JOURNAL OF ENGINEERING AND TECHNOLOGY, Vol. 7 No. 2 December 2009 ISSN 1684-4114
http://www.iutoic-dhaka.edu/jet

Object-Oriented Hydro Information System (OHIS) for the Estimation and Visualization of Vegetation Water Content

S. Shams* and J. Huang**

Received 10 February 2009; Accepted after revision 24 November 2009

ABSTRACT

Vegetation indices are important parameters in remote sensing. In many scientific investigations, successful vegetation properties estimation heavily relies on proper use of vegetation indices. Vegetation water content (VWC) is one of the most important biophysical properties of vegetation that can be used in hydrological modeling, agricultural planning, forestry administration and even the climate change studies. However, to estimate VWC using remote sensing satellites data has been heavily relying on vendor dependent specific software rather than open source software. Hence, the paper highlights the development of a java-based open source information system integrated with MySQL Database to store, retrieve, query and analyze remote sensing data to estimate VWC. The hydro information system is platform independent, which enhances its great potential to be accessed from any computers, anywhere through the web browser. Taking its advantages, the system evaluated the performance of various vegetation indices in VWC estimation systematically. The short-wave infrared (SWIR) bands were proved to be superior to the other bands when corn and soybeans VWC are concerned under the similar environmental condition of SMEX02 (Soil Moisture Experiment in 2002).

Keywords: *Hydro Information System, GIS, Java, Vegetation Water Content.*

* *Civil and Environmental Engineering Department, Islamic University of Technology (IUT), Bangladesh. Email: s-shams@iut-dhaka.edu*

** *Rosenstiel School of Marine and Atmospheric Science, University of Miami, United State. Email: jhuang@rsmas.miami.edu*

1 INTRODUCTION

The environmental studies involve many types of parameters and variables related to water, soil and air, which are all dependant on each other's. A hybrid approach integrating remote sensing and GIS is one of the best approaches to look into the characteristics of such environmental parameters. GIS and Remote Sensing have been very beneficial in mapping [1] and [2] and data analysis [3], and hence greatly aided in the understanding and decision-making [4] in resource management. GIS has been intensively used in land registration, hydrology, land evaluation, planning or environmental observation [5], [6], [7], [8], [9], [10] and [11]. GIS has been widely used in water quality and impact assessment [12] and [13].

Geo-data comes in many different forms, such as maps or images and regarded as a remote sensing data when images are taken from the air or from the space. Remote sensing technology has many attributes that would be beneficial to detecting, mapping and monitoring the change of environment in a wider aspects [14], [15] and [16]. For example, the changes in vegetation contents through a growing season can be detected as radiance or reflectance at the sensors mounted on the satellites. The post processing products, including vegetation indices, can be used to determine the vegetation health conditions in a particular region in both spatial and temporal manners. As a very important parameter in hydrological process, vegetation water content (VWC) estimation using remote sensing had been attempted in decades. Vegetation indices, including the broadly used normalized difference vegetation index (NDVI), are the major indicators to model VWC. NDVI is based on the red (RED) and near infrared (NIR) bands, which are located in the strong chlorophyll absorption region and high reflectance plateau of vegetation canopies respectively. Using satellite spectrometers and the Airborne Visible/Infrared Imaging Spectrometer (AVIRIS) data, it has been found that vegetation indices (VI) based upon NIR and SWIR are better than those employing VIS and NIR when retrieving leaf water content information [17], [18], [19], [20] and [21].

Recent research demonstrated the potential of shortwave infrared band in VWC estimation and the utilization of operational MODIS data [22]. However, from the implementation point of view, users may have to rely on the remote sensing data processing packages heavily. Unfortunately most of the applications are platform and expensively software dependent and it seriously restricts the application of scientific findings of research. In this study, the use of different bands to form vegetation indices has been further tested systematically using object-oriented Hydro Information System (HIS) with an open source database such as MySQL. Object-orinted programming language uses objects to interact with each others to design the applications. The crucial point of developing an object-oriented system resides in the setting up of a standard communication protocol between the different components and these standard communication protocols can be identified as interfaces. An interface is defined as a set of methods and properties that can be accessed by other objects. The advantage of using this approach is that once an object is implementing a standard interface, it will immediately be ready for coupling with any other objects that comply with this standard. The HIS not only use these advantages of object-oriented programming but also illustrates the advantages of the SWIR based vegetation indices in VWC estimation. The HIS introduces a user friendly open source information system which has a great potential to process the remote sensing data and visualize user interested results locally or remotely with excellent platform or software independency.

2 THE STUDY AREA AND DATA SOURCE

The study area is the Walnut Creek Watershed in Iowa, US, where the SMEX 02 (Soil Moisture Experiment in 2002) [23] campaign took place. The corn and soybeans occupied 73.4% of the total area (39.5% corn and 33.9% soybeans), 12% as urban areas and roads, 14% as grasses and trees, 0.6% in trace pixels of other classes [22] and [24]. The MODIS (Moderate Resolution Imaging Spectro-radiometer) datasets, which are free of use to date, provides the remote sensing data source for this study. The suitability of MODIS data resolution for the study area had been discussed by Chen et al. [22]. Another reason for using MODIS dataset is its daily availability, which will significantly enhance the use of the current developed HIS with online database for real time implementation worldwide. The seven bands of MODIS data were retrieved for a DOY 182 (182^{nd} day of a year) and were used for vegetation indices calculation and VWC estimation thereafter.

3 DEVELOPMENT OF HIS AND ITS INTEGRATIONWITH VEGETATION MODELS

The methodology consists of development of an Information System capable of storing hydro-based GIS data compatible to OGIS (Open Geo-data Interoperability Specification). OGIS project, underway since June 1993, is an attempt to design methods that provide an object oriented architectural framework for access to geo-data, independent of the specific data structures and file formats used to model the data. The three major requirements for OGIS are interoperability, sharing data space, and availability of heterogeneous resource browser. The Java lies at the core of the system integrated with a MySQL

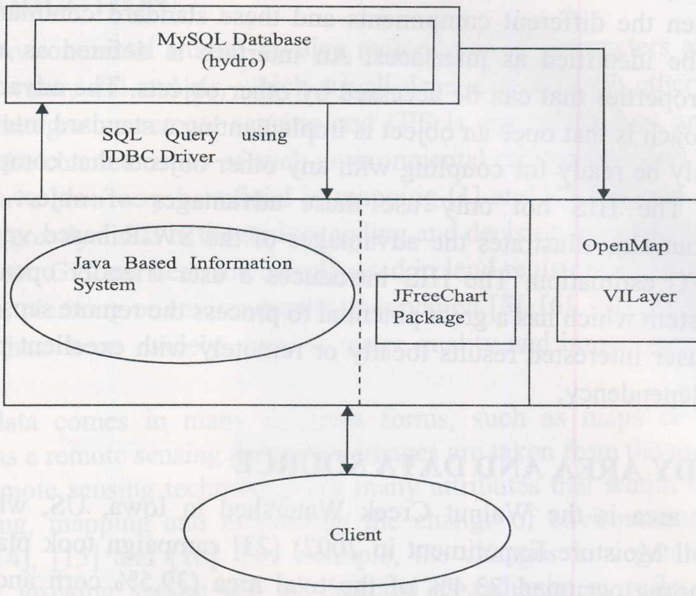

Figure 1: Basic Architecture of the System.

database, Open-Map [26] and J-Free-Chart package that is capable of doing analysis as done with ENVI [27] and Mat Lab software. The basic architecture of the system is shown in the **Fig.1**.

MySQL supports spatial extensions following the specification of the *Open GIS Consortium* (OGC) to allow the generation, storage, and analysis of geographic features. The specifications describe a set of SQL geometry types, as well as functions on those types to create and analyze geometry values. Therefore, MySQL has been used for spatial extention to demonstrate how remote sensing data can be used for estimating vegetation water content. A database name hydro was created using MySQL database features. The MySQL databasetable band1, band2, band3, band4, band5, band6 and band7 were loaded into the hydro database. Each band consists of pixel values for 70 rows and 191 columns. A vegeIndexFrame.java class was created which connects to the database and retrieves data in matrix format ($m \times n$, where m is number of rows and n is number of columns) and obtains value for each individual pixel from seven MODIS bands (centered at 648 *nm*, 858 *nm*, 555 *nm*, 860 *nm*, 1240 *nm*, 1640 *nm*, 2130 *nm* from band 1 to 7) and then calculate vegetation index such as the MODIS-derived Normalized Difference Water Indices ($NDWI_{1640}$, Chen et al. [22] using SWIR (Short Wave Infra Red) band (1640 *nm*) by the following formula:

$$NDWI_{1640} = (band2_{858nm} - band6_{1640nm}) / (band2_{858nm} + band6_{1640nm}).$$

The analysis of NDWI consists of values both for corn and soybeans. Since the focus of our studies is based on corn only, each pixel value is designated as 1 for corn pixels and 0 for the rest of the pixels. Classification procedures were conducted as in Huang [24] and Chen et al. [22]. The classified data is stored into the hydro database as a classification table. The NDWI values only for corn are calculated using the following formula:

$$NDWI_{corn} = NDWI * \text{ classification values}$$

The each individual pixel is recorded with specific latitude and longitude and is stored into the hydro database table, which consists of an id, location (spatial column that stores data in geometry point format as specified by OGC). There is no single command in MySQL that can store a series of point data as required for the above task. Therefore following two- step tasks are carried out to achieve this goal.

Step1: Import file in a "non-spatial" table (with coordinates as DOUBLE or INT columns), such as using the SQL command. The following statements use the codes for SQL command.

```
CREATE TABLE myImportTable (ID INT AUTO_INCREMENT, X DOUBLE,
Y DOUBLE, PRIMARY KEY (ID));
LOAD DATA INFILE 'myFileWithCoordinates.txt' INTO TABLE
myImportTable;
```

Step2: Insert values in a spatial column by constructing the 'spatial' string. The following statements use the codes for SQL command.

```
CREATE TABLE mySpatialTable (ID INT AUTO_INCREMENT,
myGeometryColumn GEOMETRY, PRIMARY KEY (ID));
INSERT INTO mySpatialTable (myGeometryColumn) SELECT
GeomFromText(CONCAT('POINT(',X,' ',Y,')')) FROM myImportTable;
```

The table's id_location and value_ndwi are joined together and the data are inserted into a new table result_ndwi which records values as well as pixels location. The java based GUI is connected with MySQL database and the user is prompted with a dialog box containing a password, once connected to the database the submenus are activated such as NDWI, Vegetation Water Content (VWC), VWC Comparison, Corelation and RMSE.

4 RESULTS AND DISCUSSION

The VWC for each individual pixel was calculated using the following formula [22] and [24]:

$$VWC = 9.44 * NDWI_{1640} + 1.37$$

When the user selects the NDWI sub-menu, an Interface pops up showing the band types to be chosen as shown in the **Fig.2**. Once the bands are selected the user clicks the NDWI button and it calculates NDWI values. Using Reset Button the user can deselect and get back to the initial band set up.

When the user clicks the sub-menu "Vegetation Water Content (VWC)" of the menu "Vegetation Index" it shows data for VWC in tabular format as shown in **Fig.3**.

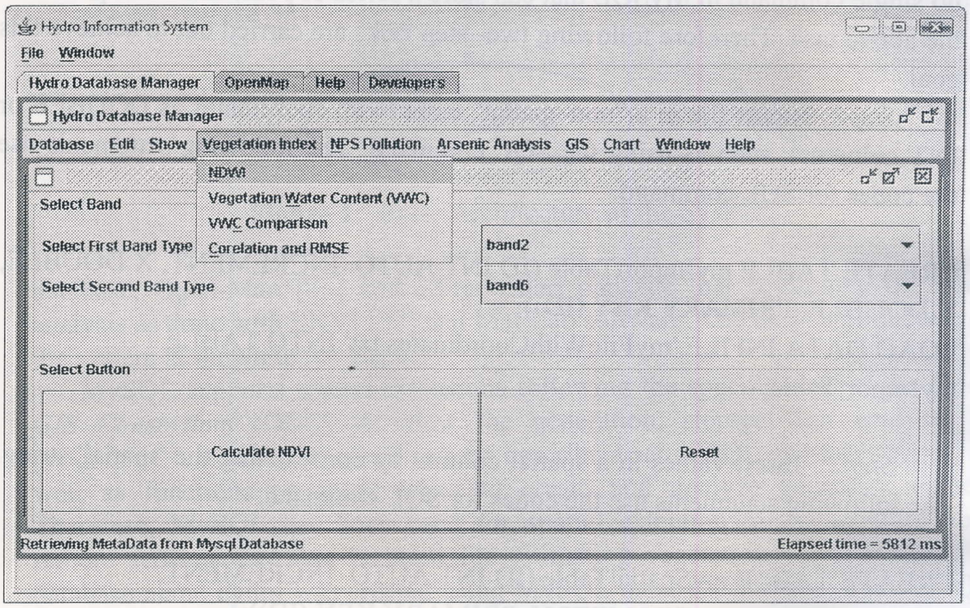

Figure 2: Showing the Interface when user clicks the NDWI sub-menu.

To display VWC, the system is linked with Open-Map which allows users to visualize and manipulate geospatial information. This is accomplished by creating a layer called VI which access data from the table named value VWC of MySQL's hydro database and displays the band spectrum of the pixel values as shown in **Fig.4**. The Open-Map calls the open-map VWC properties file to for layout features of the Open-Map with relevant data as shown in **Fig.4**. The pixels were displayed with grids using create-Box-Shape method and pixel values were classified using Java RGB (a mixture of three additive primaries:

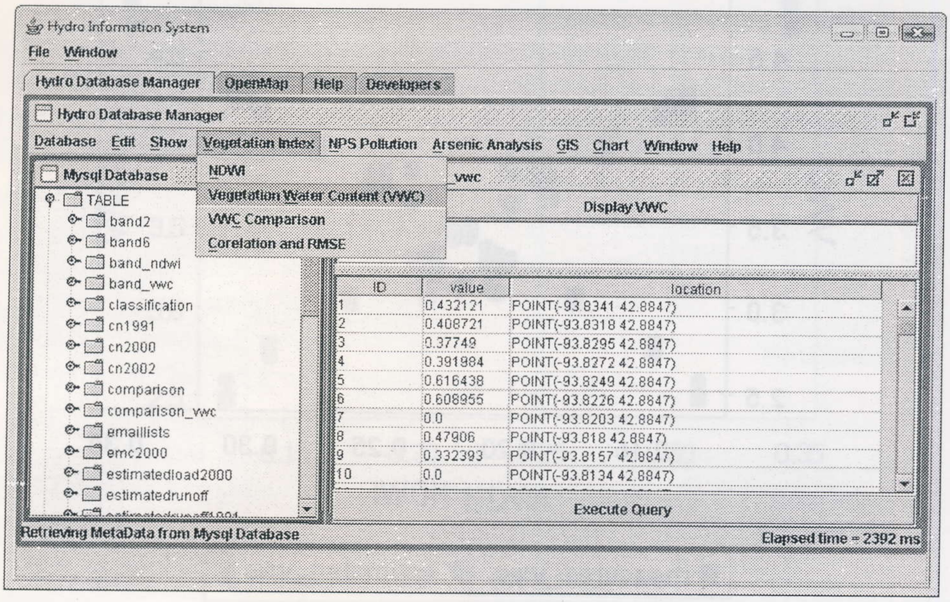

Figure 3: Showing measured and estimated VWC data in tabular format with its specific location as geo-referenced point.

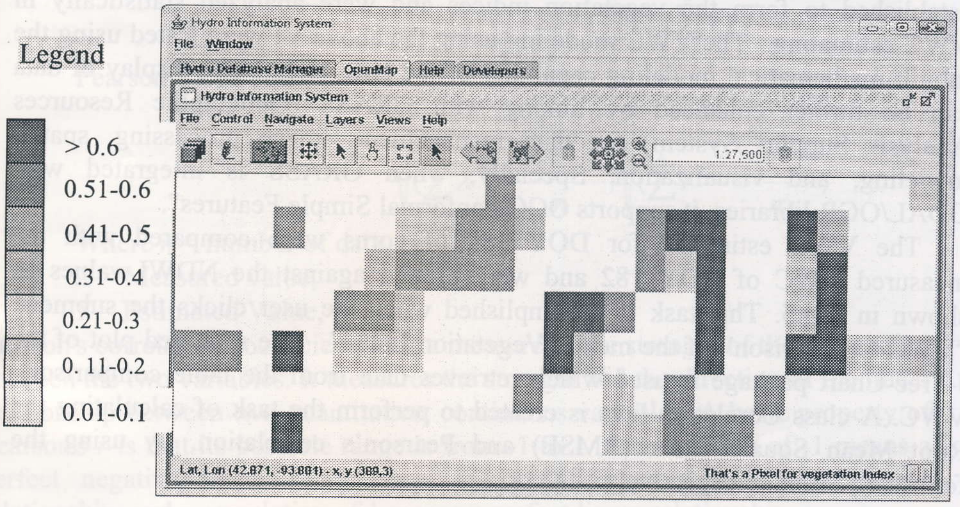

Figure 4: Showing VWC values, the dark color indicates higher VWC, clicking on the pixel; it shows the pixel id and calculated value.

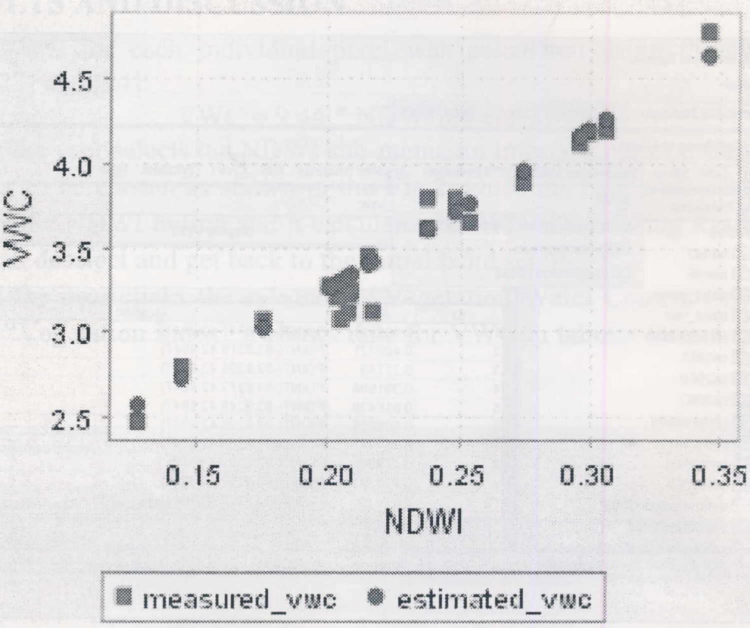

Figure 5: Showing comparisons between measured and estimated VWC.

red, green and blue) model. All the combinations among the seven bands were established to form the vegetation indices and were analyzed statistically in VWC estimating. The VWC modeling using the above VI were tested using the inbuilt mathematical modeling capability of the HIS itself. "The display of data can be further enhanced by linking with GRASS (Geographic Resources Analysis Support System) for data management, image processing, spatial modeling, and visualization. Specially, when GRASS is integrated with GDAL/OGR libraries, it supports OGC-conformal Simple Features".

The VWC estimated for DOY 182 of corns were compared with the measured VWC of DOY 182 and were plotted against the NDWI values as shown in **Fig.5**. This task is accomplished when the user clicks the submenu "VWC Comparison" of the menu "Vegetation Index". The scattered plot of the J-Free-Chart package is used which retrieves data from the table comparison_ VWC. A class Corelation.Java is created to perform the task of calculating the Root Mean Square Error (RMSE) and Pearson's correlation, by using the following formula respectively.

$$\text{Root Mean Square Error, } RMSE = \sqrt{\frac{\sum_{1}^{n}(X-Y)^2}{n}}$$

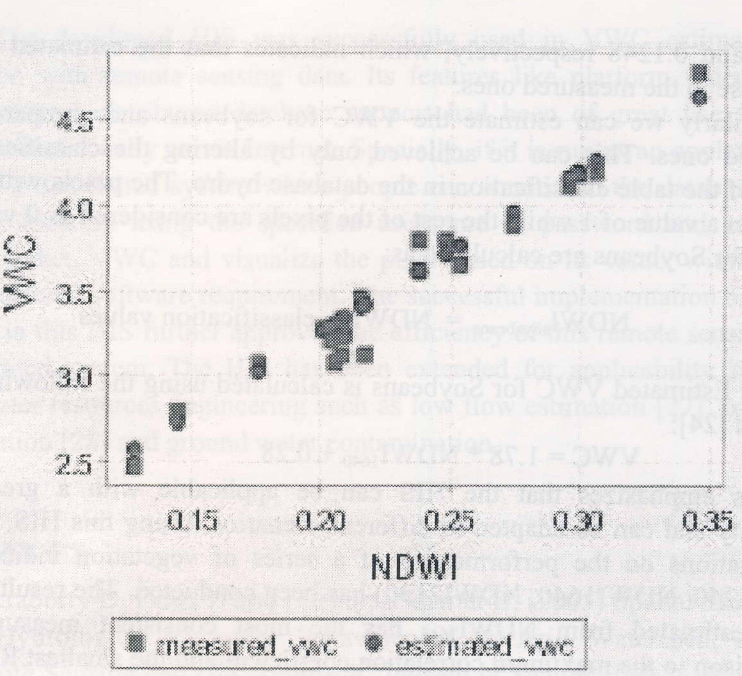

Figure 6: Showing the relationship between the measured and estimated VWC based on calculated correlation coefficient and RMSE.

Pearson's Correlation, $$r = \frac{\sum_{1}^{n} XY - \frac{(\sum_{1}^{n} X)(\sum_{1}^{n} Y)}{n}}{\sqrt{(\sum_{1}^{n} X^2 - \frac{(\sum_{1}^{n} X)^2}{n})}\sqrt{(\sum_{1}^{n} Y^2 - \frac{(\sum_{1}^{n} Y)^2}{n})}}$$

Where n = number of data,
X = Measured value,
Y = Estimated Value,

Pearson's correlation coefficient *(r)* is a measure of the strength of the association between the two variables. It measures the strength and the direction of the linear relationship between two quantitative, continuous variables. A basic property of Pearson's r is that its possible range is from -1 to 1. A correlation of -1 means a perfect negative linear relationship, a correlation of 0 means no linear relationship, and a correlation of 1 means a perfect linear relationship.

When the user clicks the sub-menu "Correlation and RMSE" under the menu "Vegetation Index" it appears with a frame consisting a "Calculate" button and by pressing this button it calculates RMSE and correlation coefficient as shown in **Fig.6**. The RMSE and Pearson's Correlation, r for the VWC are calculated as

0.9713 and 0.1248 respectively, which indicates that the estimated values are very close to the measured ones.

Similarly we can estimate the VWC for soybeans and compare with the measured ones. This can be achieved only by altering the classification pixel values of the table classification in the database hydro. The pixels with soybeans are given a value of 1 while the rest of the pixels are considered as 0 values. The NDWI for Soybeans are calculated as:

$$NDWI_{soyabeans} = NDWI * \text{ classification values}$$

The Estimated VWC for Soybeans is calculated using the following formula [22] and [24]:

$$VWC = 1.78 * NDWI_{1640} + 0.28$$

This emphasizes that the HIS can be applicable with a great deal of flexibility and can be adapted to different scenarios. Using this HIS, systematic investigations on the performances of a series of vegetation indices (NDVI, NDWI1240, NDWI1640, NDWI2130) has been conducted. The results indicated VWC estimated from $NDWI_{1640}$ has the most consistent measurements in comparison to the maximum correlation coefficient and the smallest RMSE. This further confirmed the results discovered in Huang [24] and Chen et al. [22that SWIR band (1640 *nm*) has better potential to estimate VWC with simple linear models. "However, it is recommended to use Enhanced Vegetation Index (EVI) as it corrects for some distortions in the reflected light caused by the particles in the air as well as the ground cover below the vegetation. The EVI data also does not become saturated as easily as the NDVI when viewing rainforests and other areas of the Earth with large amounts of chlorophyll".

5 CONCLUSIONS

The developed system is based on Open Source tools like Open-Map and MySQL database therefore any new development can be easily incorporated by modifying the source code. The developed information system can be connected to any types of databases (MySQL, PostGreSQL, Oracle, Sysbase) and data can be queried simultaneously and visualized in tabular format. The system can be replaced as an alternative of using any closed or vendor dependent software likes ENVI and MATLAB to do matrix analysis and to calculate vegetation index. The HIS turns out to be a single environment where VWC analysis, estimation and visualization all are computed under a single information system saving considerable time and money which would have been spent to perform similar tasks under a multiple environments and software systems. From the methodology point of view, the analysis can be extended to various crop types, or vegetation species. Corn and soybean are the examples first tested in this system. The developed HIS was successfully used in VWC estimation in coordination with remote sensing data. Its features like platform independent open source and distributed database support had been of great use in such modeling, either locally or via Internet. Specially, if it is run as an applet, it can be used as a distributed system where people can connect to the database from the remote location using the specified username or password and calculate vegetation index, VWC and visualize the pixel based on its value, without any local professional software requirement. The successful implementation of VWC estimation in this HIS further approved the efficiency of this remote sensing and GIS integrated system. The HIS has been extended for applicability in other areas of water resources engineering such as low flow estimation [27], pollutant load estimation [28] and ground water contamination.

REFERENCES

[1] Chakraborty D, Dutta D and Chandrasekharan H. (2005) Spatial Modelling for Hydrological Response Behaviour of an Arid Watershed, India - Remote Sensing and GIS Approach. *Journal of Spatial Hydrology*, Vol. 5, no. 1: pp. 46-66.

[2] Yang X and Liu Z. (2005) Using satellite imagery and GIS for land use and land cover change mapping in an estuarine watershed. *International Journal of Remote Sensing,* Vol. 26, no. 23: pp. 5275-5296.

[3] Benz UC, Hofmann P, Willhauck G, Lingenfelder I and Heynen M. (2004) Multi-resolution, object-oriented fuzzy analysis of remote sensing data for GIS-ready information. *Journal of Photogrammetry & Remote Sensing,* Vol. 58: pp. 239– 258.

[4] Bojorquez-Tapia LA, Diaz-Mondragon S and Ezcurra E. (2001) GIS-based approach for participatory decision making and land suitability assessment. *International Journal of Geographical Information Science,* Vol. 15 no. 2: pp.129 – 151.

[5] Bourletsikas A, Baltas E and Mimikou M. (2006) Rainfall-Runoff Modeling for an Experimental Watershed of Western Greece Using Extended Time-Area Method and GIS. *Journal of Spatial Hydrology,* Vol. 6, no. 1: pp. 93-104.

[6] Jain MK, Kothyari CU and Raju KGR. (2004) *A GIS based distributed rainfall-runoff model. Journal of Hydrology*, Vol. 299: pp. 107-135.

[7] Karssenberg D and De Jong K. (2005) Dynamic environmental modelling in GIS: 1. Modelling in three spatial dimensions. *International Journal of Geographical Information Science,* Vol. 19, no. 5: pp. 559-579.

[8] Kiker GA, Bridges TS, Varghese A, Seager TP and Linkovjj I. (2005) Application of Multicriteria Decision Analysis in Environmental Decision Making. *Integrated Environmental Assessment and Management,* Vol. 1, no. 2: pp. 95–108.

[9] Li X and Yeh AG. (2002) Neural-network-based cellular automata for simulating multiple land use changes using GIS. *International Journal of Geographical Information Science*, Vol. 16, no. 4: pp. 323-343.

[10] Verburg PH and Veldkamp A. (2005) Introduction to the Special Issue on Spatial modeling to explore land use dynamics. *International Journal of Geographical Information Science,* Vol. 19, no. 2: pp. 99-102.

[11] Zerger A and Wealands S. (2004) Beyond Modelling: Linking Models with GIS for Flood Risk Management. *Natural Hazards,* Vol. 33, no. 2: pp. 191-208.

[12] Shams S, Langaas S and Bhattacharya P. (2004) A Prototype Spatial Expert System for Evaluation and Mitigation of Groundwater Contamination: the Case of Arsenic in Bangladesh. *Proceedings from the 4th International Symposium on Environmental Hydraulics & 14th Congress of the Asia & Pacific Division of the International Association of Hydraulic Engineering and Research on December 15-18*, 2004, pp. 789-794. Hong Kong: The University of Hong Kong.

[13] Shams S, Bhattacharya P, Langaas S, Jacks G, Ahmed KM, Thunkvist E-L, Chen D and Granlund A. (2006) Mapping and Interpretation of field data for Evaluation and Mitigation of Groundwater Arsenic Contamination in Bangladesh. *International Journal of Hydroinformatics,* Vol. 8, no. 1: pp. 25-36.

[14] Baban SMJ and Yusof KW. (2001) Mapping land use/cover distribution on a mountainous tropical island using Remote Sensing and GIS. *International Journal of Remote Sensing*, Vol. 22, no. 10: pp. 1909-1918.

[15] Koukoulas S and Blackburn GA. (2005) Mapping individual tree location, height and species in broadleaved deciduous forest using airborne LIDAR and multi-spectral remotely sensed data. *International Journal of Remote Sensing*, Vol. 26, no. 3: pp. 431-455.

[16] Weng Q. (2002) Land use change, analysis in the Zhujiang Delta of China using satellite remote sensing, GIS and stochastic modeling. *Journal of Environmental Management*, Vol. 64: pp. 273–284.

[17] Ceccato P, Flassee S, Tarantola S, Jacquemoud S and Gregoire JM. (2001) Detecting vegetation leaf water content using reflectance in the optical domain. *Remote Sensing of Environment*, Vol. 77: pp. 22–33.

[18] Rollin EM and Milton EJ. (1998) Processing of high spectral resolution reflectance data for the retrieval of canopy water content information. *Remote Sensing of Environment*, Vol. 65: pp. 86– 92.

[19] Serrano L, Ustin SL, Roberts DA, Gamon JA and Penuelas J. (2000) Deriving water content of chaparral vegetation from AVIRIS data. *Remote Sensing of Environment*, Vol. 74: pp. 570–581.

[20] Sims DA and Gamon JA. (2002) Estimation of vegetation water content and photosynthetic tissue area from spectral reflectance: A comparison of indices based on liquid water and chlorophyll absorption features. *Remote Sensing of Environment*, Vol. 84: pp. 526–53.

[21] Ustin SL, Roberts DA, Pinzon J, Jacquemoud S, Gardner M and Scheer G. (1997) Estimation canopy water content of chaparral shrubs using optical methods. *Remote Sensing of Environment*, Vol. 65: pp. 280– 291.

[22] Chen D, Huang J and Jackson TJ. (2005) Vegetation water content estimation for corn and soybeans using spectral indices derived from MODIS near and short-wave infrared bands. *Remote Sensing of Environment*, Vol. 98, no. 2-3: pp. 225-234.

[23] Soil Moisture Experiment in 2002, *http://hydrolab.arsusda.gov.smex02/* (accesed on Jauary 29, 2004).

[24] Huang J. (2006) *Vegetation properties relationships from Spectral bands and vegetation indices from operational satellites*. PhD thesis, The University of Manchester, Manchester, UK.

[25] OpenMap. *http://openmap.bbn.com/* (accesed on Jauary 22, 2005).

[26] ENVI - The Remote Sensing Exploitation Platform. *http://www.ittvis.com/envi/* (accesed on March 9, 2007).

[27] Leone A, Shams S and Chen D. (2006) An Object-Oriented and OpenGIS Supported Hydro Information System (30-HIS) for Upper Mersey River Basin Management. *Journal of River Basin Management,* Vol. 4 no. 2: pp. 1-9.

[28] Shams S, Zakzok E, Chen D and Huang J. (In Press). Estimation and monitoring non-point source pollutant loads: an object oriented hydro information approach. *J. Environment and Waste Management* (Acceptance of Manuscript Number: IJEWM-07-2906-Cc).