errors in the data. Engineers at OWML can manually correct these data and generate a data file. This file contains all the data being modified, to be inserted, and to be deleted. As the new data are synchronized, these corrected data update the slave database automatically.

IV. SYSTEM DEMONSTRATION

Our system can display both real-time and historical surface water monitoring data, such as flow and stage, upon request. These data are collected at a frequency of every 15 minutes. For rapidly changing situations, the data can be collected as often as every 5 minutes. We have accumulated more than 30 year of data for the Occoquan Watershed. From the 30-year historical data, some interesting patterns can be identified by using various visualization methods. Our system can present the flow and stage information in both tabular and graphical formats. By comparing the data collected from different stations, some associations can be discovered. In addition, statistics extracted from the historical data, such as minimum, average or maximum flow can help identify time periods when unusual weather patterns occur. Moreover, aggregation queries can support the presentation of data summaries in various granularities.

The functionalities of our system are demonstrated by the following examples..

Flow and stage illustration: Figure 6 shows the flow and stage fluctuations detected from station ST01 between Sept 15, 2004 and Sept 22, 2005. The X-axis denotes the time period and the Y-axis denotes the flow/stage value. There is a peak around September 20th, lasting from 1AM September 19 to 18PM September 21. During this period, Hurricane Isabel passed close by the Occoquan watershed. The accompanied rainfall contributed to the high flow values.

Data comparison between different stations: For stations located along the same river, the flow information collected by the upstream stations may have a close relationship with readings of the downstream stations. For example, Figure 7

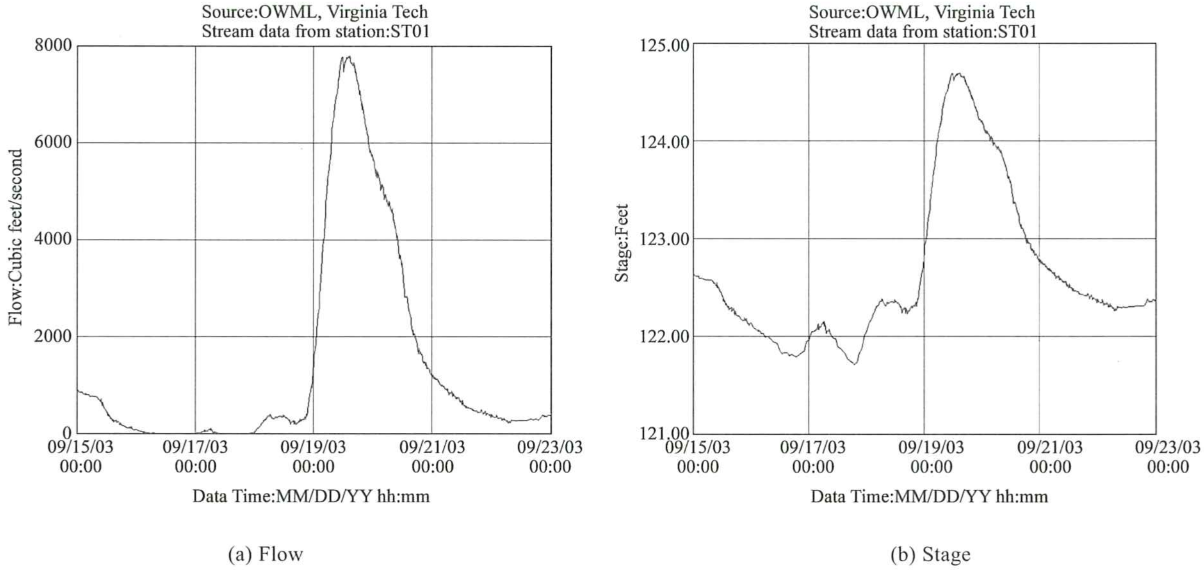


Figure 6. The flow and stage information for ST01

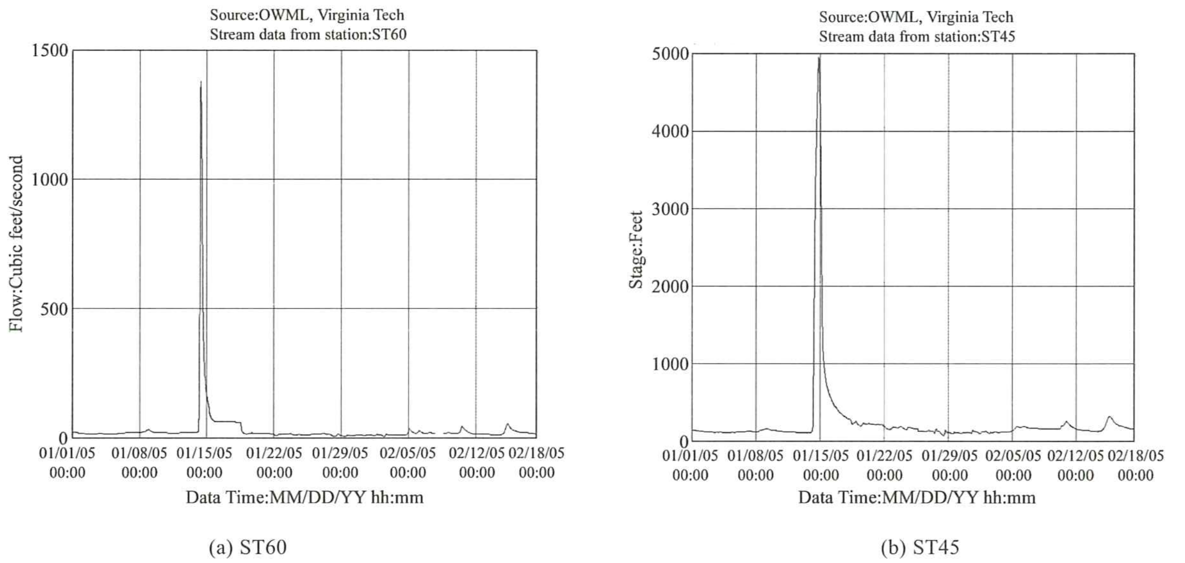
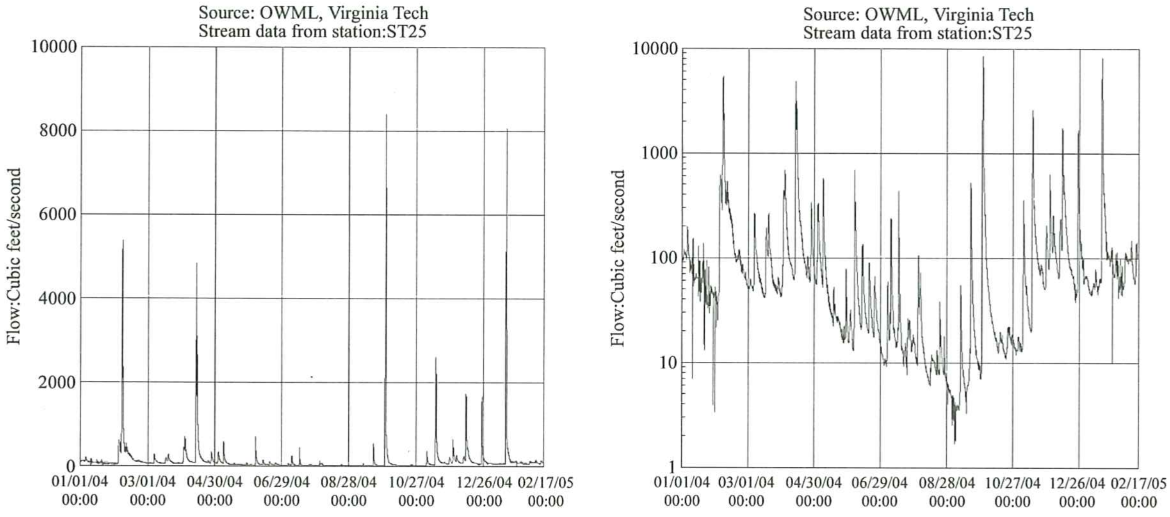


Figure 7. The flow comparison between ST60 and ST45

shows a flow comparison between station ST60 and station ST45. ST60 is located upstream and ST45 is located downstream of Bull Run. The X-axis denotes the time period and the Y-axis denotes the flow value. We can clearly observe that the peak flow detected by ST60 occurs at around 7AM Jan 14, 2005, whereas the peak flow detected by ST45 occurs at around 5PM Jan 14, 2005, a 9 hour gap. This relationship is useful for flood predictions. By collecting and analyzing data from different stations along the river, we can develop a model to predict future flow to the downstream and estimate the chance of flooding.

Log scale vs. Linear scale: For stream flow data over a long period, it is difficult to observe the details because the difference between the maximum and minimum values is so large that the data curve is extremely compressed. To address this problem, an option of displaying the data in log scale is provided. Figure 8 shows the comparison between linear scale and log scale for flow data collected by station ST25 between Jan 1, 2004 and Feb 17, 2005. The X-axis denotes the time period and the Y-axis denotes the original flow value (in linear scale) or the logarithm of the flow value (in log scale). With the linear scale, only the largest data values can be clearly shown, while the intermediate values are displayed in a small region and cannot be clearly differentiated. With a log scale, the data values are more evenly distributed in the figure, thus revealing more details. For example, the data values between 1AM June 29 2004 and 1AM August 28 2004 in Figure 8(b) can show

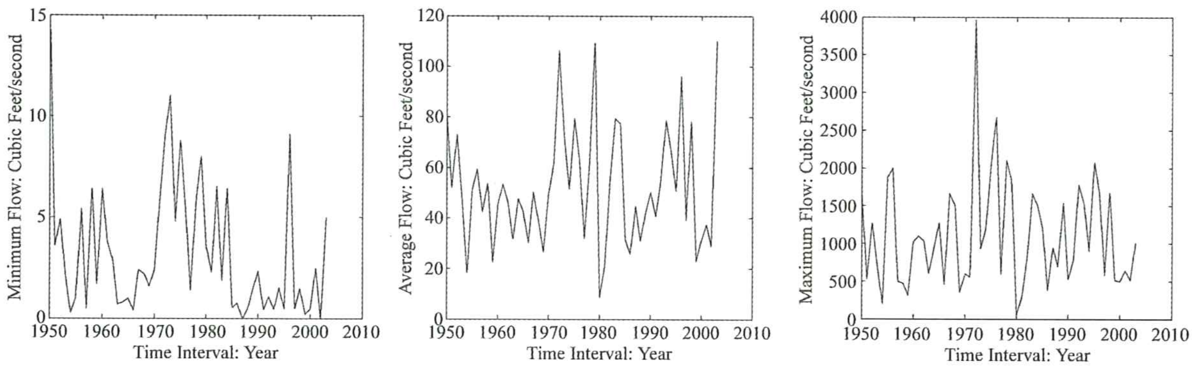


(a) Linear (b) Log

Figure 8. The linear and log scale display

considerably more details than that in Figure 8(a).

Illustration of Data Statistics: In some cases, statistics from the historical data can help identify interesting patterns. For example, which years were unusually dry and which years were particularly rainy. Figure 9 shows the 30-year statistics of the flow information collected by station ST70, including the minimum flow, the average flow, and the maximum flow. We can clearly recognize that for northern Prince William County where ST70 is located, 1980 was a dry year because its average flow was very low, as well as its minimum flow and maximum flow. 1973 was a rainy year considering its high values in the minimum, average, and maximum flow statistics.

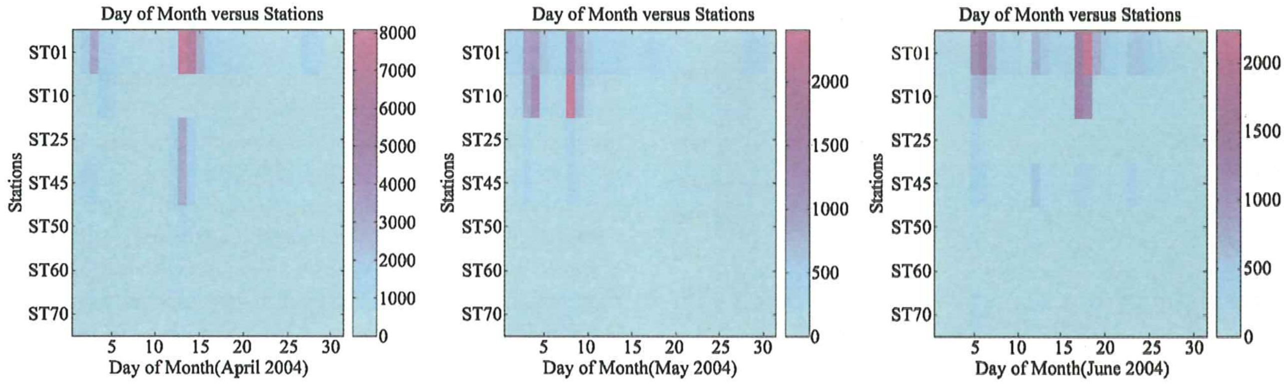


(a) Minimum flow (b) Average flow (c) Maximum flow

Figure 9. The minimum, average, and maximum flow for station ST70

Data cube: By integrating the data cube concept with an interactive graphic user interface, the spatial patterns and temporal trends can be identified easily. The concept of the data cube is the engine behind our system. A data cube is used to generate the union of a set of alphanumeric summary tables corresponding to a given hierarchy. For example, the flow data of a specific station can be aggregated to "Time of Day", T_{TD}, "Day of Month", T_{DM}, and "Month of Year", T_{MY}. Also, the different monitoring stations can be viewed along with another dimension. By combining S with the time dimension, we can recognize changes in the data in multiple stations simultaneously. Since different stations represent different geographical areas, the temporally aggregated data for various stations may reveal interesting spatial temporal association. Via rollup or drilldown operations, users can visualize data at different levels of the aggregation hierarchy. For example, aggregating the historical data to different granularities and displaying them in a single figure can help reveal interesting correlations among data from different time periods. Several data cube examples are described as follows.

- ST_{DM} **(Stations vs. Day of Month):** Figure 10 shows the relationship between different stations and different days in a particular month. The X-axis denotes the days of a



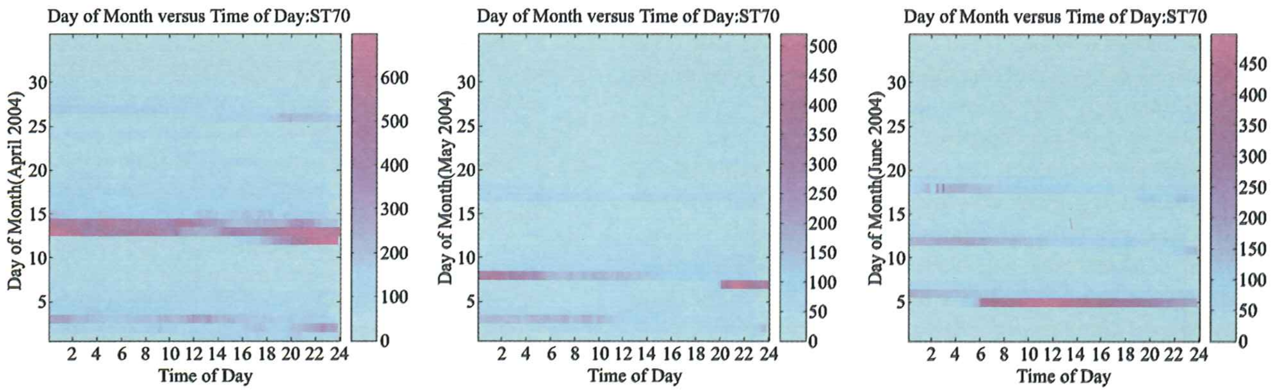
(a) April 2004 (b) May 2004 (c) June 2004

Figure 10. The data cube: stations versus day of month

month and the Y-axis denotes stations, from ST01 to ST70. The sub-figures (a), (b), and (c) illustrate the data cubes for April, May, and June in 2004, respectively. As shown in Figure 10(a), high flow values were detected by ST01, ST25 and ST45 on April 13th and 14th. In May 2004, Figure 10(b) shows that high flow values clustered in two periods: May 3rd to 4th, May 8th to 9th. These high values were detected by stations ST01 and ST10. Figure 10(c) shows that on June 17th and 18th, flow was very high for station ST01 and ST10, whereas other stations did not report high flow values. On June 5th, a fairly high flow was detected by all the stations.

- $T_{TD}T_{DM}$ **(Time of Day vs. Day of Month):** Figure 11 shows the flow information for station ST70 in two dimensions: Time of Day and Day of Month. The X-axis denotes the time of a day, and the Y-axis denotes the days of the month. The data cubes for April, May and June in 2004 are illustrated in sub-figures (a), (b) and (c), respectively. Figure 11(a) shows that in April 2004, ST70 detected comparatively high flow on three days, the 12th, 13th and 14th. On April 12th, the high flow occurred between 20PM and 24PM. On April 13th, the high flow occurred over two time periods, 0—3AM and 19—24PM. On April 14th, the high flow values were evenly distributed between 0AM and 12AM. In Figure 11(b), it can be seen that high flow values clustered around two time periods, 20—24PM on May 7th and 05AM on May 8th. Figure 11(c) shows that a high flow was on June 6th, 2004, starting from 6AM. The flow reached a peak at 12AM, and returned to a low level by 24PM. In addition, this figure reveals that in June 2004, there were four days with relatively high flow values, the 5th, 6th, 12th, and 18th.
- $T_Y T_{DY}$ **(Years vs. Day of Year):** Figure 12 shows a range of years (1950–2003) and the daily flow in each year for ST70. The X-axis shows the 365 days of a year and the Y-axis lists the years. In this figure, we can easily identify on a specific day in a year when a particularly high flow occurred. For example, near the 170th day in year 1973, there was an extremely high flow which was represented by the dark color.

The data cube has obvious advantages over ordinary visualization methods, such as curves, especially when we need to combine two different dimensions to obtain aggregated views of the original data. There are many possible



(a) April 2004 (b) May 2004 (c) June 2004

Figure 11. The data cube: time of day versus day of month(ST70)

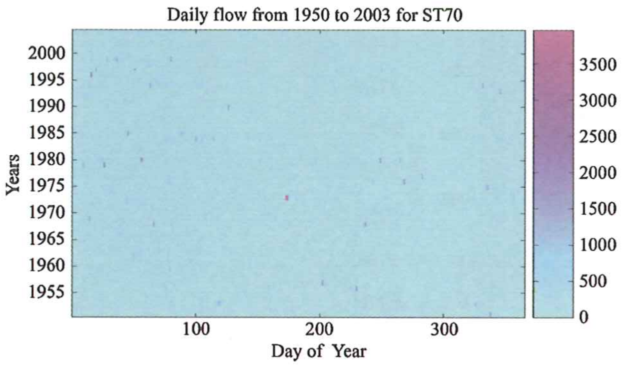


Figure 12. The data cube: years versus day of year

combinations of dimensions, for example, two different time granularities, or the spatial dimension versus the temporal dimension. With the help of data cubes, interesting patterns or relationships can be identified effectively.

V. DISCUSSION

Real-time data processing systems must collect and analyze data in a very short period of time. The quick response of a Web-GIS depends on two key factors: network speed and database processing speed. As the number of users increases, the performance of the network and database server may degrade. To reduce system delay effectively, users' requests need to first be analyzed so that different strategies can be used to handle different tasks. If the request involves large amount of data but the query result is small in size, server-side operations are preferable. It minimizes the cost of transmitting huge amount of data through the Internet, although it requires some computation resources on the server side. If the request is "computation intensive", requires small amount of data but long CPU time, a client-side scheme will be preferred. For example, as the demand for 3D visualization grows, we may need to transfer more data processing to the client side to reduce the server's burden. Our future plans include the provision of a dynamic request classification method in order to decide which scheme to use, client side or server side. This can be achieved by analyzing the user requests, for example, the requested data size and the desired output formats. The importance of real-time data processing lies not only on online monitoring and visualization, but also on trend analysis and prediction. It is important to be able to build a prediction model which can explore the correlations among multiple stations based on historical data, and extract practical association rules to predict the flood possibility or the impact of a pollution event. Spatial data mining techniques such as spatial outliers detection and spatial association rule identification will be employed to analyze the relationship between spatial attributes (location, altitude, etc.) and non-spatial attributes (e.g., flow, oxygen density). Also, weather data sets should be integrated into the prediction model, since precipitation also plays an important role in flood prediction.

Spatial data are much more complex than traditional transaction data. It contains not only non-spatial attributes such as numbers and characters but also spatial information such as location, size, and outline. Thus efficient visualization methods are in great demand for accurate and intuitive data representation. Currently, our system provides only 1D and 2D data visualization. An effort is underway to develop 3D representations, for example, using a data cube for TimeOfDay-DayOfMonth-MonthOfYear. Also, pie charts and histograms will be supported to illustrate the statistics of various attributes. A zooming function is currently under development, in order to provide multiple-resolution views for different purposes. For online pollution monitoring, detailed data may be required for every 5 minutes, whereas to see the flow pattern of a particular year, we need only daily flow data. A more intelligent visualization system should be developed so that the accuracy level of data can be dynamically adjusted according to the user's query requests.

In addition to the current stations, more monitoring equipment can be installed and more useful attributes can be collected for analysis and visualization. A multi-parameter water quality monitor will shortly be installed at one of the stations on Bull Run. This monitor will measure dissolved oxygen, water temperature, specific conductance, pH, and nitrate concentrations in Bull Run. In addition a buoy-mounted multi-parameter monitor may be installed in the Occoquan Reservoir in the near future. The buoy mounted system will be capable of making measurements through the water column to provide near real-time water quality depth profiles in the reservoir. Lastly, the OWML maintains a network of twelve recording rain gauges. Most of these are located in remote locations where phone lines are not available. With the appropriate wireless communication equipment, these gauges could be queried by the data collection server and have their information displayed on the web.

REFERENCES

[1] J. Gray, S. Chaudhuri, A. Bosworth, A. Layman, D. Reichart, M. Venkatrao, F. Pellow, H. Pirahesh, 1997, Data cube: A relational aggregation operator generalizing group-by, cross-tab, and sub-totals. *Data Min. Knowl. Discov.*, 1(1):29–53.

[2] R. Peng, 1997, An assessment of the development of internet GIS.

[3] R. H. Guting, 1994, An introduction to spatial database systems. *The VLDB Journal*, 3(4):357–399.

[4] S. Tu, X. He, X. Li, J. J. Ratcliff, 2001, A systematic approach to reduction of userperceived response time for gis web services. In GIS '01: Proceedings of the 9th ACM international symposium on advances in geographic information systems, ACM Press, pp.47–52.

[5] Z. K. Wei, Y. H. Oh, J. D. Lee, J. H. Kim, D. S. Park, Y. G. Lee, H. Y. Bae, 1999, Efficient spatial data transmission in web-based GIS. In WIDM '99: Proceedings of the 2nd international workshop on Web information and data management, ACM Press,

pp.38–42.

[6] S. Flick, 1996, An object-oriented framework for the realization of 3D geographic information systems. In JECGI '96: Proceedings of the second joint European conference & exhibition on geographical information (Vol. 1) : from research to application through cooperation, IOS Press, pp.187–196.

[7] V. Coors, S. Flick, 1998, Integrating levels of detail in a web-based 3D-GIS. In GIS '98: Proceedings of the 6th ACM international symposium on advances in geographic information systems, ACM Press, pp. 40–45.

[8] R. R. Vatsavai, T. E. Burk, B. T. Wilson, S. Shekhar, 2000, A web-based browsing and spatial analysis system for regional natural resource analysis and mapping. In GIS '00: Proceedings of the 8th ACM international symposium on advances in geographic information systems, ACM Press, pp.95–101.

[9] A. Stockus, A. Bouju, F. Bertrand, P. Boursier, 1999, Integrating GPS data within embedded internet GIS. In GIS '99: Proceedings of the 7th ACM international symposium on advances in geographic information systems, ACM Press, pp.134–139.

[10] C. du Mouza, P. Rigaux, 2002, Web architectures for scalable moving object servers. In GIS '02: Proceedings of the 10th ACM international symposium on advances in geographic information systems, ACM Press, pp.17–22.

[11] S. Tu, L. Xu, M. Abdelguerfi, J. J. Ratcliff, 2002, Achieving interoperability for integration of heterogeneous cots geographic information systems. In GIS '02: Proceedings of the 10th ACM international symposium on advances in geographic information systems, ACM Press, pp.162–167.

[12] O. Boucelma, J. Y. Garinet, Z. Lacroix, 2003, The virgis wfsbased spatial mediation system. In CIKM '03: Proceedings of the twelfth international conference on information and knowledge management, ACM Press, pp.370–374.

[13] J. Han, J. Y. Chiang, S. Chee, J. Chen, Q. Chen, S. Cheng, W . Gong, M. Kamber, K. Koperski, G. Liu, Y. Lu, N. Stefanovic, L. Winstone, B. B. Xia, O. R. Zaiane, S. Zhang, H. Zhu, 1997, Dbminer: a system for data mining in relational databases and data warehouses. In CASCON '97: Proceedings of the 1997 conference of the Centre for Advanced Studies on Collaborative research, IBM Press, pp.8.

[14] J. Han, K. Koperski, N. Stefanovic, 1997, Geominer: a system prototype for spatial data mining. In SIGMOD '97: Proceedings of the 1997 ACM SIGMOD international conference on management of data, ACM Press, pp.553–556.

[15] K. Koperski, J. Han, 1995, Discovery of spatial association rules in geographic information databases. In SSD '95: Proceedings of the 4th International Symposium on Advances in Spatial Databases, Springer Verlag, pp.47–66.