The Design and Development of Spatial Database Management Systems (SDMS) for Hydrographic Studies using coupled Open-Source GIS and Relational Database

A. Boluwade[1] and A Ferdinand

[1]Department Bioresource Engineering, Faculty of Agricultural and Environmental Science, McGill University, Ste Anne de Bellevue, QC, H9X 3V9, Canada.
Email: alaba.boluwade@mail.mcgill.ca

ABSTRACT

In this paper, we seek to design, develop and explore a typical geo-database for hydrographic purposes through the aid of some Open-Source Geographical Information System (GIS) and relational database packages that had been confirmed and certified to be adequate for various GIS tasks. One of the recent and reliable Open-Source software in GIS analysis is gvSIG. gvSIG is a tool oriented to manage geographic information. This GIS software had been integrated with another Open- Source tool database called POSTGIS which is adequate for handling spatial data. Geo-database starts from the design of the data model using the Object-oriented Unified Modified Language (UML) taking into account the primitive data types. The complexities in setting up a geo-database for hydrographic studies are considered. We examined some rivers, lakes and the countries associated with these water bodies. The database will also include countries and its associated regions. Various complex queries can be performed and visualized using these GIS and database tools. Conclusively, this paper demonstrated the potentials, efficiency and optimal benefits involved in using Open-Source tools like gvSIG and POSTGIS for major GIS projects in developing countries like Vietnam.

1. INTRODUCTION

Hydrography focuses on the measurement of physical characteristics of water and marginal land. In the general usage, “hydrography” pertains to measurement and description of any waters. With that usage, oceanography and limnology are subsets of hydrography. In specialized usage, the term applies to those measurements and descriptions of navigable waters necessary for safe navigation of vessel.

In this paper, we are going to design Geo-database on part of the hydrograph for Africa as the term is very broad. Particularly, we are going to deals with some rivers, lakes and the countries associated with these water bodies and some cities of African continent. The database will also include countries and its associated regions.

This Spatial Geo-Database (SGD) would be designed to facilitate storage and access of data through the use of gvSIG and POSTGIS. Data access refers to viewing and downloading data from the SGD. The viewing of data will be aided with gvSIG software. The data are of two kinds, spatial and tabular data and will be maintained in the form of a relational database.

1.2 Brief History of gvSIG and PostGIS

On one hand, gvSIG started in 2002 when the Regional Ministry of Infrastructure and Transport (CIT) of Valencia, Spain proposed the process of changing its computer system to an open source system (Anguiz and Diaz, 2008). The software supports raster and vector data and provides, similar to uDig, several database connectors. Raster functionality is developed based on algorithms of the SAGA GIS project. The programming language for the development of gvSIG is JAVA. The primary goal of gvSIG is to provide software with the functionality of ESRI's ArcView (3.X) (which almost been reached), and, in some aspects, is exceeded by gvSIG. Although the project is under Spanish lead, the user community can be considered international (Steinigher and Boher, 2008). In India, one of the first introductions to gvSIG took place in the WebGIS workshop at IIITMK-Trivandrum (Oct 28—30, 2008), where Prof Dietrich Schroeder of the Stuttgart University of Applied Sciences, Germany used it to demonstrate it spatial analysis capabilities. The audience found its capabilities inspiring and comparable to or even bettering those of ArcGIS. gvSIG has multilingual interface, can be run in Windows, Mac and Linux operating systems, and has implemented the Open Geospatial Consortium (OGC) standards so that gvSIG is compatible with many other GIS software.

On the other hand, PostGIS is an extension to the PostgreSQL object-relational database system which allows GIS objects to be stored in the database. PostGIS includes support for GiST-based R-Tree spatial indexes, and functions for analysis and processing of GIS objects. The GIS objects supported by PostGIS are a superset of the "Simple Features" defined by the OpenGIS Consortium (OGC). As of version 0.9, PostGIS supports all the objects and functions specified in the OGC "Simple Features for Spatial Query Language (SQL)" specification. PostGIS has been developed by Refractions Research as a project in open source spatial database technology. PostGIS is released under the GNU General Public License. They continue to develop PostGIS, and added user interface tools, basic topology support, data validation, coordinate transformation, programming APIs and much more (http://postgis.refractions.net/). PostGIS is an excellent way to bring tabular and spatial data together into a common management environment. The Spatial Query Language (SQL) functions in PostGIS include buffer, intersection, within, distance, and more. These functions take geometric data from columns in PostGIS tables and return new geometries or other information.

2.0 RESEARCH METHODOLOGY

2.1 Area of study

Africa as a continent has been taken as a case study in this paper. In geographical terms, Africa is a complex objects which consists various countries and they are considered as polygons. The integration of various tools will be demonstrated. The procedures described here can also be applicable for any sustainable projects especially for geoinformatic projects in a country like Vietnam.

2.2 Data Base Model

Database modeling is the first step to consider in any geo-database application (see Figure 1). Each of the steps in the work flow follows sequential order in the database development.

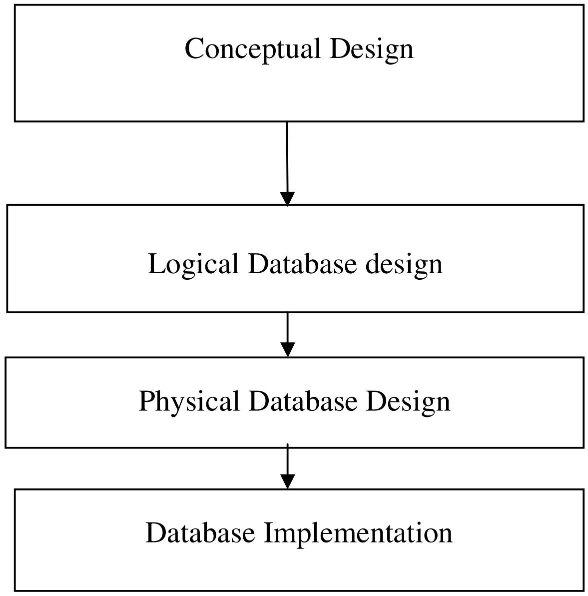

Figure 1: The Database Development Procedures.

2.3 Application of Data Model

There would be need to develop the data model for any particular application. Each of the objects under consideration is considered a "class" and has to be related to other classes with the correct associations and cardinalities. These classes are then translated to be tables which have to be populated with the required attributes. As seen in Figure 2 below; Africa region, countries and cities are related as composition. In practice, these tables and relationships could be developed using a modelling tool called the Unified Modelling Language (UML)

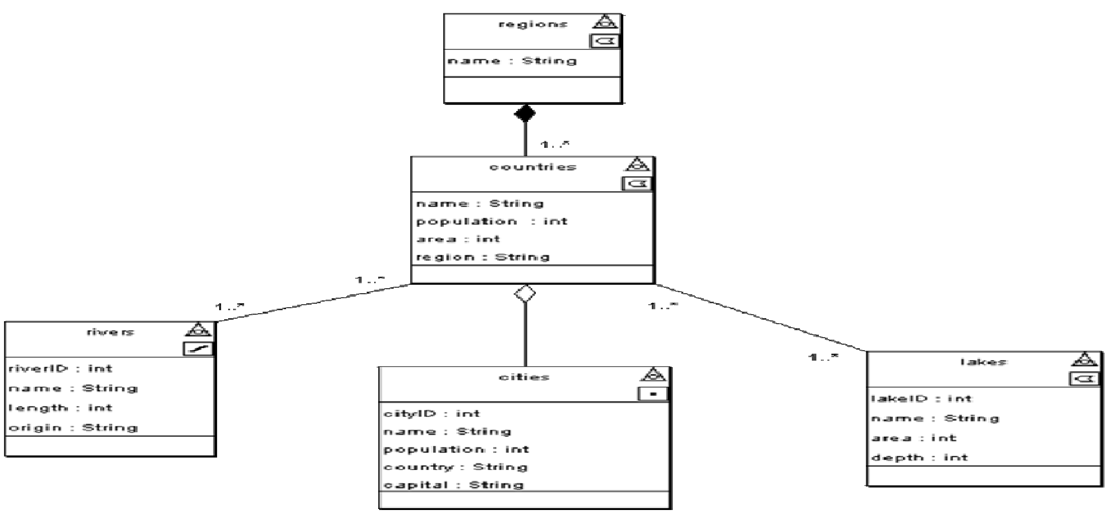

Figure 2: The Data Model of the Hydrographic Application in Africa.

The resulting logical database models, including its SQL implementation have been described below:

2.3.1 Logical Database Model

The following POSTGIS codes give the translation of “Classes” in Figure 2 into “Relational Tables”:

project_regions (name)
project_countries (name, population, project_region.name)
project_cities (cityID, name, population, capital, project_countries.name)
project_rivers(riverID, name, system)
project_rivers_project_countries (project_countries.name, riverID,project_countries, riverID)
project_lakes (lakeID, name, depth)
project_lakes_project_countries (lakeID,project_countries.name,lakeID, project_countries.name)

The underlined green words are primary key while the red are foreign keys.

2.3.2 SQL Implementation

The following shows the SQL processing of the tables in section 2.3.1:

```
/*
CREATE TABLE "project_regions" (
name VARCHAR (40),
PRIMARY KEY(name)
);
*/
--SELECT AddGeometryColumn('project_regions','the_geom',-1,'MULTIPOLYGON',2);

/*
CREATE TABLE  "project_countries" (
name VARCHAR (40),
population INTEGER,
region VARCHAR (40),
CONSTRAINT pk_project_countries PRIMARY KEY(name),
CONSTRAINT fk_project_countries FOREIGN KEY(region) REFERENCES
project_regions (name)
);
*/
--SELECT AddGeometryColumn('project_countries','the_geom',-1,'MULTIPOLYGON',2);

/*
CREATE TABLE  "project_cities" (
cityID SERIAL PRIMARY KEY,
name VARCHAR  (40),
```

```
population INTEGER,
capital CHAR (1),
country VARCHAR (40),
CONSTRAINT fk_project_cities FOREIGN KEY (country) REFERENCES
PROJECT_COUNTRIES (name),
CONSTRAINT capital_values CHECK (capital IN('Y','N'))
);
*/
--SELECT AddGeometryColumn('project_cities','the_geom',-1,'POINT',2);

/*
CREATE TABLE  "project_rivers" (
riverID SERIAL PRIMARY KEY,
name VARCHAR (40),
system VARCHAR (40)
);
*/
--SELECT AddGeometryColumn ('project_rivers','the_geom',-1,'MULTILINESTRING',2);

/*
CREATE TABLE  "project_rivers_project_countries" (
riverID INTEGER,
name VARCHAR (40),
CONSTRAINT pk_project_rivers_project_countries PRIMARY KEY (riverID,name),
CONSTRAINT fk_project_rivers_project_countries1 FOREIGN KEY (riverID)
REFERENCES project_rivers (riverID),
CONSTRAINT fk_project_rivers_project_countries2 FOREIGN KEY (name)
REFERENCES project_countries (name)
);
*/

/*
CREATE TABLE  "project_lakes" (
lakeID SERIAL PRIMARY KEY,
name VARCHAR  (40),
depth INTEGER
);
*/
--SELECT AddGeometryColumn ('project_lakes','the_geom',-1,'MULTIPOLYGON',2);

/*
CREATE TABLE  "project_lakes_project_countries" (
lakeID INTEGER,
name VARCHAR (40),
CONSTRAINT pk_project_lakes_project_countries PRIMARY KEY (lakeID,name),
CONSTRAINT fk_project_lakes_project_countries1 FOREIGN KEY (lakeID)
REFERENCES project_lakes (lakeID),
CONSTRAINT fk_project_lakes_project_countries2 FOREIGN KEY (name)
REFERENCES project_countries (name)
);*/
```

3.0 RESULTS AND DISCUSSION

If all the SQL implementations in Section 2 are successful, then the next step would be to develop queries to test the developed geodatabase.

3.1 Queries Used to Test the Model

The following queries were used to test the model which includes the simplest to the most complex. In essence, the purpose is to query the geodatabase for any specific information that would be of interest.

- ❖ Which countries are Nile river crossing?
- ❖ Which countries are touching Lake Victoria?
- ❖ The five largest lakes based on surface area.
- ❖ Ten capital cities with the largest population
- ❖ Three lakes with the highest depth.
- ❖ Three countries with the highest population densities in the north African region
- ❖ The origin of river Niger?

3.2 Screenshots of the Resulting gvSIG Interface

Few of the queries and the results which will be visualized in gvSIG (as a layer) will be shown in this section.

(i) Which countries are Nile river crossing?

```
/*
CREATE VIEW nile_crossing AS
SELECT pc.name,pc.the_geom
FROM  project_countries pc, project_rivers pr
WHERE ST_Crosses(pc.the_geom,pr.the_geom)
AND
pr.system = 'Nile';
*/
--INSERT INTO geometry_columns VALUES ('','public','nile_crossing','the_geom',2,-
1,'MULTIPOLYGON');
```

Figure 3 shows the visualization of the query.

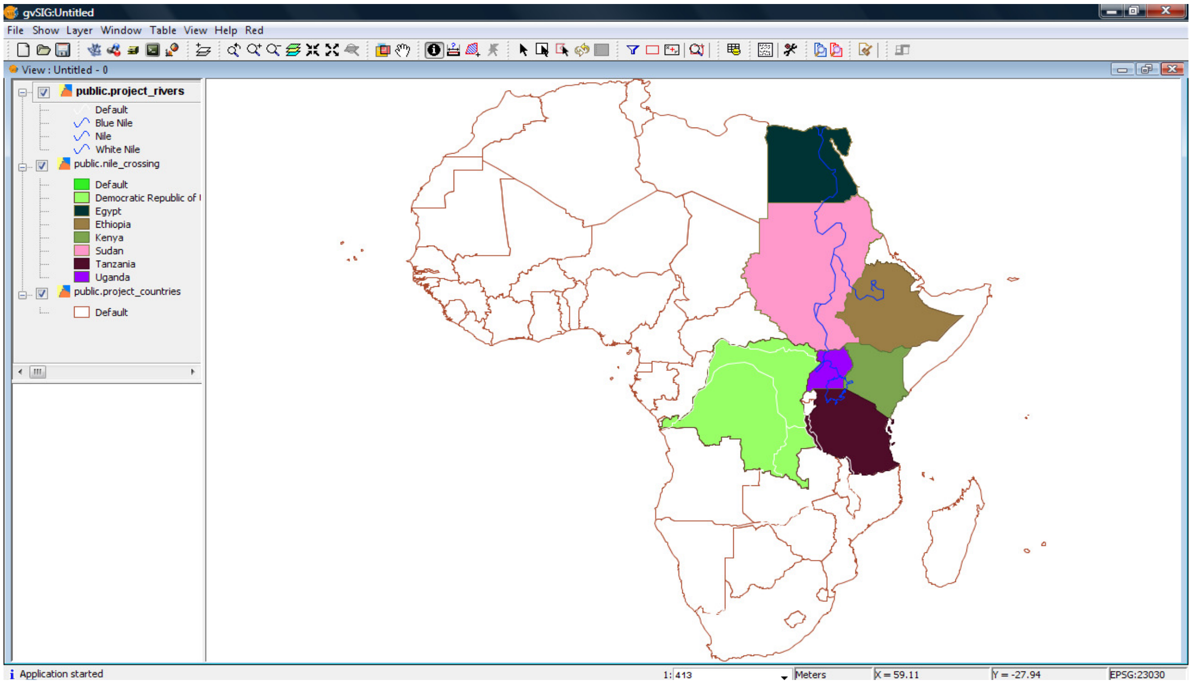

Figure 3: Countries in which river Nile is crossing.

(ii) Which countries are touching Lake Victoria?

```
/*
CREATE VIEW border_victoria AS
SELECT pc.name,pc.the_geom
FROM  project_countries pc, project_lakes pl
WHERE ST_Intersects(pc.the_geom, pl.the_geom)
AND
pl.name = 'Lake Victoria';
*/
--INSERT INTO geometry_columns VALUES ('','public','border_victoria','the_geom',2,-1,'MULTIPOLYGON');
```

In Figure 4, the result of the query can be seen.

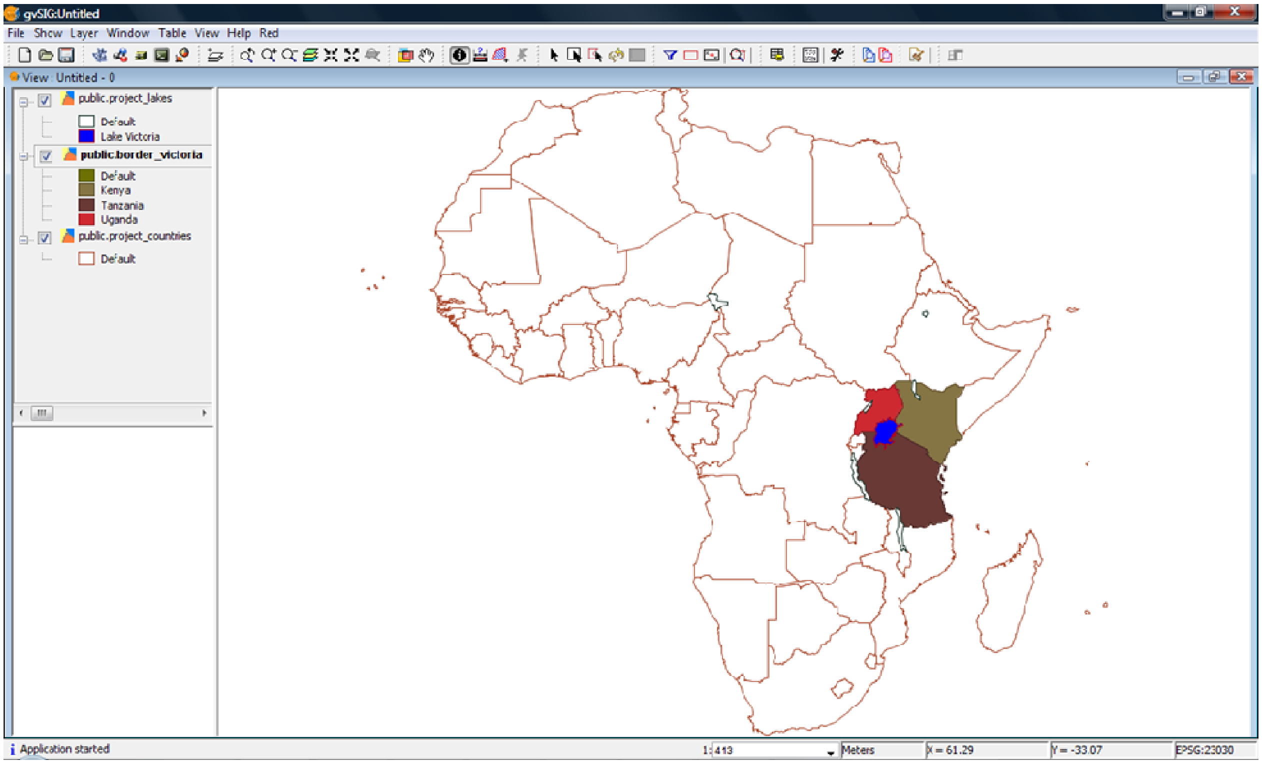

Figure 4: Countries which borders Lake Victoria.

(iii) The five largest lakes based on surface area.

```
/*
CREATE VIEW five_lakes AS
SELECT pl.name, ST_area(pl.the_geom) AS surface_area, pl.the_geom
FROM project_lakes pl
ORDER BY 2 DESC
LIMIT 5;
*/
--INSERT INTO geometry_columns VALUES ('','public','five_lakes','the_geom',2,-
1,'MULTIPOLYGON');
```

The result is shown in Figure 5.

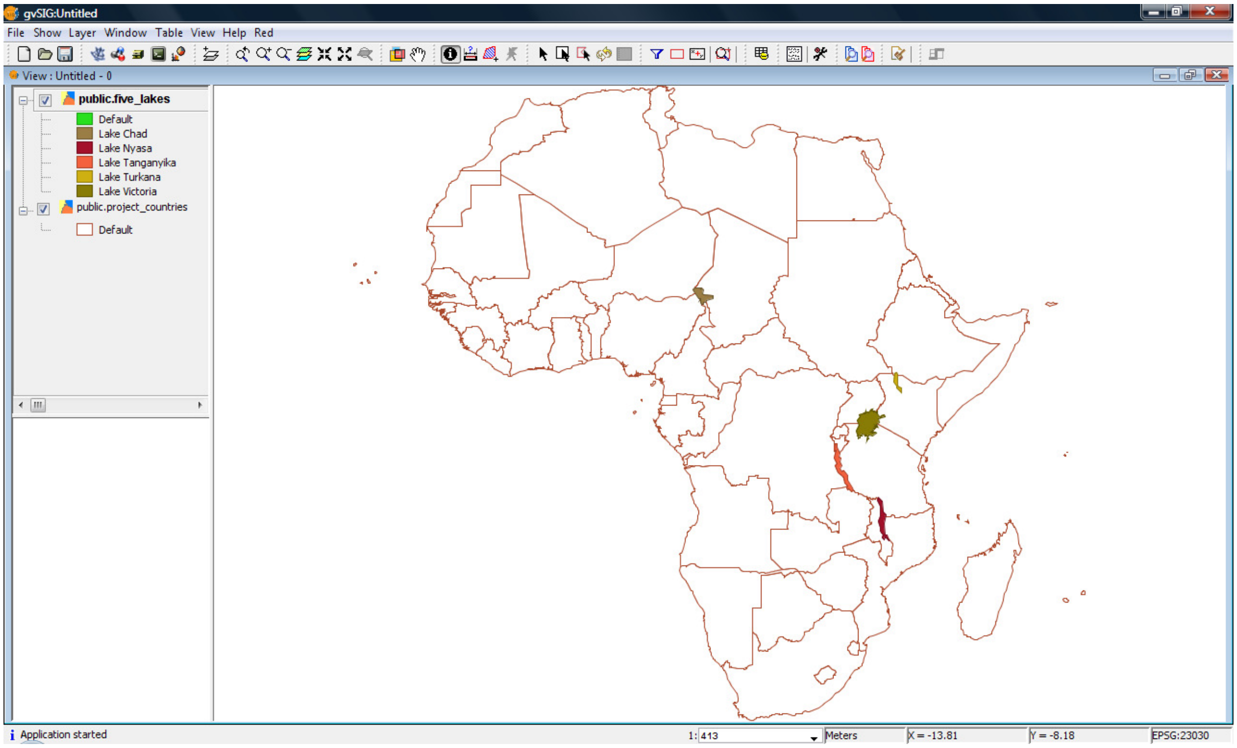

Figure 5: Five largest lakes based on surface area.

4.0 CONCLUSION AND RECOMMENDATION

In conclusion, the design and development of geo-database for Africa had been done. In this paper, valuable information for various hydrographic activities in Africa has been provided. This would be an asset for African and Non-Africa researchers. Developing countries' researchers and GIS experts like Vietnam can also make use of these efficient and sustainable open source GIS tools for various geoinformatic projects without having to acquire commercial GIS software. The various stages of database development had been followed sequentially from logical to physical stages. As a recommendation, there would still be need to explore some other minor and major rivers and lakes for a more holistic hydrographic study.

5. REFERENCES

Anguix A. and Diaz L. 2008. *Review. gvSIG: A GIS desktop solution for an open SDI*. Journal of Geography and Regional Planning Vol. 1(3), pp. 041-048.

Bas V, Ulrike R, Adam R and M. Loudon. 2008. *Open source GIS for HIV/AIDS management.* International Journal of Health Geographic 7:53

Dravis, P. 2002 .*A Survey on Open Source Software.* The Dravis Group, San Francisco.

Ghosh, R. A., B. Krieger, R. Glott, and G. Robles. 2002. *Open Source Software in the Public Sector: Policy within the European Union*, International Institute of Infonomics, University of Maastricht, Maastricht.

Steinigher S. and Bocher E. 2008. Review Article. *An Overview on Current Free and Open Source Desktop GIS Developments.* Int. Journal of Geographical Information Science, 1. Revision, Sept. 5th 2008.