HDI4NVN – Hydrological Data Infrastructure for North Vietnam

Maria A Brovelli[1], Truong Xuan Quang[1], Gérald Fenoy[2]
[1] DIIAR, Politecnico di Milano – Como Campus (Como), ITALY.
maria.brovelli@polimi.it, xuan.truong@mail.polimi.it
[2] GeoLabs SARL, Futur Building 1, 1280, Avenue des Platanes 34970 Lattes, FRANCE.
gerald.fenoy@geolabs.fr

Keywords: Web Processing Service. Open Geospatial Consortium. istSOS. Sensor Observation Service. ZOO Project. Web Processing Service. Open Source Geospatial foundation

ABSTRACT

Global climate change is increasing the risk of exposure of million of people all over the world at water stress and/or extreme events of flood and drought. The research aims at building a spatial infrastructure for monitoring the hydrological status of an area in the Northern part of Vietnam. The area is of particular interest, being affected by natural calamities like droughts, floods, landslides and so on.

In this paper we present the intelligent geoportal and the OGC Web Services (WMS, WFS, SOS, WPS) we implemented to manage, analyze and visualize about 50 years of hydrological and environmental data. The intelligent geoportal, which is a gateway to the hydrological data infrastructure, is completely based on Free and Open Source Software (FOSS); it makes available the following functionalities: (1) storing and managing base maps such as land used and land cover generated form free Landsat TM, DEM, aspect, hillshade from ASTER and SRTM and sensor observations, (2) searching and querying data or observations including register sensors, insert observation values of hydrology and environmental data to a database management system, (3) getting information such as statistical analysis and statistical charts from hydrological and environmental sensors at daily, monthly and yearly intervals, (4) comparing the rainfall data collected by means of rain gauges at ground level with those remote-sensed (PERSIANN).

1. INTRODUCTION

Climate change has been affecting the higher frequency and intensity of natural disaster in the world. It is remarkable that major disasters occurred in developing countries, where there is lack of information, facilities, data acquisitions and people often face with other limitations such as: forecasting technologies, lack of sensor networks and networks of ground-based hydro-meteorological observations frequently sparse.

National Centre for Hydro-Meteorological Forecasting, which is the Vietnam Institute of Meteorology Hydrology and Environment (NCHMF), aims at meeting all the requirements for the disaster prevention and preparedness, the socio-economic development and national security in Vietnam. NCHMF used Mapper.js library (at http://www.netzgesta.de) in order to add automatic area highlighting to image maps and it also supports users to request hydro-meteorological data and weather forecast at daily, monthly intervals (see at NCHMF). However, the NCHMF Web-GIS cannot satisfy the requirements of many end-users and particular scientists because there is no way for users to access data in the past, and it is also missing statistical tools, flexibility search and query. Otherwise there is need to improve quality and quantity of base map. Moreover, there is the nees to build management, analysis, and visualization tools for historical hydrological and environmental data.

The Open Geospatial Consortium (OGC) is a non-profit, international, voluntary consensus standards organization that is leading the development of standards for geospatial and location based services. OGC offers standardized OGC Web Services (OWS) such as Web Map Service (WMS) providing dynamic generation of maps, Web Feature Services (WFS) for requesting features. Beside the most known web services above, OGC Sensor Web Enablement (SWE) presents many opportunities for connecting real-time of heterogeneous sensor data to the internet. Otherwise, one of the most recent interoperability standards is

OGC Web Processing Service (WPS). WPS provides rules for standardizing input and output data, presents how client can request the execution of a process and output can be handled for a same process.

2. DATA USED

The main study area is in the Northern part of Vietnam; it involves the Red River Basin, which is the second largest basin in Vietnam. The data used can be subdivided into two distinct groups: the former group namely is remote sensed data (RS) and GIS (Geographic Information System) data, the second group consists of measured data.

Type	Name	No of Senses/maps	Note
Base map	Landsat TM 7	12	30m
	ASTER	17	30m
	SRTM	4	90m
ESRI shapefiles	Thematic maps	25	
Precipitation	PERSIANN	--	0.25 °
Gauge or Sensor (GoS)	Water level	6	daily
	Discharge	9	daily
	Rainfall	19	daily
	Evaporation	15	daily
DAM	Inflow	1	daily

Table1. Data used

Almost RS and GIS data are free data they are from various sources; measured data was issued by the Institute of Water Resource Planning of Vietnam (WRP). The data used can be summarized in the (Table 1). PERSIANN was downloaded at http://chrs.web.uci.edu/ with last visited May 2012; PERSIANN (Precipitation Estimation from Remotely Sensed Information using Artificial Neural Networks). The PERSIANN system for rainfall estimation is under development at the Center for Hydrometeorology and Remote Sensing at The University of California, Irvine.

3. APROACH

To generate our base maps from a set of satellites images and increase speed of map rendering we use image pyramid and geo-web cache. Historical of rainfall, discharge, evaporation, water level, inflow data from 1956 to 2004 at daily, monthly and yearly intervals measured by gauges/sensors at hydrological or hydro-meteorological stations are registered as a sensor into PostGIS, in addition to data corresponding to sensors are also inserted as observations into PostGIS. Users can request any type of data at any time periods; their requests are answered by means of maps and statistics charts. We not only used historical rain gauges/environmental sensors but also precipitation data from PERSIANN; however PERSIANN data need to be monitored.

3.1 Display satellite image map and Search

We used various scenes of satellite images for whole study area; the data amount of the mosaic images from 12 Landsat TM and 4 SRTM scenes are quite large data. There is need to reduce the processing time for each time rendering map at different zoom level. In this case we used pyramid image function in Geoserver version 2.1.2 and gdal_retile library to build pyramid image. As most Web mapping clients render WMS data at every time they are queried, they can result in unnecessary and increased waiting times. To deal with this problem, we used *GeoWebCache* to optimize above experience by saving (caching) map images. Currently, to query map features, we used via *GetFeatureInfo* request of WMS in order to build searching functionality for whole 25 ESRI shapefiles. It is quite simple to use WFS *GetFeature*; because WFS GetFeature can accept query of feature within a bounding box and return the query results in GML format (standard encoding developed by OGC). Otherwise, we used Styled Layer Descriptor (SLD) as a filter to search for specific places. SLD is a extension the WMS standards to allow user to ability of controlling the visual portrayal of the geospatial data.

3.2 Sensor Observation Services (SOS)

In recent years, SWE standards have been applied in various projects: Chung et al. [2009] presented standard way to access observations form sensors. M. A. Brovelli et al. [2010] built up Web GeoServices for monitoring the environmental impact of large structures allowing the visualisation of sensor data both with respect to time and space. The system developed supports the implementation of SOS in the standard way developed by M. Cannata [2010].

The SOS Software implementation used for test proposes is istSOS version 1.1 M. Cannata [2010], istSOS is new software for implementing SOS, OGC standard. Our system implements not only *GetCapabilities*, *DescribeSensor*, *GetObservation* but also optional transactional profile. We need to register new sensors and their observation data also; transactional profile provides an access point for data uploading M. Cannata [2010]. The *RegisterSensor* operation allows adding into the SOS system a new sensor by supplying the sensor description. The *InsertObservation* facilitate the sensor in registering new observations for one or more sensors already registered in SOS system and the O&M encode elements containing and presenting the measured values. To get a result from a specific sensor in a period of time such as day, month, year, *GetObservation* request allows obtaining desired observation value.

3.3 Web Processing Service (WPS)

Previous research on WPS such as M.A. Brovelli et al. [2010] used WPS as a powerful tool for their MAGO SOS client and server side. On the other hand, in the work related swapping existing off-line algorimths/tools in open source projects to make them available to process as a WPS, we mention ZOO Project G.Fenoy et al., [2012], it presented possibility to make connection to the GRASS and invoke GRASS command to process data.

A. WPS software implementation

A WPS software implementation used for processing our services is ZOO-Project. ZOO-Project includes ZOO Kernel and ZOO Service Provider. ZOO Kernel is the main core of ZOO Project. It is a server side implemented in C language. Kernel makes possible to create, manage and chain WPS 1.0.0 compliant Web Services by loading dynamic libraries, handling them on-demand. Otherwise it can easily connect to geospatial libraries, scientific models, common cartographic engines and spatial databases. ZOO Service Provider is a couple of Service Shared Objects (SSO) and one metadata ZOO configuration file (.zcfg) per each provider service.

B. Using WPS Interface to design a sequence of Web Services

In [3.2] we saw that using SOS we can insert sensors into PostGIS; the second step is to use those data as input data to create new web services. Otherwise PERSIANN data in the ESRI ASCII raster format are also available to use as the input data. We used WPS for providing group of functionalities such as: *statistical functionality* (B.1) is required to better understand, analyze, and interpret the data. PERSIANN data can be *visualized* (B.2) and added over the Web-GIS. Daily rainfall data measured at the various stations can be *interpolated* (B.3). Finally PERSIANN and gauges data can be *compared* (B.4).

B.1. Statistical functionality

The WPS interface can request geospatial Web Service, such as WFS, WMS, etc., by calling GET operations of these services. Two different implementations can be considered: user through the web interface makes his/her request, the client sends the request to SOS, collects the results and then send them to WPS to be processed. This implementation is easier at server side but it has at least two drawbacks: it needs to invoke SOS service, in this case there is no way to avoid redundant data base in some case e.g.: we need to know only max/min of all data set; it can cause delay in case user request large of data set. To build statistical functionality we implemented the new WPS service. In this service, users need only sending query parameters to the server, service will directly collect data from database and calculate all statistics's factors such as: max, min, average, mean, standard deviation.

B.2 Visualization PERSIANN over to Web-GIS

Daily precipitation data can be downloaded though internet at the web site of PERSIANN product (see at http://chrs.web.uci.edu/persiann/data.html). Name of each file representing each day's precipitation was coded YYDOY.asc YY: year, being DOY the day of year. The idea for making visualization of PERSIANN data in the easiest way is use new enhancement of ZOO Project support for W*S output (see at http://www.zoo-project.org/trac/ticket/34) last update Jul, 2011. This implementation using MapServer integration into ZOO-Kernel to support WMS/WFS/WCS output by very simple way, users need only to modify the zcfg without any changes in SSO.

B.3 Interpolation

As the reason in B.1, WPS can invoke data by connecting to the PostGIS. From 19 rainfall gauges data set, it is possible to "estimate" of rainfall spatial distribution within an area, the method we use to estimate is traditional method named inverse distance. Therefore, we apply GRASS in order to avoid rewriting a same interpolation function instead of it needs only calling GRASS commands. ZOO Service is supported for Python language, Python code can invoke GRASS functionality into ZOO service's code. The specific GRASS command to build interpolation map is v.surf.idw command. The boundary box of interpolation map is within 20.612° and 22°82 N, 102.65° E and 105.67°10E. (Figure 1.B) shows result of the interpolation precipitation in day 25-Jan-2002.

B.4 Comparison

In addition to rain gauges data, the data required for 19 stations can be extracted from PERSIANN data. We wrote a small program to extract the data require for a specific regions, Data, after their extraction, are registered by mean *RegisterSensor* and inserted by using *InsertObservation* request. This implementation is processed analogously to the detail in section 3.2. Our system exists two database sources one from rain gauge data (RG) and the other from PERSIANN data (PN). From two data sources, there are able to create WPS service to analyze correlation between them; the correlations in short time (daily interval) are extremely low. However, correlation of monthly accumulation between gauge and PERSIANN data are almost indicate strong linear relationship (almost correlation coefficient of 19 stations are larger than 0.7), in particular for example in 2002, at Tuan Giao station, the correlation coefficient is about 0.95. (Table 2) illustrates the results of correlation between

RG and PN during three year data from 2000 to 2002 and some of differences between RG and PN.

Station	2000			2001			2002		
	Corr	P (%)	Ave	Corr	P (%)	Ave	Corr	P (%)	Ave
Dien Bien	0.56	64.1	1.7	0.75	75.25	0.96	0.94	72.5	1.2
Nam Muc	0.73	72.3	1.8	0.87	75.6	1.08	0.79	75.8	3.21
Than Uyen	0.91	73.6	1.5	0.90	78.75	1.4	0.93	69.9	2.54
Nam Giang	0.92	74.1	4.1	0.85	76.5	2.98	0.72	76	4.0
Tuan Giao	0.67	67.2	1.9	0.75	76.2	1.56	0.95	71.4	1.8
Hoa Binh	0.59	66.1	-0.6	0.81	64.5	3.67	0.71	71.4	-1.9
Lai Chau	0.81	72	4.1	0.88	78.4	3.5	0.93	75	7.24
Muong Te	0.81	72.8	5.4	0.98	75.5	-2.1	0.92	71.8	4.88
Sin Ho	0.83	71.2	6.5	0.87	75.4	4.9	0.77	73.1	5.2
Quynh Nhai	0.81	68.5	2.2	0.70	75.1	2.09	0.91	72.58	1.31
Son La	0.81	70.6	0.16	0.75	77.25	-0.87	0.82	75.4	-0.48

Table 2: Three years data (2000- 2002) of difference between monthly rain gauge (RG) and PERSIANN data (PN); Corr: Correlation between PN and RG; P: percent of sum of no-rain or rain they are simultaneous between PN and RG divide to number of observations; *Ave*: average difference between RG and PN (in mm).

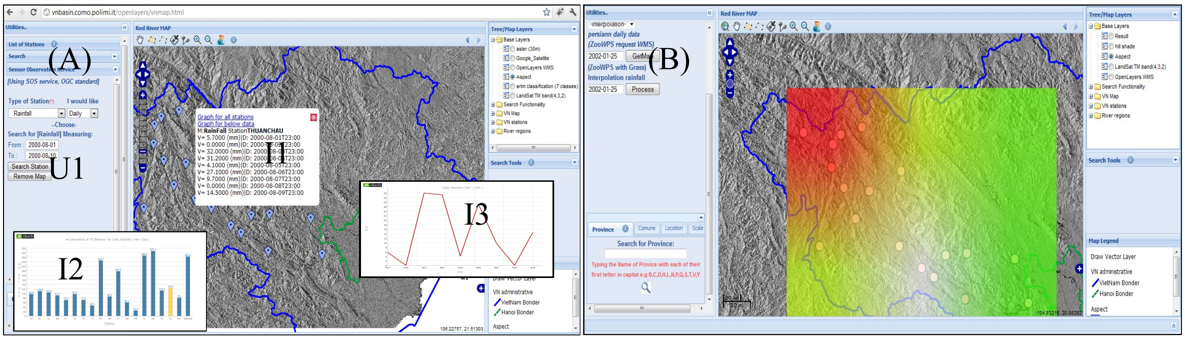

Figure 1: (A) WebGIS with SOS's interface to query data and their results, *I1* is the popup showing value of interval time corresponding to selection from user's interface (*U1*). *I2* is comparison graph for all sensors have a same measure's properties. *I3* is graph for showing different values of observation time at the station when user clicks on. (B) Map of precipitation interpolation using inverse distance method.

4. RESULTS AND CONCLUSION

In this paper we present the intelligent geoportal named HDI4NVN. The aim is not only to improve WebGIS existing but also to extend functionality such as: supporting numerous base maps, search and query, manage and visualize data from gauge and sensor for long periods of time, presenting data in tables and charts. Otherwise WPS plays an important role for building statistical functionality, presenting series of daily rainfall from remote-sensed based for around 10 years over WebGIS, making processing tool to interpolate rainfall/evaporation for around 50 years of data.

Free source remote sensing images available via internet, after some pre-processing are available to build up a base map of the geoportal. Beside, image pyramid and GeoWebCache are powerful tools for increasing the speed of rendering time. Historical data of hydro-meteorological data can be managed by SOS service relying on PostGIS. The use of *GetObservation* request allows clients to obtain the desired observations and from these results of request it is possible to create statistic charts or tables.

Interpolation results still have several drawbacks such as number of points is sparse and study area is large, to overcome this problem, PERSIANN data can be used but it is clear that satellite data need to be examined. We realize that correlation between PERSIANN and gauge data are low in case users request a few days of comparison; on the opposite correlation meets expectation value when we investigate monthly data.

In the future work we would like to select and apply some of the hydraulic models to enrich functionality of the geoportal and also investigate again our PERSIANN data.

ACKNOWLEDGEMENTS

We would like to give special thanks to Prof. Andrea Castelletti and Dr. Xuan Quach Thi for having kindly provided the hydrological measurement.

5. REFERENCES

CHRS, Center for Hydrometeorology & Remote Sensing, University of California, Irvine. http://chrs.web.uci.edu (accessed Jun, 2012).

G. Fenoy, N. Bozon & V. Raghavan, ZOO-Project: *the open WPS platform*. Applied Geomatics, Springer 2012.

Lalonde, W., Geowebcache web site http://geowebcache.org/ (last accessed Jun 2012)., 2002. *Styled Layer Descriptor Implementation Specification*. Version 1.0.0. OGC 02-070.

Lan-Kun Chung, Yao-Min Fang, Yin-Huei Chang, Tien-Yin Chou, Bing Jean Lee and Baranski, B. (2009). *A SOA based debris flow monitoring system - Architecture and proof-of-concept implementation*. Presented at The first symposium on Enterprise Architecture and Information Technology, Taipai, Taiwan.

M. A. Brovelli , G. Giori, M. Mussin, M. Negretti., *Web geoservices for monitoring the environmental impact of large structure*. International Society for Photogrammetry and Remote Sensing (ISPRS), 2010.

M. Cannata, M. Antonovic, ISTSOS: *Investigation of The Sensor Observation Service. WebMGS 2010:* 1st International Workshop on Pervasive Web Mapping, Geoprocessing and Services. 26.-27. August 2010. Como, Italy.

NCHMF, The National Center for Hydro-Meteorological Forecasting, http://www.nchmf.gov.vn/web/en-US/105/92/Default.aspx.