JOURNAL OF THE KOREAN SOCIETY OF AGRICULTURAL ENGINEERS, SEP. 2007, pp. 81-90

한국농공학회논문집
제49권 제5호

# Web-based GIS for Real Time Hydrologic Topographical Data Extraction for the Geum River Watershed in Korea

## Web 기반 GIS를 이용한 금강유역의 실시간 수문지형인자 추출

**Nam, Won Ho* · Choi, Jin Yong**,† · Jang, Min Won*** · B. A. Engel******

**남원호 · 최진용 · 장민원 · B. A. Engel**

**Abstract**

Watershed topographical information is required in hydrologic analysis, supporting efficient hydrologic model operation and managing water resources. Watershed topographical data extraction systems based on desktop GIS are abundant these days placing burdens for spatial data processing on users. This paper describes development of a Web-based Geographic Information Systems that can delineate the Geum River sub-basins and extract watershed topographical data in real time. Through this system, users can obtain a watershed boundary by selecting outlet location and then extracting topographical data including watershed area, boundary length, average altitude, slope distribution about the elevation range with Web browsers. Moreover, the system provides watershed hydrological data including land use, soil types, soil drainage conditions, and NRCS(Natural Resources Conservation Service) curve number for hydrologic model operation through grid overlay technique. The system operability was evaluated with the hydrological data of WAMIS(Water Management Information System) with the government operation Web site as reference data.

*Keywords : Web-based Geographic Information Systems, World Wide Web, Watershed Delineation, Hydrologic Topographical Data*

# I. Introduction

Collecting watershed topographical information to support watershed management is the first step for simulating hydrologic models and for calculating hydrologic components for understanding a watershed. The watershed topographical data collection for a watershed management is a quite time-consuming process when it is implemented using the traditional manual methodologies. Computerized mapping systems, called GIS (Geographical Information Systems) based on desktop computers, nowadays, have been used frequently for hydrologic data preparation, hydrologic

* Graduate Student, Dept. of Rural Systems Engineering, Seoul National University

** Assistant Professor, Dept. of Rural Systems Engineering, Seoul National University

*** Researcher, Research Institute for Agriculture & Life Sciences

**** Professor, Dept. of Agricultural and Biological Engineering, Purdue University

† Corresponding author. Tel.: +82-2-880-4583
Fax: +82-2-873-2087
E-mail address: iamchoi@snu.ac.kr

modeling and model implementation based on digital maps, and have become a generalized approach for hydrologic analysis(Turcotte et al., 2001). Nevertheless, the demands of real time Web-based GIS have risen due to the burdens of digital data gathering and processing, meanwhile Choi and Engel (2003) developed an efficient algorithm that can be used in the Internet environment.

The high-speed broadband Internet proliferation has stimulated operation of hydrologic models via this communication environment(Al-Sabhan et al., 2003). Internet-based hydrologic model operation is a useful way to serve scattered potential users who are requiring hydrologic impact analysis, and the analysis typically requires hydrologic topographical data that are bounded by a watershed. Therefore, to truly support and realize the Internet operation of hydrologic models in a straightforward manner from data preparation to interpretation of results, real time watershed topographical information extraction including watershed delineation functionality is required.

Development and operation of a Web-based watershed delineation system presents difficulties inherited from the Internet environment in terms of network communication speed and Web browser interface environment capabilities. Considering these limitations, realization of a Web-based watershed topographical information extraction system requires acceptable speed and accuracy while dealing with dynamic HTML pages for data capture, a large amount of geo-referenced spatial data processing and display, and two-way communication of data including coordinates of the geographic entities. Hence, a Web-based operational hydrologic topographical data extraction should not only integrate applications written in multiple languages including HTML, Java, C, and PERL and the large number of digital maps, but it should also overcome network bandwidth limitations to provide acceptable access time without sacrificing user interface convenience.

This paper describes a Web-based GIS that was developed to support watershed management using Web-GIS capability for internet map browsing, online watershed delineation application and hydrologic topographical data extraction tool through the Internet in real time, and demonstrates the system operability tested with the Geum River watershed.

## II. SYSTEM DEVELOPMENT

### 1. System Configuration

The typical Web-based system is comprised of a user interface, a server side engine application and client and server interfaces as described in Table 1. The user interface for this Web-based system was developed using the MapServer Web-GIS application(MapServer, 2001), HTML(Hyper Text Markup Language), JavaScript, and Java Applet. The Web user interface has menus that are using for graphical display control and links to other functions including print map, and query mode. The interface also uses HTML form protocols to submit Transverse Mercator coordinates with MapServer CGI(Common Gateway Interface) variable values. The watershed delineation and the hydrologic topographical data extraction capability were developed using the C language to manipulate the flow direction grid data and to facilitate rapid execution.

Most Web-based user interfaces can be regarded as an operating environment on the client side. When menu buttons are selected, the system communicates to the server via the network with a specified protocol. Therefore, the Web-

**Table 1 Description of the Web-based GIS components**

| System | CGI and programming language | Type | Functionality |
|---|---|---|---|
| user interface in Web browser | MapServer CGI, HTML, Java Applet, JavaScript | CGI, Web pages | - watershed selection<br>- outlet point selecting<br>- display map control<br>- print map, query mode |
| server side engine application | C | executable application | - watershed delineation<br>- hydrologic topographical data extraction |
| client and server interface | MapServer CGI, PERL | CGI application | - file management<br>- application operation<br>- HTML standard out |

based system requires communication to interface between the application and server, and naturally a network application that follows network protocols should be developed to complete the Web-based system. CGI applications usually implement the server-side operation including data capture from client, server program executions and result sending to client browser. Although several languages have been used for the data exchange between the client and server, in this study, PERL(Practical Extraction and Report Language) was adopted to connect the client and server side application. PERL is a typical language for CGI and supports the protocol for HTML form input and HTML standard out.

## 2. Web-GIS User Interface Using Mapserver

### MapServer CGI

The MapServer Web GIS application was selected as the CGI engine for developing the Web GIS user interface. The MapServer was originally developed by the University of Minnesota ForNet project in cooperation with NASA(National Aeronautics and Space Administration) and the MNDNR(Minnesota Department of Natural Resources). The MapServer, a CGI, is an open source development environment for building spatially enabled Internet applications. The CGI, running on the server side, provides a light weight page for the client side. Thus, if the server is powerful enough to control the processes from multiple connections within a reasonable connection time, it is the preferable choice to support potential users since concerns regarding client side computer capability and connection speed are minimized.

### Web-GIS User Interface

The MapServer CGI works in a similar manner as general CGI applications. Although the functionalities of MapServer are described in detail at http://mapserver.gis.umn.edu, the basic flow of how it works with HTML, digital maps and the network is described in Fig. 1. An HTML page with a query string with the HTML GET method sends variable values to the server, and the MapServer CGI parses the variables, reads digital maps described in a map file, draws maps and ancillary graphics, prepares an HTML page for publishing, and sends the HTML page to the client browser. This interaction continues for zooming in, zooming out, panning, and turning on/off map layers to locate the watershed. The user interface obtains coordinates for the outlet point and submits the coordinates to the watershed delineation CGI for watershed boundary generation and hydrologic topographical data extraction.

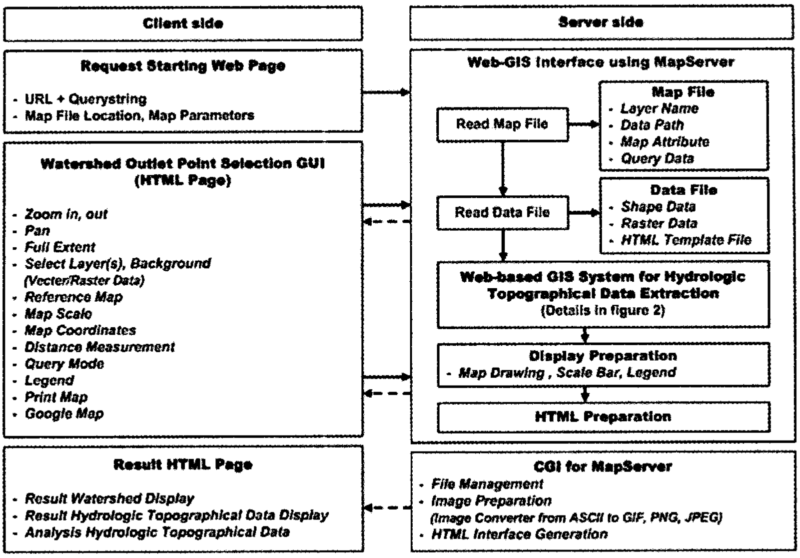


Fig. 1 Schematic diagram of the Web-based watershed delineation system development

## 3. Development of Hydrologic Topographical Data Extraction System

### Data Structure

Most data structures for topographical analysis involving hydrologic topographical data extraction and watershed generation are grid format, because the DEM(Digital Elevation Model), typical grid data for surface elevation, is commonly used for 3 dimensional data processing(Fairfield and Leymarie, 1991; ESRI, 1992; Jenson and Domingue, 1988; Martz and Garbrecht, 1992; Morris and Heerdegen, 1988). Hence, the ASCII(American Standard Codes for Information Interchange) grid format proposed by ESRI (1992) was adopted to facilitate generalization. The ASCII grid data contains a header section with basic information about the grid data including number of rows, number of columns, x and y coordinates of one of four corners, a cell size and the no data value. The header data are essentially used to interpret the grid data.

The flow direction grid data should be generated from a DEM that has been created by the sink filling function. Errors in DEMs are usually classified as either sinks or peaks. A sink is an area surrounded by higher elevation values, and is also referred to as a depression or pit. This is an area of internal drainage. Some of them are natural, but in many cases sinks occur due to imperfections in the DEM. Similarly, a spike or peak is an area surrounded by cells of lower value. These are commonly occurring features and are less detrimental to flow direction calculations(ESRI, 1992). The sinks are generally removed using the sink-filling processes of commercial and public GIS software(ESRI, 1992; GRASS, 1999; Lee et al., 1999; Kim et al., 2000).

### Watershed Delineation System

Various hydrologic topographical data are typically summarized based on delineated watershed boundaries. In this study, to delineate a watershed boundary, the double seed array replacement scheme developed by Choi and Engel (2003) was adopted and the algorithm requires an outlet point and flow direction grid as input(as shown in Fig. 2).

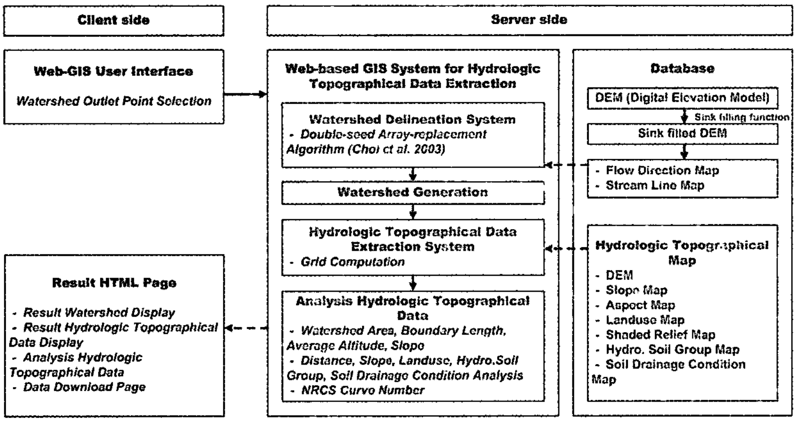


Fig. 2 Detail description of the Web-based watershed delineation including data and procedures (Choi et al., 2003)

Selecting the outlet point from an image in a Web browser is not as precise as from the display window of a desktop GIS, so a point search algorithm is needed to prevent the failure of the watershed delineation. In this study, the overlay map of stream networks and roads was used as the supplemental geographical information to help the user in selecting the outlet point. Even if the outlet point selected is not exactly on the stream network, the outlet search algorithm embedded in the watershed delineation program can locate the nearest stream, and the watershed delineation program uses the point as the outlet for the watershed boundary estimation.

### Watershed Topographical Data Extraction System

The watershed topographical data are required to implement hydrologic analysis, and if the data can be obtained automatically for a watershed that will greatly simplify the user's efforts in conducting hydrologic analyses. In fact, the results of the watershed delineation can be applied to clip the other grid data sets. To accomplish this, grid clipping applications and CGI were developed. The CGI runs the data extraction applications and results are sent by HTML standard out to the client browser.

To apply the results of watershed delineation within hydrologic analyses, a grid clipping module for the basic data extraction of watershed characteristics including boundary length, average altitude, average slope, areas about the reaches, slope distribution about the elevation, land use, shaded relief, hydrologic soil types, soil drainage conditions and NRCS(Natural Resources Conservation Service) curve number for direct runoff calculation.

The data prepared from the extraction applications were described in Table 2. The DEM, slope, land use, hydrologic soil group and soil drainage condition data were prepared and used. The watershed area and boundary length can be obtained though the C program using delineated watershed data, and the average altitude and average slope can be obtained using prepared DEM and slope data. The land use, hydrologic soil group and soil drainage condition data are created right after the watershed delineation utilizing grid computation, and the NRCS curve number can be created though the overlay these data. Users can download these results in ASCII grid format for their own analysis purposes using desktop GIS.

**Table 2 Extractable data sets**

| Data set | Source | Items |
|---|---|---|
| watershed area | delineated watershed | - |
| boundary length | delineated watershed | - |
| average altitude | used DEM | - |
| average slope | used slope data | - |
| DEM | sink filled DEM | altitude |
| flow direction | generated from a DEM | eight flow direction grid |
| shaded relief | generated from a DEM | - |
| land use | National Geographic Information Institute land use | seven land use categories including : water, fallow, paddy, uplands, pasture, forest and residential |
| hydro. soil group | generated from land use | hydrologic soil group : A, B, C and D |
| soil drainage condition | generated from land use | six soil drainage condition |
| NRCS curve number | combination of land use, hydrologic soil group and soil drainage condition that are extracted | NRCS curve number |

# III. APPLICATION OF SYSTEM

## 1. Data Preparation

### Site Description

To test the applicability of the system, the Geum River watershed in Korea was selected. The Geum River watershed is a watershed among the four big watersheds in Korea including fourteen sub-basins. The watershed is about 9,914 $km^2$ in size and located in the middle of the Korean Peninsula flowing from the east to the west.

### Data Description and Preparation

The GIS data including DEM, flow direction, slope, land use, hydrological soil group and soil drainage condition were prepared to implement the system and consisted of an ASCII grid with a 30m cell size. The DEM that was supplied by the Ministry of Environment was processed to remove the sinks, and the flow direction was created. The land use data was provided by the Ministry of Environment and was classified with seven categories: water, fallow, paddy, uplands, pasture, forest and residential. The soil drainage condition data was provided by the Rural Development Administration and was classified with six soil drainage condition: very poor, poor, mild poor, mild good, good, very good. The hydrological soil group A, B, C, and D indicates high, moderately high, moderately low, and low infiltration rate for rainfall, respectively.

## 2. Operation Description

### Web-GIS Interface

The hydrologic topographical data extraction system starts by the request of the URL http://ruralwater.snu.ac.kr from the client for the menu as shown in Fig. 3. The Web-GIS interface supports selected background and layers, map display control, reference map, legend and print map button. The layers included in the map display are important to provide location information in selecting the outlet point. For the system map interface, the layers are divided into two groups of data: background and foreground. The background data includes the DEM, shaded relief, land use map and hydrologic soil group map,

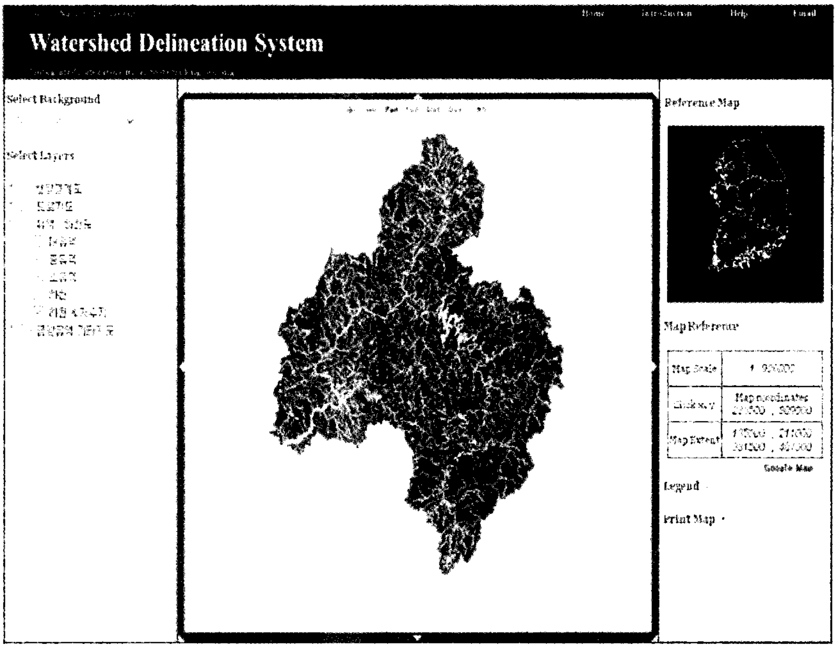


**Fig. 3 Map user interface of the real time watershed delineation system on the Web**

each of which can be used as a map background, and the foreground data includes roads, rivers, county boundaries and watershed boundaries.

Using the menu, users can choose the Geum River watershed and then select the outlet point of interest as shown in Fig. 3. After the outlet point is selected for Yong-Dam watershed, 3001 watershed number, the submit button is selected, and users obtain the generated watershed image and information including hydrologic topographical data maps(as shown in Fig. 4). Also, using the result button, users can obtain topographical data including watershed area, boundary length, average altitude, slope distribution about the elevation within the Web browser and hydrological data including land use, soil types, soil drainage conditions, and average NRCS curve number for hydrologic model(as shown in Table 3). Users can examine the result data that are prepared in tables and figures though their Web browser.

The watershed area in Yong-Dam watershed is about 926.66km$^2$ and the boundary length is 196.35km. The average slope of the watershed is about 16.82°, and the average altitude is about 508m. According to analysis of soil drainage condition, the watershed in Yong-Dam has good soil drain condition in about 84.61% of its area. The watershed consists of 70.52% forest, 10.87% paddy field, 10.49% upland, and 8.12% others and consists of 61.97% hydrologic soil group A, and 23.65% hydrologic soil group B. The watershed average CN is 48 under AMCⅡ.

## 3. Performance Evaluation

### System Performance Evaluation

Hydrologic topographical data extraction processing is often one of the more time consuming jobs of hydro-geographical analyses using GIS, and, in fact, the time required has been a significant reason why the application has not

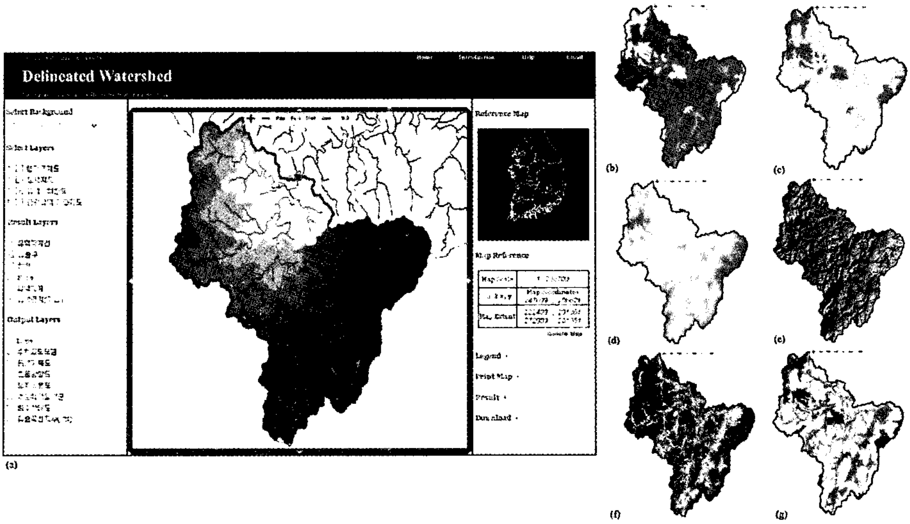


**Fig. 4 Web page for the delineated watershed output and extracted hydrologic topographical output display (a) Delineated watershed map display and map interface Web page (b) Soil drainage condition (c) Hydro. soil group (d) DEM (e) Shaded relief (f) Land use (g) NRCS curve number**

been available for Internet use(Pandey et al., 2001). Therefore, the time required for hydrologic topographical data extraction was obtained to determine the system performance via the Web. The running time is on the order of minutes from selecting the watershed delineation button in the Web page to receiving the results page includes the network connection, reading data, writing results, file conversion and sending and displaying HTML standard output. The total

**Table 3 Extracted results for hydrologic topographical data in Yong-Dam watershed**

| Topographical data | |
|---|---|
| area (km$^2$) | 926.66 |
| boundary length (km) | 196.35 |
| average altitude (m) | 508.00 |
| average slope (degree) | 16.82 |
| average NRCS curve number (under AMC II ) | 48.34 |

Analysis of soil drainage condition (Fig. 4. (b))

| soil drainage condition | very poor | poor | mild poor | mild good | good | very good |
|---|---|---|---|---|---|---|
| area (km$^2$) | 126.54 | 3.94 | 12.08 | 126.87 | 373.36 | 283.86 |
| rate (%) | 13.66 | 0.43 | 1.30 | 13.69 | 40.29 | 30.63 |

Analysis of hydro. soil group (Fig. 4. (c))

| hydro. soil group | A | B | C | D |
|---|---|---|---|---|
| area (km$^2$) | 574.23 | 219.13 | 6.75 | 126.54 |
| rate (%) | 61.97 | 23.65 | 0.73 | 13.66 |

Analysis of land use (Fig. 4. (f))

| land use | residential | paddy | uplands | forest | pasture | fallow | water |
|---|---|---|---|---|---|---|---|
| area (km$^2$) | 22.16 | 100.71 | 97.10 | 653.47 | 30.60 | 12.36 | 10.18 |
| rate (%) | 2.39 | 10.87 | 10.49 | 70.52 | 3.30 | 1.33 | 1.10 |

time is critical when a system is operated via the Internet because of network congestion. To determine the real time and overall performance of the system via the Web, the time data were measured under Internet operation conditions rather than the CPU time, since this is a more practical measure of overall program performance. The time variation with watershed size mainly came from the time required to write the delineated watershed file which depends on the number of grids in the delineated watershed.

### Comparison of Watershed Parameters

A comparison of watershed area and boundary length gathered from WAMIS (Water Management Information System, http://www.wamis.go.kr) and extracted values from the real time watershed delineation program were compared to evaluate the watershed delineation system accuracy. In this study the values from the WAMIS were regarded as the reference values. For the evaluation, seven watersheds were selected in the Geum River sub-basins to estimate the overall magnitude of error for watershed delineation. The areas, boundary lengths and differences for the selected watersheds are depicted in Table 4. The differences of areas vary from 0.21% to 5.47% with 1.32% average and the differences of boundary lengths vary from 2.69% to 9.56% with 4.99% average. The differences in area and boundary length estimation came from the resolution of the topographic data for the watershed selected. The Geum River watershed is making it difficult to determine the correct boundary given with 30m resolution DEM. Nevertheless, the results show a 98.6% mean accuracy in area estimation demonstrating the possibility of real time applicability.

**Table 4 Comparison of watershed parameters**

| watershed number | area[1] ($km^2$) | boundary length[1] (km) | area[2] ($km^2$) | boundary length[2] (km) | error of area[3] (%) | error of boundary length[4] (%) |
|---|---|---|---|---|---|---|
| 3001 | 930.35 | 188.73 | 926.66 | 196.35 | 0.40 | 4.04 |
| 3003 | 464.08 | 107.82 | 462.89 | 112.32 | 0.26 | 4.17 |
| 3005 | 664.62 | 196.20 | 628.27 | 204.94 | 5.47 | 4.45 |
| 3007 | 553.56 | 160.49 | 552.41 | 164.80 | 0.21 | 2.69 |
| 3009 | 648.98 | 142.99 | 647.58 | 147.85 | 0.22 | 3.40 |
| 3011 | 1855.83 | 270.79 | 1843.86 | 288.70 | 0.64 | 6.61 |
| 3013 | 666.10 | 179.87 | 679.91 | 197.07 | 2.07 | 9.56 |
| average | - | - | - | - | 1.32 | 4.99 |

1) Water Management Information System (http://www.wamis.go.kr)
2) Watershed Delineation System
3) |(colume 2) -(colume 4)| / (colume 2) * 100
4) |(colume 3) -(colume 5)| / (colume 3) * 100

## Ⅳ. SUMMARY AND CONCLUSIONS

A system for Web-based hydrologic topographical data extraction that can be operated in real time through the Internet was developed. The system is comprised of a program for hydrologic topographical data extraction, common gateway interface, user interface using MapServer CGI and geographical digital data. The double seed array replacement scheme was used to develop the watershed delineation program that delineates a watershed boundary using an outlet point and a flow direction grid data. The watershed generation scheme adopted identifies watershed boundaries through on-line operation via Internet using an outlet point selected from the Web-GIS graphical user interface. The Web-based watershed delineation system and the hydrologic topographical data extraction system were tested for the Geum River watershed. The speed and results were acceptable for use as a real time Internet system. The time variation depends on the number of grids contained in the delineated watershed. The time required for hydrologic topographical data extraction system was quite acceptable for application on the Web. The area differences between the watersheds delineated and WAMIS varied from 0.21% to 5.47% with a 1.32% average and the differences of boundary length vary from 2.69% to 9.56% with a 4.99% average for seven watersheds in the Geum River watershed. Also, users can obtain hydrologic topographical data including land use, soil types, soil drainage conditions, and average NRCS curve number for direct runoff calculation through the overlay of land use, soil type and soil drainage conditions data utilizing grid computation for the watershed. The system would be useful for the users who are willing to operate hydrologic models for the various watershed management purposes in terms of basic digital topographic data providing and preparing, especially in the Geum River watershed.

## References


1. Al-Sabhan, W., M. Mulligan & G. A. Blackburn, 2003, A real-time hydrological model for flood prediction using GIS and the WWW, Computers, Environment and Urban Systems, 27, pp.9-32.
2. Choi, J. Y. and B. A. Engel, 2003, Real Time Watershed Delineation System Using Web-GIS, Journal of Computing in Civil Engineering, American Society of Civil Engineers, ASCE, 17(3), pp.189-196.
3. ESRI, 1992, Environmental Systems Research Institute, Inc. 1992. GRID Command References.
4. Fairfield, J. and P. Leymarie, 1991, Drainage Networks from Grid Digital Elevation Models, Water Resources Research, 27, pp.709-717.
5. GRASS, 1999, GRASS 4.3 Reference Manual, Center for Applied Geographic & Spatial Research, Baylor University, USA.
6. Jenson, S. K. and J. O. Domingue, 1988, Extracting Topographic Structure from Digital Elevation Data for Geographic Information System Analysis, Photogrammetric Engineering and Remote Sensing, 54(11), pp.1593-1600.
7. Lee, D. I., G. S. Lee, H. J. Lee & G. S. Cho, 1999, A study on Watershed Auto-Delineation of Low Topographic Relief Landscape Using Topogrid, Korea Society of Surveying, Geodesy, Photogrammetry and Cartography, 17(3), pp.245-256 (in Korean).
8. MapServer, 2001, <http://mapserver.gis.umn.edu> (Nov. 1. 2001).
9. Morris, D. G. and R. G. Heerdegen, 1988, Automatically Derived Catchment Boundary and Channel Networks and Their Hydrological Applications, Geomorphology, 1(2), pp. 131-141.
10. Martz, L. W. and J. Garbrecht, 1992, Numerical Definition of Drainage Network and Subcatchment Areas from Digital Elevation Models, Computers and Geosciences, 18(6), pp.747-761.
11. Pandey, S., J. Harbor & B. A. Engel, 2001, Internet Based Geographic Information Systems and Decision Support Tools, Urban and Regional Information Systems Association, Park Ridge, IL, USA.
12. Turcotte, R., J. P. Fortin, A. N. Rousseau, S. Massicotte & J. P. Villeneuve, 2001, Determination of the Drainage Structure of a Watershed Using a Digital Elevation Model and a Digital River and Lake Network, Journal of Hydrology, 240, pp.225-242.