Preliminary Data Compilation for the Nile Basin Decision Support System:

INCEPTION REPORT

Report FutureWater: 85

Commissioned by
Nile Basin Initiative

Authors
P. Droogers
W.W. Immerzeel

Preliminary Data Compilation for the Nile Basin Decision Support System

INCEPTION REPORT

February 2009

FutureWater Report 87

Commissioned by
Nile Basin Initiative

Authors
P. Droogers
W.W. Immerzeel

FutureWater
Costerweg 1G
6702 AA Wageningen
The Netherlands

+31 (0)317 460050

info@futurewater.nl

www.futurewater.nl

Preface

The Nile Basin Decision Support System will provide the necessary knowledge base and analytical tools to support the planning of cooperative joint projects and the management of the shared Nile Basin water resources on an equitable, efficient and sustainable manner. The Water Resources Planning and Management Project (WRPM) is one of the eight projects under the Shared Vision Program of NBI. The WRPM project is preparing the development of the Nile Basin Decision Support System (NB-DSS) in 3 Work Packages:

1. WP1: Software Development and Implementation
2. WP2: Data Compilation and Pilot Test Applications
3. WP3: Supervision and Monitoring

The bulk of the data compilation activities shall be carried out under Work Package 2. However, the project envisages the need to conduct preliminary data collection and compilation. This data collection and compilation is the basis for the subsequent, more extensive, work on data compilation under WP2. A consultancy is undertaken focusing on this preliminary data collection and compilation under the title "Support Data Compilation for the Development of the Nile Basin DSS". The consultancy has been divided into three phases:

1. Inception phase
2. Analysis phase
3. Synthesis phase

This report describes the results of the first phase of the project (Inception phase) and is a result of the following activities:

- Data requirements
- Data availability
- Gap identification
- DBMS options
- Projection options

The consultants wish to acknowledge the support, fruitful discussions and useful comments from NBI staff and in particular: Dr. Abdulkarim Seid, Mr. Ephrem Getahun, Dr. Solomon Demissie, Dr. Mekuria Beyene, Dr. Elnaser Abdelwahab.

Table of contents

1 Introduction

1.1 The Nile Basin

The Nile, which is home and source of livelihood to approximately 160 million people, is the longest river in the world having a total length of about 6700 km, traversing an extremely wide band of latitude, from 4-degree south to 32-degree north. The area draining into the Nile river system, about 3 million km^2 extends over 10 African countries.

The two main river systems that feed the Nile are the White Nile, with its sources on the Equatorial Lake Plateau (Burundi, Rwanda, Tanzania, Kenya, Democratic Republic of the Congo and Uganda) and fed by substantial flow from the Baro-Akobo-Sobat system that originates in the foothills of southwest Ethiopia, and the Blue Nile, with its sources in the Ethiopian highlands. Tekeze-Setit-Atbara system contributes to the flow further downstream of Khartoum.

Annual runoff potential of the Nile Basin is estimated to be approximately 85 billion cubic meters (BCM). Such estimates of average annual runoff may vary depending upon the length of records used for the estimation. Compared to other major river basins, the Nile Basin's disparity in water availability differs sharply among sub-basins. Arid portions (perhaps one-third of the area of the Basin) yield negligible flows; whereas, the Highland of Ethiopia, comprising perhaps as little as 15–20 percent of the land area of the overall Basin, yields 60–80 percent of the annual flow in the lower Nile.

1.2 The Nile Basin Initiative

The NBI is a partnership initiated and led by the riparian states of the Nile River through the Council of Ministers of Water Affairs of the Nile Basin states (Nile Council of Ministers, or Nile-COM). The NBI seeks to develop the river in a cooperative manner, share substantial socioeconomic benefits, and promote regional peace and security. The NBI started with a participatory process of dialogue among the riparian countries that resulted in their agreeing on a shared vision: to "*achieve sustainable socioeconomic development through the equitable utilization of, and benefit from, the common Nile Basin water resources*".

The NBI involves complex multi-country projects under the two programs, namely, the Shared Vision Program, which deals with projects aimed at creating enabling environment and capacity development for advancing cooperation among riparians, and the Subsidiary Action Programs (SAP), the Eastern Nile and the Nile Equatorial Lakes – Subsidiary Action Programs (EN – and NELSAP), which deal with cooperative investment oriented projects. Considerable progress has been made in the implementation of the various projects under the two NBI programs.

1.3 The Water Resources Planning and Management Project

The water resources planning and management (WRPM) project is one of the eight projects under the Shared Vision Program of NBI. The primary objective of this project is to enhance analytical capacity for a basin-wide perspective to support the development, management, and protection of Nile Basin water resources in an equitable, optimal, integrated, and sustainable manner. The project has the following three technical components:

1. Water Policy Good Practice Guides and Support: The output of this component is to strengthen capacity to formulate and implement effective national policies and strategies for integrated water resources management (IWRM) in Nile Basin countries. This will be a country- and needs-driven component aimed at enabling all basin countries to operate on an equal footing.
2. Project Planning and Management Good Practice Guides and Support: The output of this component is the enhanced capacity in Nile Basin countries for planning and managing multi-country projects, thus contributing to an improved IWRM in the region. These skills will become particularly important as NBI cooperation grows and cooperative investment projects are developed through the subsidiary action programs.
3. Nile Basin Decision Support System: The output of this component is an operational Nile Basin DSS supported by trained staff. The Nile Basin DSS will provide a common, basin-wide platform for communication, information management, and analysis of Nile Basin water resources. Coupled with human resources development and institutional strengthening, the Nile Basin DSS will provide a framework for sharing knowledge, understanding river system behavior, evaluating alternative development and management schemes, and supporting informed decision-making from a regional perspective, thus contributing to sustainable water resources planning and management in the basin.

The WRPM project is managed from the Project Management Unit (PMU) hosted by the Government of Ethiopia and located at Addis Ababa.

1.4 The Nile Basin Decision Support System

The Nile Basin DSS, which is a component of the Water Resources Planning and Management Project, is expected to provide the necessary knowledge base and analytical tools to support the planning of cooperative joint projects and the management of the shared Nile Basin water resources on an equitable, efficient and sustainable manner.

The primary objective of the Nile Basin DSS is to develop a shared knowledge base, analytical capacity, and supporting stakeholder interaction, for cooperative planning and management decision making for the Nile River Basin. An essential feature of the Nile Basin DSS should be that it is an agreed upon tool that will be accepted and used by all riparians in the management of the shared Nile water resources.

To support the development and continued use of the Nile Basin DSS, a Nile Basin Regional Decision Support System Center (DSS Center) is established at the Project Management Unit (PMU), in Addis Ababa, Ethiopia. The Regional DSS Center is responsible for developing and operational use of the Nile Basin Decision Support System. The regional DSS Center is supported by national DSS units in every NBI member country.

A recently finished consultancy described the design and requirements for the DSS based on a rigorous process of stakeholder consultations, analysis of available models and technology and an assessment of different areas of concern related to water resources in the riparian countries (Fedra, 2008). The DSS is designed to contain three major components:

- An information management system that provides a common and shared information basis for the planning and decision making processes, locally, sub-regionally, and basin wide, directly accessible for all stakeholders;
- A modular river basin modeling and economic evaluation system built around a dynamic water budget and allocation model, that helps to design and evaluate possible interventions, strategies and projects in response to the problems and challenges identified and prioritized in the stakeholder consultations;
- Tools for a participatory multi-criteria analysis to rank and select alternative compromise solutions for win-win strategies.

In the initial first phase of the DSS development are designed to address a basic set of main concerns and priority issues comprising efficient water resources management and allocation, water quality, extreme events (floods and droughts), agriculture, hydropower, and navigation as well as watershed management and erosion, considering simultaneously hydrological, environmental and socio-economic criteria and objectives. This shall be extended to include other areas of concern in subsequent phases.

The DSS shall be implemented at a central location (NBI PMU), two-sub-regional locations (covering the Eastern Nile and the Nile Equatorial Lakes sub-regions) and at the country levels. The data structure and DSS components are identical at all scales, however the contents may vary between the central, sub-regional and national installations of the DSS. Clearly this requires a well designed generic database structure that can deal with a variety of thematic fields, across different spatial and temporal scales and feeding different types of models.

The development of the DSS is organized in three different work packages:

- Work Package 1 (WP1): Software Development and Implementation
- Work Package 2 (WP2): Data Compilation and Pilot Test Applications
- Work Package 3 (WP3): Supervision and Monitoring coordinates/synchronizes these parallel activities and organizes quality assurance processes

The work packaging of the development of the Nile Basin DSS mainly aims at obtaining a high quality product (software + data) in a comparatively short period and delivering the first release of the Nile Basin DSS with the core functionality at an early stage. In order to meet these requirements all relevant data need to be compiled prior to the DSS development insofar as to expedite data compilation activities in WP2 and avoid problems that could arise in data compilation during execution of WP2. The bulk of the data compilation activities shall be carried out under Work Package 2. However, the project envisages the need to conduct preliminary data collection and compilation.

This data collection and compilation is the basis for the subsequent, more extensive, work on data compilation under WP2 and has been divided into three phases:

1. Inception phase
2. Analysis phase

3. Synthesis phase

This report describes the results of the first phase, Inception, as was carried out during December 2008 and January 2009.

2 Data Requirements

2.1 Identification of required data types

For the core models as identified in Fedra (2008) data can for a large part be derived and extracted from public domain data sources. In this consultancy the focus is primarily on a preliminary analysis of required data types to sustain the development of the core models at the level of the entire Nile Basin.

This set of core models will address the following areas of concern, with water quality and climate change being cross-cutting issues:

1. Water resources development
2. Optimal water resource utilization
3. Coping with floods
4. Coping with droughts
5. Energy development
6. Rainfed and irrigated agriculture
7. Watershed and sediment management
8. Navigation

The range of requirements for such a tool is complex. As a result, to be realistic, the DSS development shall evolve gradually over time in an adaptive, phased approach with initial efforts focusing on the priority concerns. Time and budgetary limitations also require that the DSS has to be implemented in an adaptive, step-wise fashion. This suggests the need for an immediate target, defining what is a reasonable expectation of the first phase; defining in other words what is achievable while responding nonetheless to clear priority needs of the riparians. It was, therefore, decided that the immediate target of the first DSS development exercise shall be "A water balance and allocation model, linked to a set of core models relevant to the priority areas of concern in the Nile Basin, and integrated with an information system and decision support tools for multi-criteria analysis (MCA)".

A range of different models from meteorological, rainfall-runoff, hydrological, hydraulic to water quality models will address these areas of concerns. However, despite the model type, in modeling the following stages can be distinguished each with its own data requirements.

<u>Model schematization and parameterization</u>
The first step in building a model is the spatial stratification of the model domain. This generally includes a sub basin delineation based on topographic features, generation of hydrological response units from combinations of land use and soil and stream network and reservoir definition. Secondly these units of calculation need to be parameterized with for example hydraulic and hydrologic properties, land surface characteristics, reservoir storage volumes.

<u>Model forcing</u>
Once the model is parameterized it is forced with meteorological data. These can be either observed data from meteorological stations, modeled data from weather models, remote sensing data or a combination of these as well as stochastically generated sets of time series.

<u>Model calibration and validation</u>

The calibration process includes modification of model parameters in such a way that model outputs best match observations of for example stream flow or water quality. Validation refers to process of checking whether the calibrated model mimics reality well in a time slice outside the calibration period.

Scenario analysis

The final step is the actual analysis using the calibrated models. This may entail climate change scenarios, management scenarios or land use change scenarios depending on the area of concern that is addressed.

Each of these essential stages has it own data requirements and a detailed overview is given in Figure 1. This Figure is based on experiences with DSS development of other river basin authorities such as

- Mekong River Commission (Droogers, 2008)
- Hai Basin Commission (http://www.iwlearn.net/iw-projects/Fsp_112799469189)
- Krishna Basin (http://cwc.nic.in/regional/hyderabad/welcome.html)
- Rio Bravo Basin (http://www.ibwc.state.gov/) and
- Yellow River Conservancy Commission (http://www.yellowriver.gov.cn).

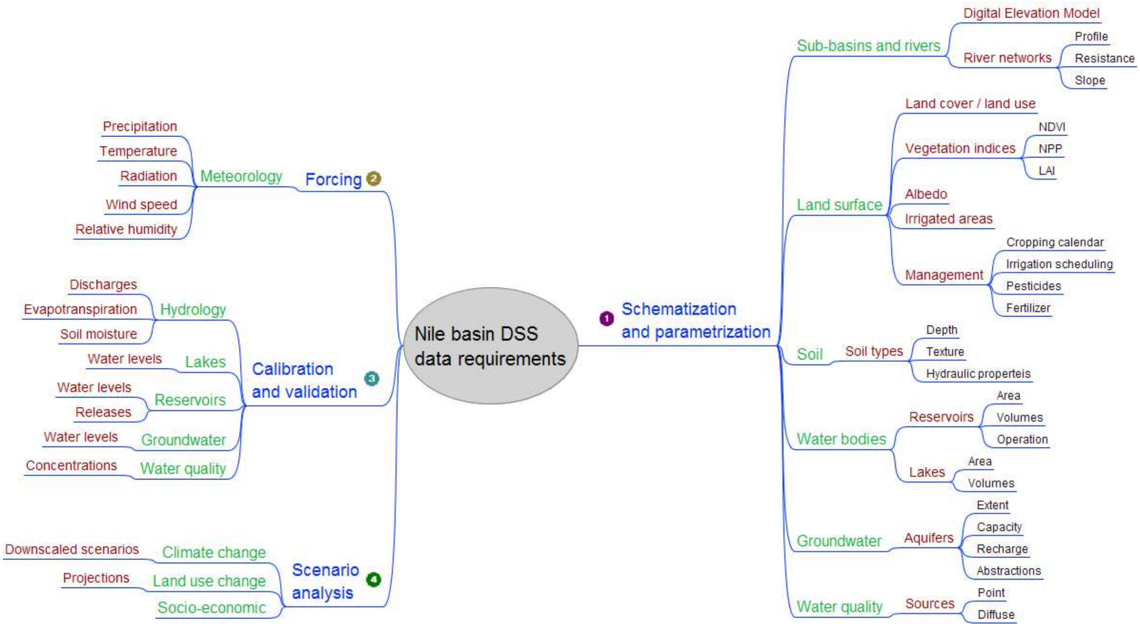

Figure 1 Data requirements for the Nile Basin DSS

3 Datasets available

3.1 Local sources

During the inception phase of the project local data were obtained from only the Ethiopian part of the Nile Basin (Figure 2). A total of 383 stations are present in the local database (Table 1). Other countries will be included during the analysis phase.

Table 1. Number of stations with climatic and hydrological information.

SubBasin	Climate	Hydrologic	Total
Abbay	173	28	201
Bara Akobo	42	41	83
Blue Nile	0	9	9
Tekeze	57	32	89
Total	273	110	383

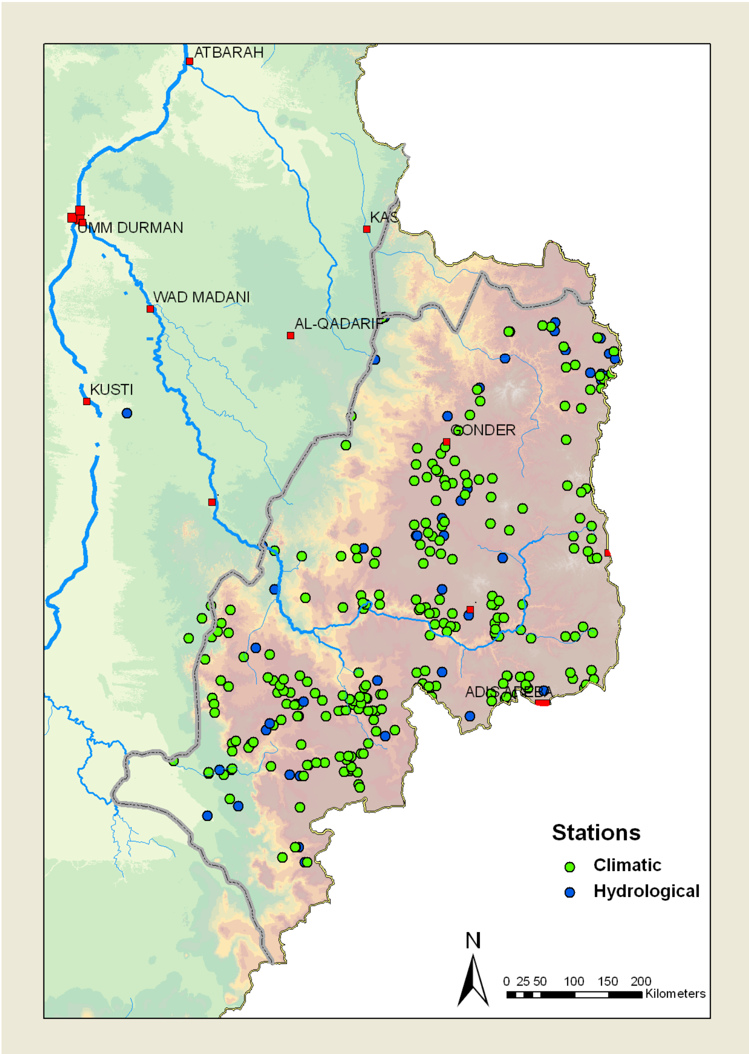

Figure 2. Local data obtained during the inception phase.

3.2 Public domain sources

3.2.1 Climate Data

3.2.1.1 CRU

The best global source of meteorological data at a relatively high spatial and temporal resolution is available from the Climatic Research Unit of the University of East-Anglia. The CRU TS 2.1 dataset comprises 1224 monthly grids, for the period 1901-2002, and covers the global land surface at 0.5°× 0.5°resolution (Mitchell *et al*. 2003). The dataset comprises: cloud cover, diurnal temperature range, precipitation, temperature and vapor pressure (note that no wind speed data are available in the CRU dataset).

The dataset is based on raw station data. Since data can be scarce in some regions or periods, a method called 'relaxation to the climatology' was used to create continuous grids. This implies that, for some areas or regions, data are less accurate.

More information on the CRU gridded datasets can be found at http://www.cru.uea.ac.uk/cru/data/hrg/.

3.2.1.2 TRMM

Rainfall can be obtained from TRMM (Tropical Rainfall Measuring Mission). TRMM is a joint space project between NASA and the Japanese Aerospace Exploration Agency (JAXA). TRMM is designed to measure tropical precipitation. The TRMM satellite has a latitudinal range from 50^oS – 50^oN. The TRMM satellite has been launched November 27, 1997 and the mission has recently been extended to 2009.

The sensors to measure rainfall consist of precipitation radar, a multi-frequency microwave radiometer and a visible and infrared (VIS/IR) radiometer. These sensors are used in a complementary way to deduct rainfall. Several orbital and gridded data products are available for download at the Goddard Distributed Active Archive Centre (DAAC). The 3B-43 product is considered as the most applicable for hydrological studies. The output is rainfall for 0.25x0.25 degree (~ 25x25 km) grid boxes for each month.

Further information can be obtained at:
http://disc.sci.gsfc.nasa.gov/data/datapool/TRMM/01_Data_Products/index.html and
Tropical Rainfall Measuring Mission (TRMM): http://trmm.gsfc.nasa.gov/

3.2.1.3 Data Archives

There are a couple of data archives available that store and distribute local data. In principle, these archives store only data that have been obtained from national weather services and one would expect that those data is already possessed by the national NBI offices. However, if the national NBI offices fail to gather these data, one could try the public domain data archives. The most relevant are:

- GSOD database NOAA/WMO
  - http://www.ncdc.noaa.gov/cgi-bin/res40.pl?page=gsod.html
- Weather Underground

- o http://www.wunderground.com

3.2.1.4 Recommendation

It should be emphasized here that climate data and especially rainfall is of paramount importance for the DSS. It is recommended to follow a three level approach in information. The most reliable information is data from local stations. If these data are not available (or are not resealed by national weather agencies) for a certain area or period one could go for TRMM as far as precipitation concerns. If this also fails the CRU data set can be used.

3.2.2 Streamflow

3.2.2.1 GRDC

The Global Runoff Data Center holds the Global River Discharge Database. GRDC's role is to serve as a facilitator between data providers and data users. It serves under the auspices of the World Meteorological Organization (WMO) and has been established at the Federal Institute of Hydrology (BfG), Germany. For the Nile Basin data from 77 stations are available. However, most data is only available at a monthly base and most recent data is from 1990. For 8 stations, all in Ethiopia, data are available at daily time scale for the period 1978-1980.
http://grdc.bafg.de

3.2.2.2 UNH

The University of New Hampshire developed a global runoff database. The database is developed using a global water balance model and fine-tuned using observations. These observations can be downloaded, but includes only monthly averages from historic data and are identical as the ones from the GRDC (http://www.grdc.sr.unh.edu).

3.2.2.3 River Discharge Database

SAGA is the Center for Sustainability and the Global Environment of the University of Wisconsin-Madison. SAGA holds a database with global river discharges. Data from 24 stations[1] along the Nile are available. Data is a identical as the GRDC but with a few less stations. In contrast to the GRDC data can be downloaded directly from their website.
http://www.sage.wisc.edu/riverdata/

3.2.2.4 RivDis database

The Global River Discharge Database development efforts represent the first step in a continually evolving compilation of river discharge information. One of the primary sources of information for the database development was the UNESCO river archives and the series of publications entitled "The Discharge of Selected Rivers of the World" which were provided, in book form from 1969 through 1984. The series served as an important source of information on

[1] Searching the database on the word “Nile” provides also two stations in the USA, Kansas.

approximately 1000 stations. RivDis v1.0 provides discharge data from the original UNESCO publication series in a digital format that can be easily acquired and analyzed by researchers and planners in the water sciences community. For Africa data of 269 stations are available of which 27 are located inside the Nile Basin. More details can be obtained from: http://www.rivdis.sr.unh.edu/

3.2.2.5 LEGOS

The GOHS (Géophysique, Océanographie et Hydrologie Spatiales) group of LEGOS (Laboratoire d'Etudes en Géophysique et Océanographie Spatiales) in Toulouse collects and distributes river, lake and reservoir levels based on satellite observations. The water level time series are based on altimetry measurements from Topex/Poseidon, Jason-1, ERS-2, ENVISAT and GFO satellites. The database includes water levels for over 130 lakes and man-made reservoirs, 250 virtual stations on rivers and about 100 sites on flooded areas.

Most river water level time series are based on Topex/Poseidon observations and start in January 1993. The lake water level time series are multi-satellite data combinations. They also start in January 1993. The time series are regularly updated and the number of sites increases regularly.

During the analysis phase of the project a complete inventory of available data for the Nile Basin will be made and relevant data will be included in the database.

3.2.3 Digital Elevation Models

3.2.3.1 SRTM

The most commonly public domain DEM dataset is the SRTM dataset which is acquired with the space shuttle during an 11-day mission in 2002. More information on the SRTM can be found at http://www2.jpl.nasa.gov/srtm/. The dataset can be downloaded from various sources. Probably the most user-friendly one is provided by the CGIAR (version 4) and has been obtained in tiles of 5 x 5 degrees for the entire Nile basin. This dataset has been slightly modified from the original version provided by the USGS (seamless.usgs.gov) and data gaps are filled using an automated procedure. For more information reference is made to http://srtm.csi.cgiar.org/.

3.2.3.2 ASTER

A new global data based on data acquired with the ASTER satellite with an improved resolution of 30 m is expected to be released in spring 2009. A summary of the characteristics as well as a link to the data sources can be found in Table 2.

Table 2 Public domain DEM sources

	ASTER G-DEM	**SRTM3**
Data source	ASTER	Space shuttle
Generation and distribution	METI/NASA	NASA/USGS
Release year	2009 (planned)	2003
Data acquisition period	2000 - to date	2000

DEM resolution	30 m	90 m
DEM accuracy	~ 7 m	~10 m
DEM coverage	83°N - 83°S	60°N - 56°S
Area of missing data	Areas with permanent cloud cover	Topographically steep terrain
Link to data	http://www.ersdac.or.jp/GDEM/E/index.html	http://seamless.usgs.gov http://srtm.csi.cgiar.org/

3.2.4 *Soils*

3.2.4.1 FAO/UNESCO

The most commonly used global dataset on soil is the FAO/UNESCO Soil Map of the World at scale 1:5,000,000 (FAO, 1974). Based on this map the Digital Soil Map of the World (DSMW) was published. One of the drawbacks of this dataset is that the information is qualitative rather than quantitative. However, procedures have been developed to transform the original classes into information needed by models (Droogers et al., 2001).

3.2.4.2 SOTER

SOTER's objective is to characterize all soils from all corners of the world under a single set of classification rules. As there was no universally accepted system for world-wide classification of terrain, SOTER has designed its own system (Van Engelen and Wen, 1995). However, regarding the Nile Basin Countries only for Kenya a detailed soil map is available. SOTER has also published a CD-ROM with "Soil and Terrain Database for northeastern Africa", but this dataset has the same soils information as the FAO/UNSECO global dataset.

3.2.4.3 Harmonized World Soil Database

In 2008 a new global dataset was developed under the name "Harmonized World Soil Database" existing out of a 30 arc-second raster database with over 15000 different soil mapping units that combines existing regional and national updates of soil information worldwide (SOTER, ESD, Soil Map of China, WISE) with the information contained within the 1:5 000 000 scale FAO-UNESCO Soil Map of the World (FAO, 1974).

The resulting global raster database consists of 21600 rows and 43200 columns, which are linked to harmonized soil property data. The use of a standardized structure allows for the linkage of the attribute data with the raster map to display or query the composition in terms of soil units and the characterization of selected soil parameters (organic Carbon, pH, water storage capacity, soil depth, cation exchange capacity of the soil and the clay fraction, total exchangeable nutrients, lime and gypsum contents, sodium exchange percentage, salinity, textural class and granulometry).

3.2.4.4 IUSS

The International Union of Soil Sciences is starting a new initiative to develop an updated digital soil map, with special emphasis on the need of the modeling community (http://www.globalsoilmap.net/). It will take however years before the first results will emerge.

3.2.4.5 Recommendation

Based on the previous sections the Harmonized World Soil Database can be considered as the most appropriate one to be used for the DSS.

3.2.5 Land cover / land use

3.2.5.1 USGS

The USGS (United States Geological Survey) has been active in developing and distributing land cover information. The Africa Land Cover Characteristics Data Base Version 2 is a subset of a global land cover characteristics data base that was developed on a continent-by-continent basis and are based on 1-km AVHRR data spanning April 1992 through March 1993 (Loveland and others, 1999). There are seven different legends available depending on the need of the user. A drawback of this dataset is that it is based on information from over 15 years ago.

3.2.5.2 GLCF

The Global Land Cover Facility (GLCF) is a center associated to the University of Maryland. GLCF provides earth science data and products to help everyone to better understand global environmental systems. In particular, the GLCF develops and distributes remotely sensed satellite data and products that explain land cover from the local to global scales. The UMD Land Cover Classification global land cover was based on imagery from the AVHRR satellites acquired between 1981 and 1994. The database includes fourteen land cover classes and is available at three spatial scales: 1 degree, 8 kilometer and 1 kilometer pixel resolutions. A drawback of this dataset is that it is based on information from over 15 years ago.

3.2.5.3 GLC2000

The Land Cover map of Africa is one regional component of the GLC2000 exercise, conceived and coordinated by the European Commission's Joint Research Centre. The GLC2000 maps are based on daily observations made from 1st November 1999 to 31st December 2000 by the VEGETATION sensor on the SPOT 4 satellite. The Africa map's legend pays special attention to the forest and savannah biomass. The map shows specific land-cover features as the irrigated agriculture, the ribbons of secondary forest of the swamp forests at a spatial detail never achieved before. The current version is 05 (Mayaux et al., 2004).

GLC2000 has a legend with 27 classes and can be considered as the best land cover dataset currently available for the entire continent. The spatial resolution is 1 x 1 km.

3.2.5.4 AfriCover

The purpose of the Africover Project is to establish a digital georeferenced database on land cover and a geographic referential for the whole of Africa. Currently data from the following ten countries are available:

- Burundi
- DR Congo
- Egypt

- Eritrea
- Kenya
- Rwanda
- Somalia
- Sudan
- Tanzania
- Uganda

This dataset is probably the best available land cover data set, but a major drawback is that not all Nile Basin countries are included.

3.2.5.5 Global Irrigated Area Map

The Land and Water Division of the Food and Agriculture Organization of the United Nations and the Johann Wolfgang Goethe Universität, Frankfurt am Main are co-operating in the development of a global irrigation mapping facility. The first global digital map of irrigated areas on the basis of cartographic information and FAO statistics has a resolution of 0.5 degree and was developed in 1999. Since 1999 the methodology to produce the map has been improved which made it possible to increase the spatial resolution of the map to 5 minutes (about 10 km at the equator). The objective of the co-operation between the Johann Wolfgang Goethe Universität and FAO is to develop global GIS coverage of areas equipped for irrigation and to make it available to users in the international community. The data collected through the AQUASTAT surveys was used to improve the overall quality and resolution of the information. (Siebert et al, 2006).

3.2.5.6 Recommendation

It is recommended to use the GLC2000 dataset as this is the most recent one and constructed using a uniform approach for the entire Nile Basin. This map can be overlayed with the Global Map of Irrigated Areas to ensure that irrigation will be included explicitly.

3.2.6 Vegetation Indices

3.2.6.1 MODIS

MODIS (or Moderate Resolution Imaging Spectroradiometer) is a sensor aboard the Terra (EOS AM) and Aqua (EOS PM) satellites. Terra's orbit around the Earth is timed so that it passes from north to south across the equator in the morning, while Aqua passes south to north over the equator in the afternoon. Terra MODIS and Aqua MODIS are viewing the entire Earth's surface every 1 to 2 days, acquiring data in 36 spectral bands.

Various vegetation characteristics are derived from MODIS and are available to be included in the NBI-DSS. The most relevant products are:

- MOD 13 - Gridded Vegetation Indices (Max NDVI & Integrated MVI)
- MOD 15 - Leaf Area Index & FPAR
- MOD 17 - Net Photosynthesis and Primary Productivity

Most of these products are available at spatial resolution of 250, 500 and 1000 m and at temporal resolutions of 16 days and one month. (http://modis.gsfc.nasa.gov)

Various sites are available for downloading data. The primary site to obtain MODIS data is LP DAAC (Land Processes Distributed Active Archive Center) at https://lpdaac.usgs.gov/. Data can be obtained using three tools:

- Data Pool: The Data Pool is the publicly available portion of the LP DAAC online holdings. Data Pool provides a more direct way to access files by foregoing their retrieval from the nearline tape storage devices
- WIST: The Warehouse Inventory Search Tool (WIST) is a web-based client to search and order earth science data from various NASA and affiliated centers
- GloVis: The USGS Global Visualization Viewer (GloVis) is an online search and order tool for selected satellite data

3.2.6.2 Landsat

The Landsat program is the longest running enterprise for acquisition of imagery of Earth from space. The first Landsat satellite was launched in 1972; the most recent one, Landsat 7, was launched on April 15, 1999. Landsat 7 data has eight spectral bands with spatial resolutions ranging from 15 to 60 meters. Since May 2003 Landsat 7 has a scan line failure, but images are still useable.

Original Landsat images can be downloaded from various sites. For the purpose of the NBI-DSS it is recommended to use the so-called GeoCover2000 images. Specific detailed studies might benefit from the original high-resolution Landsat images.

The GeoCover2000 images are distributed by various sites, but most recommended is the Earth Science Data Interface of the Global Land Cover Facility (http://glcfapp.umiacs.umd.edu). Alternatively data can be obtained from the NASA site: https://zulu.ssc.nasa.gov/mrsid/mrsid.pl.

3.2.6.3 Spot

SPOT (Satellite Pour l'Observation de la Terre) is a high-resolution, optical imaging Earth observation satellite system operating from space. It is run by Spot Image based in Toulouse, France. SPOT4 was launched in 1998 and SPOT5 in 2002. SPOT5 has two high resolution geometrical instruments and has a resolution of 2.5 to 5 meters in panchromatic mode and 10 meters in multispectral mode. A drawback of SPOT images is the very high price. Depending on the resolution, areal extent and date of images various prices are asked. Purchasing images for the entire Nile Basin is therefore no option and only for specific detailed projects one could consider purchasing images.

SPOT offers also free NDVI images as 10 days composites at a resolution of 1 km. These images are already georeferenced and cloud-filtered. Data can be obtained from http://free.vgt.vito.be/

3.2.6.4 Recommendation

Vegetation indices are primarily used to support land cover classification and as input to hydrological models. For visualization purposes vegetation maps (sometimes combined with elevation) are used as well. For the Nile Basin DSS it is recommended to use MODIS images as these provides also derived products as LAI and NPP. Landsat GeoCover2000 images can be used for visualization using three bands. For specific local projects original Landsat images (starting from 2009 for free) can be used.

3.2.7 Evapotranspiration

There is a growing consensus that evapotranspiration is a valuable calibration parameter for hydrological models. Various SEBAL analysis has been undertaken for the Nile Basin over the last years and should be included in the Nile Basin DSS.

3.2.8 Groundwater

Groundwater information can be divided into aquifer characteristics (location, KD values) and groundwater levels. Public domain datasets and remote sensing can hardly be used to get this information. There are however interesting developments on two fronts. First of all groundwater, or better phrased soil moisture contents, can be observed using various satellites. [Ref Martine Rutten]

GRACE (Gravity Recovery And Climate Experiment) is a twin-satellite mission, developed to measure changes in the Earth's time-variable gravity field (Tapley et al., 2004). GRACE data are available since May 2002. GRACE data products are expressed in mm equivalent water. Two factors are important when evaluating results. Firstly, no distinction between snow cover, soil moisture and deep groundwater storage can be made. Secondly, results are given relative to the long-term average from Apr-2002 to Apr-2006. This means that no absolute values of water storage can be provided and that no spatial differences in water stored can be observed. In other words GRACE detects only changes in stored water. GRACE data is resolution independent, but data can only be used at resolutions of at least 100 x 100 km^2.

4 Data Gaps

One of the objectives of this Inception Report is to determine data gaps, defined here as data that requires substantial efforts to obtain. Based on the data requirements, as described in Chapter 2, and the dataset available (Chapter 3) this data gap identification is presented in Table 3.

The first basin-wide DSS development shall primarily be focused on water balance and allocation. In order to reflect this in the data prioritization Table 3 includes as well a column with priority, ranging from 1 (necessary for the first phase), 2 (required for the 8 priority areas as defined in Chapter 2), and 3 (advanced future capabilities for the DSS). Note that most of the data in Table 3 has priority 1 as the focus of this consultancy was already geared towards the preliminary data collection.

Table 3. Data gap identification. Numbers in "source" refer to the sections in this report, letters refer to additional comments.

Availability:

- *+: available*
- *o: not clear*
- *-: requires substantial efforts*

Priority

- *1: required for the first phase of the DSS*
- *2: required for the complete DSS*
- *3: required for future DSS applications*

	Availability	Source	Priority
Schematization and parameterization			
Sub-basins and rivers			
DEM	+	3.2.3.1	1
River networks	+	3.2.3.1 (A)	1
Profile	-	(B)	1
Resistance	-	(B)	1
Slope	+	3.2.3.1	1
Land Surface			
Land cover / land use	+	3.2.5.6	1
Vegetation indices			
NDVI	+	3.2.6.1	2
NPP	+	3.2.6.1	2
LAI	+	3.2.6.1	1
Albedo	+	3.2.6.1	1
Irrigated areas	+	3.2.5.5	1
Management			
Cropping calendar	-	(C)	2
Irrigation scheduling	-	(C)	2
Soil			
Soil types			
Depth	+	3.2.4.5	1
Texture	+	3.2.4.5	1
Hydraulic properties	o	3.2.4.1	1
Water bodies			
Reservoirs			
Area	+	3.2.6.2	1
Volume	-	(D)	1
Operation	-	(D)	1
Lakes			
Area	+	3.2.6.2	2

Volumes	-	(E)	2
Groundwater			
Aquifers			
Extent	-	(F)	3
Capacity	-	(F)	3
Recharge	-	(F)	3
Abstractions	-	(F)	3
Water quality			
Sources			
Point	-	(G)	3
Diffuse	-	(G)	3
Forcing			
Meteorology			
Precipitation	+	3.2.1.2	1
Temperature	+	3.2.1.1	1
Radiation	+	3.2.1.1	1
Wind speed	o	3.2.1.1	1
Relative humidity	+	3.2.1.1	1
Calibration and validation			
Hydrology			
Discharges	o	3.2.2 (H)	1
Evapotranspiration	+	3.2.7	1
Soil moisture	o	(I)	1
Lakes			
Water levels	o	3.2.2.3	
Reservoirs			
Water levels	o	3.2.2.3	1
Releases	-	(D)	1
Groundwater			
Water levels	-	(D)	3
Water quality			
Concentrations	-	(G)	3

Additional comments on letters in the data gap identification table (Table 3):

- A: Derived from the delineation of streamflow and subbasins based on DEM.
- B: Requires local data gathering and information.
- C: Requires local information from extension services.
- D: Ministries of Water Resources/Irrigation.
- E: Publications, local sources.
- F: Ministries of Water Resources and large scale trends from GRACE.
- G: Ministries of Environment.
- H: Additional data from local sources: Ministries of Water Resources.
- I: Innovative satellite information from SMOS to be explored

5 Data organization and archiving

The entire design and development of a database management system will be undertaken independently under WP1 and WP2 of the DSS development. In this Chapter reference is made to these aspects that should be considered during this preliminary data collection, processing and archiving.

5.1 Definition

A relational database is a collection of data items organized as a set of formally-described tables from which data can be accessed or reassembled in many different ways without having to reorganize the database tables. The relational database was invented by E. F. Codd at IBM in 1970. The standard user and application program interface to a relational database is the structured query language (SQL). SQL statements are used both for interactive queries for information from a relational database and for gathering data for reports. In addition to being relatively easy to create and access, a relational database has the important advantage of being easy to extend. After the original database creation, a new data category can be added without requiring that all existing applications be modified. A relational database is a set of tables containing data fitted into predefined categories. Each table (which is sometimes called a relation) contains one or more data categories in columns. Each row contains a unique instance of data for the categories defined by the columns. When creating a relational database, you can define the domain of possible values in a data column and further constraints that may apply to that data value. For example, a domain of possible customers could allow up to ten possible customer names but be constrained in one table to allowing only three of these customer names to be specifiable. The definition of a relational database results in a table of metadata or formal descriptions of the tables, columns, domains, and constraints.

5.2 Requirement for the NBI database

Fedra (2008) has described the requirements for the databases underlying the DSS. The information system is a crucial component of the DSS and should consist of data bases and GIS data that manage background information including reports, observation data including optionally real-time monitoring, and the model inputs, outputs and scenarios.

The data base will be used to

- Store all shared, common basin data, GIS layers, time series data, and links to the Document Management and Information System; a set of basic object to be represented and related to the water budget Model's network architecture is described above;
- Store any local data sets beyond the common, shared core data.
- Store model scenarios, parameters, model results; while primarily produced "locally" they can be exported for analysis at the next higher (sub regional and basin wide) level, or provide inputs for any downstream scenarios;
- User management and access information and control;
- Coordination between cascading models and data within a given interactive session (maintaining multiuser capabilities) using a blackboard architecture.

Fedra (2008) recommends an object oriented design on an industry standard RDBMS such as ORACLE, Sybase, MySQL or PostgreSQL.

5.3 Industry standards[1]

There is a wide variety of relational database management systems (RDBMS) either in the commercial or public domain. Here the most widely used RDBMS systems are briefly compared. Among the top commercial RDBMS are Oracle (Oracle Corporation), DB2 and Informix (IBM), Access and SQL Server (Microsoft), Adapative Server Enterprise and Advantage Database Server (Sybase), Teradata (Teradata) and Interbase (Codegear). The most important databases in the public domain are MySQL (Sun Microsystems), PostgreSQL (PostgreSQL Global Development Group) and Firebird (Firebird project).

These RDMBS are compared based on a number of criteria:

- The type of software license
- The operating systems that the RDMBS runs on.
- ACID criteria. ACID stands for Atomicity, Consistency, Isolation and Durability. Atomicity refers to the ability of the DBMS to guarantee that either all of the tasks of a transaction are performed or none of them are. The Consistency property ensures that the database remains in a consistent state before the start of the transaction and after the transaction is over (whether successful or not). Consistency states that only valid data will be written to the database. Isolation refers to the requirement that other operations cannot access or see the data in an intermediate state during a transaction. This constraint is required to maintain the performance as well as the consistency between transactions in a DBMS system. Durability refers to the guarantee that once the user has been notified of success, the transaction will persist, and not be undone. This means it will survive system failure, and that the database system has checked the integrity constraints and won't need to abort the transaction.
- Referential Integrity in a relational database is consistency between coupled tables. Referential integrity is usually enforced by the combination of a primary key or candidate key (alternate key) and a foreign key. For referential integrity to hold, any field in a table that is declared a foreign key can contain only values from a parent table's primary key or a candidate key.
- Interface defines the method for communication either through SQL, a graphical user interface (GUI) or an Application Programming Interface (API).
- The maximum size of the database
- The maximum size of an individual table
- The access control

[1] Partly based on http://en.wikipedia.org/wiki/Comparison_of_relational_database_management_systems

Table 4 Comparison between industry standard RDBMS

Database	Maintainer	Type	Operating System	ACID	Referential Integrity	Interface	Maximum size	Maximum table size	Access control	Link
Oracle	Oracle	commercial	Windows, Mac, Unix, Linux	Yes	Yes	SQL	unlimited	4 GB x block size	high	www.oracle.com/database/index.html
DB2	IBM	commercial	Windows, Unix, Linux	Yes	Yes	GUI, SQL	unlimited	512 TB	high	www.ibm.com/db2
Informix	IBM	commercial	Windows, Mac, Unix, Linux	Yes	Yes	?	?	?	?	www.ibm.com/informix
Access	Microsoft	commercial	Windows	No	Yes	GUI, SQL	2 GB	2 GB	low	www.office.microsoft.com/access
SQLServer	Microsoft	commercial	Windows	Yes	Yes	GUI, SQL	524258 TB	524258 TB	high	www.microsoft.com/sqlserver
ASE	Sybase	commercial	Windows, Unix, Linux	Yes	Yes	SQL	unlimited	16 EB	high	www.sybase.com/ase
ADS	Sybase	commercial	Windows, Linux	Yes	Yes	SQL, API	?	?	?	www.sybase.com/ads
TeraData	TeraData	commercial	Windows, Unix, linux	Yes	Yes	SQL	unlimited	unlimited	?	www.teradata.com
Interbase	Codegear	commercial	Windows, Mac, Unix, Linux	Yes	Yes	SQL	?	?	?	www.interbase.com
MySQL	Sun microsystems	public domain	Windows, Mac, Unix, Linux	Yes	Yes	SQL	unlimited	2 GB	medium	http://www.mysql.com/
PostgreSQL	PostgreSQL Global Development Group	public domain	Windows, Mac, Unix, Linux	Yes	Yes	SQL	unlimited	32 TB	high	www.postgresql.com
Firebird	Firebird project	public domain	Windows, Mac, Unix, Linux	Yes	Yes	SQL	unlimited	32 TB	medium	www.firebirdsql.org

5.4 GIS and integration with RDBMS

Over the past five years there has been a confluence of Geographic Information System (GIS) technology, RDBMS architecture and SQL standards that has fostered the implementation of spatial processing within the RDBMS. The development of spatial RDBMS in relation to GIS has undergone many changes over the last decades and this is illustrated in Figure 3.

GIS software products started out as closed systems. A GIS was an application that got all its data from files with proprietary structures. Even though most GIS applications tend to be specialized and support only small departments or workgroups, it quickly became apparent that the problems it solved and the data it used was related to the data stored in central databases. Thus, many GIS products supported client-server connectivity to databases for access to additional attributes. Sometimes through a live connection, sometimes limited to batch download of a copy of a table.

As the volume and usage of spatial data grew, the benefits of applying industrial-strength relational database management (multi-user support, data integrity, transaction support, etc.) to spatial, as well as alphanumeric, data became obvious. Moreover, the separation of spatial features and their associated attributes led to integrity problems. So GIS vendors built middleware systems that act as front ends to databases and translate the spatial queries from the GIS application into standard SQL for the database. The middleware uses BLOBs[1] to store the proprietary data and index structures; the database doesn't know understand the contents, but at least referential integrity is preserved and the spatial data participates in all the backup and storage processes of the database. In the case of ESRI, the middleware is called ArcSDE, from SDE (Spatial Database Engine) and to this day this is how it works with SQLServer and most Oracle installations.

In the next evolutionary step toward more open and interoperable systems, databases acquired spatial types, functions, and indexes, and the specialized, proprietary middleware from GIS vendors either became thinner (e.g., ArcSDE) or, in some cases (e.g., MapInfo SpatialWare), was eliminated altogether. This opens up the use of spatial functionality in applications that aren't primarily GIS. In addition, it allows for the application of spatial logic in stored procedures and triggers, so that spatial data participates fully in the business logic organizations wish to apply in the database, for reasons of performance, security, and maintainability. This in no way diminishes the role of the GIS software vendors and their products. The ability to apply spatial business logic in the server and support spatial queries in otherwise non-GIS applications does not replace but complements the GIS software and enhances its value.

---

[1] A binary large object, also known as a blob, is a collection of binary data stored as a single entity in a database management system. Blobs are typically images, audio or other multimedia objects, though sometimes binary executable code is stored as a blob.

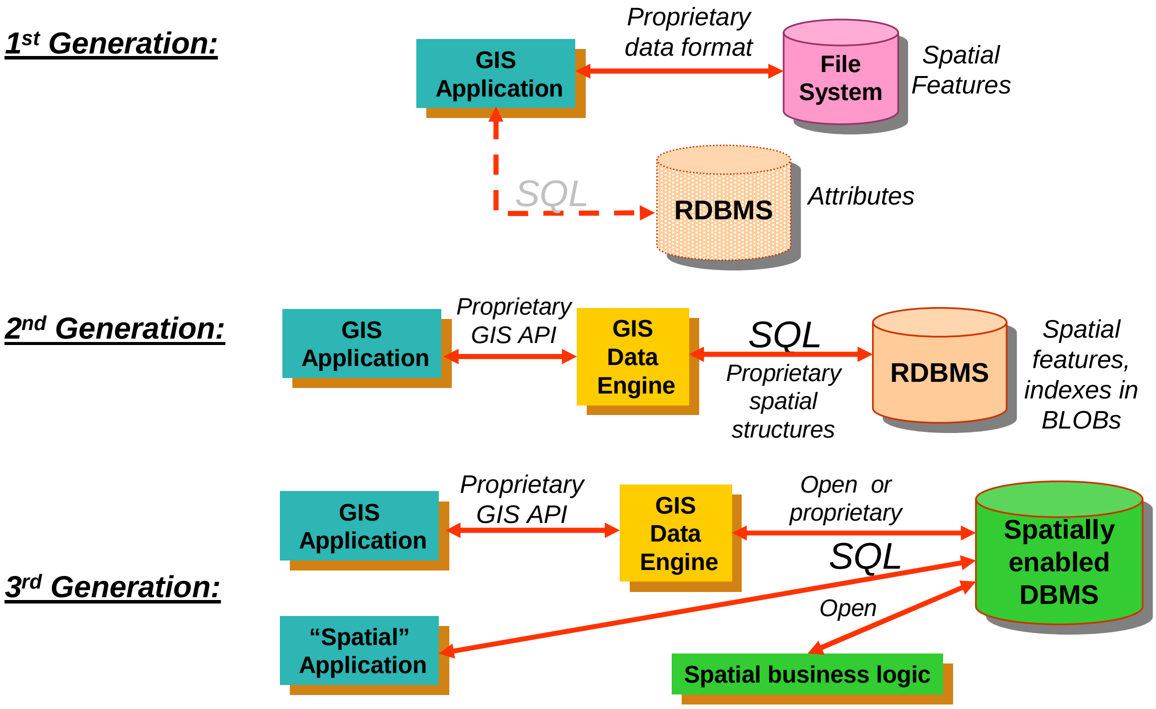

Figure 3 Architectural evolution of GIS data management (source: IBM)

Much of the data is required to build the Nile basin DSS has a spatial aspect and the three generations basically describe the different options. From a technical perspective a first generation solution is most straightforward and a 3rd generation solution is technologically complicated and expensive but offers the same database functionality (ACID, referential integrity, SQL) for spatial data as for regular data.

5.4.1 *Spatial RDBMS integrated with GIS software*

There are a number of RDBMS that are able to deal with spatial data and which are supported by mainstream GIS software. The most important ones are described here.

Oracle spatial[1]

Oracle Spatial, an option for Oracle Database Enterprise Edition, includes advanced spatial capabilities to support GIS applications, location-based services, and enterprise spatial information systems. Oracle Spatial extends the core location features included in every Oracle database. Its advanced data manipulation and spatial analysis features include buffer generation, spatial aggregates, area and length calculations, and linear referencing. Since version 10g Oracle introduced a GeoRaster datatype to store and manage image and gridded raster data and metadata. Oracle spatial is supported by all major GIS software.

DB2 spatial extender[1]

[1] http://www.oracle.com/technology/products/spatial/index.html

DB2 Spatial Extender extends the function of DB2 Universal Database with a set of advanced spatial data types that represent geometries such as points, lines, and polygons and many functions and features that interoperate with those new data types. These capabilities allow you to integrate spatial information with your business data, adding another element of intelligence to your database. DB2 Spatial Extender is supported by industry-leading spatial and GIS applications and tools such as: ESRI's ArcGIS, MapInfo, MapXtreme and GeoTools, GeoServer and uDig open source projects.

Informix spatial datablade[2]

Similar to the spatial extender the spatial datablade provide spatial support for IBMs Informix.

SQLServer[3]

Microsoft SQL Server 2008 delivers comprehensive spatial support. A new geography data type to store geodetic spatial data has been added and several operations can be performed on it. and perform operations on it. The management studio has a special spatial interface to quickly perform and view spatial queries.

PostGis for PostgreSQL[4]

PostGIS adds support for geographic objects to the PostgreSQL object-relational database. In effect, PostGIS "spatially enables" the PostgreSQL server, allowing it to be used as a backend spatial database for geographic information systems (GIS), much like ESRI's SDE or Oracle's Spatial extension. PostGis is also supported by the mainstream GIS software. PostGis is licensed under a GNU general public license and is thus similar to PostGreSQL.

5.4.2 GIS software supporting spatial RDBMS

There is a wide variety of both commercial and public domain GIS software with support spatial RDBMS. An excellent overview of public domain GIS systems is given on www.freegis.org. The most important packages that support spatial RDBMS include:

- GRASS
- GeoTools
- GeoServer
- Udig
- QGIS

The most important commercial packages are ESRI's ArcGIS (including ArcGIS Server) and MapInfo (including SpatialWare). ArcGIS server and SpatialWare are both spatial data engines that interact between frontend GIS software and backend spatial RDBMS. ArcGIS server supports Oracle, DB2, Informix, SQL server and PostGreSQL and MapInfo SpatialWare supports Oracle, DB2, Informix and SQL server.

---

[1] http://www.ibm.com/software/data/spatial/db2spatial/
[2] http://www.ibm.com/software/data/informix/blades/spatial/
[3] http://www.microsoft.com/sqlserver/2008/en/us/spatial-data.aspx
[4] http://postgis.refractions.net/

A clear distinction should be made between vector data and raster data. Vector data can be stored and retrieved using Open GIS Consortium (OGC) compliant standards, either defined by the spatial data engine or predefined by the RDBMS. The only RDBMS that has a data type defined for raster data is Oracle Spatial. However raster data can be stored in the other RDBMS in the binary large object data type (BLOB)[1]. In that case the database engine facilitates this storage including metadata.

5.5 Existing hydrological information systems

There is a number of existing information systems specifically tailored towards hydrological applications that use part of the aforementioned database technologies. However most of these systems are not state-of-the-art and up to speed with the current database and GIS technology that is available. A number of important hydrological information systems are briefly discussed here.

5.5.1 Hymos[2]

HYMOS is developed by WL Delft hydraulics and is the information system for water resources management in general, but tailored towards surface water. It covers all data storage and processing requirements for analysis, planning, design and operation of water management systems. HYMOS is time series oriented with common facilities for spatial data analysis. A wide variety of data processing and analysis features make HYMOS a powerful tool in water related studies, research and consultancy. HYMOS is compliant with major database systems such as Oracle and SQLServer.

5.5.2 ArcHydro[3]

ArcHydro is a data model developed at the Center for Research in Water Resources of the University of Texas. The Arc Hydro Data Model can be defined as a geographic database containing a GIS representation of a Hydrological Information System under a case-specific database design which is extensible, flexible, and adaptable to the user requirements. It takes advantage of the next generation of spatial data in Relational Database Management Systems (An RDBMS-based GIS System), the geodatabase model. Conceptually, it is a combination of GIS objects enhanced with the capabilities of a relational database to allow for relationships, topologies, and geometric networks.

[1] http://webhelp.esri.com/arcgisdesktop/9.3/index.cfm?TopicName=How%20raster%20data%20is%20stored%20in%20a%20geodatabase

[2] http://www.wldelft.nl/soft/hymos/int/index.html

[3] http://www.crwr.utexas.edu/giswr/hydro/ArcHOSS/model/index.cfm

5.5.3 CUAHSHI-HIS[1]

The hydrogical informations system (HIS) of the Consortium of Universities for the Advancement of Hydrologic Science (CUASHI) provides web services, tools, standards and procedures that enhance access to more and better data for hydrologic analysis. The entire systems is still under development and will be very powerful but at the same time very complex.

5.6 Conclusions

The choice of the RDBMS and GIS software is crucial in the development of the DSS. The final choice depends on the following criteria:

- Which modular modeling systems are selected and what are their data requirements?
- What is the preferred RDBMS-GIS systems architecture?
- How much financial resources are available to acquire commercial RDBMS and GIS software?
- How well are the database operators trained and experiences in the use of RDBMS and GIS technology?
- How many licenses/users are required?
- Which modeling systems will be connected to the RDBMS?

The NBI has indicated its preference for independent public domain software and therefore a choice for PostgreSQL is recommendable. PostGreSQL is a state-of-the-art database which is comparable to commercial RDBMS, but it is in the public domain.

For GIS data three options are available:

Option 1: PostGreSQL/PostGIS and public domain GIS
PostGIS can be used in combination with PostgreSQL to manage geographic data. However currently PostGIS has no datatype for raster layers, but this could change in the near future. The Geospatial Data Abstraction Library (GDAL)[1] is a public domain translator library for raster geospatial data formats that is released under an X/MIT style Open Source license by the Open Source Geospatial Foundation. It is likely that soon raster formats can be translated into a PostGIS compatible format. A public domain GIS (e.g. QGIS, UDig, GeoServer, GRASS) can then be used to display and analyze the geographic data.

Option 2: PostgreSQL/PostGIS and ESRI software
PostgreSQL is also supported by ESRI software. Since version 9.3 ArcSDE technology is now included in ArcGIS Server. PostgreSQL support is available and there is support for both the PostGIS and ESRI data types. The advantage of this choice is that ESRI has defined its own BLOB like raster data type that can be stored in the PostgreSQL database. Moreover ArcGIS is the most used GIS system in the world and can be considered as the de-facto standard.

---

[1] http://his.cuahsi.org

However, this choice will require considerable financial resources as both the enterprise edition of ArcGIS Server and the Desktop GIS ArcInfo need to purchased.

Option 3: PostgreSQL for time series and ESRI desktop for GIS data

In this option the GIS database is separated from the time series data (including model results, meteorological data, discharges, etc.). ESRI desktop GIS software can be used to manage the GIS data (both raster and vector) and metadata can be managed using the ArcCatalog software. In this case there is no need to purchase ArcGIS Server. PostGreSQL can be used to manage and store time series data. Through Open Database Connectivity (ODBC) protocols the database data can be joined to geographic datasets for visualization in GIS. Although this option is less advanced it is easier to implement from a technological perspective.

It is emphasized that independent of the technological solution the success of the Nile Basin information systems is in the quality and documentation of the data. No technological solution can circumvent the often used saying “garbage in = garbage out”. Therefore we strongly recommend focusing on consistent data acquisition and documentation in this phase of the DSS development.

6 Projections

The features on a map reference the actual locations of the objects they represent in the real world. The positions of objects on the earth's spherical surface are measured in geographic coordinates. While latitude and longitude can locate exact positions on the surface of the earth, they are not uniform units of measure; only along the equator does the distance represented by one degree of longitude approximate the distance represented by one degree of latitude. To overcome measurement difficulties, data is often transformed from three-dimensional geographic coordinates to two-dimensional projected coordinates.

Because the earth is round and maps are flat, getting information from a curved surface to a flat one involves a mathematical formula called a map projection. This process of flattening the earth will cause distortions in one or more of the following spatial properties:

- Distance
- Area
- Shape
- Direction

No projection can preserve all these properties; as a result, all flat maps are distorted to some degree. Each projection is distinguished by its suitability for representing a particular portion and amount of the earth's surface and by its ability to preserve distance, area, shape, or direction. Some map projections minimize distortion in one property at the expense of another, while others strive to balance the overall distortion.

The choice of projection is subject to a number of important considerations:

- Which spatial properties need to be preserved
- Where is the area that is mapped?
- What shape is the area that is mapped? Is it square or is it wider in the east–west direction?
- How big is the area that is mapped? On small-scale maps, where a small distance on the map represents a considerable distance on the earth, distortion may have a bigger impact, especially if you use your map to compare or measure shape, area, or distance.

Map projections can be generally classified according to what spatial attribute they preserve:

- Equal area projections preserve area. Many thematic maps use an equal area projection. Maps of the United States commonly use the Albers Equal Area Conic projection.
- Conformal projections preserve shape and are useful for navigational charts and weather maps. Shape is preserved for small areas, but the shape of a large area, such as a continent, will be significantly distorted. The Lambert Conformal Conic and Mercator projections are common conformal projections.
- Equidistant projections preserve distances, but no projection can preserve distances from all points to all other points. Instead, distance can be held true from one point (or a few points) to all other points or along all meridians or parallels.
- Azimuthal projections preserve direction from one point to all other points. This quality can be combined with equal area, conformal, and equidistant projections, as in the Albers Equal Area Azimuthal, Lambert Equal Area Azimuthal and the Azimuthal Equidistant projections.

- Other projections minimize overall distortion but don't preserve any of the four spatial properties of area, shape, distance, and direction.

In case of the Nile basin the choice of projection is very important because we deal with an extremely large area (around 3 million km^2) that is irregularly shaped and is smaller in east-west direction than in north south direction. It is also an equatorial area and for modeling and hydrological analysis purposes conservations of area is more important than distance, shape or direction.

Given these prerequisites two candidate projections are proposed:

1. Sinusoidal projection. This is an equal area pseudo–cylindrical projection where all parallels and the central meridian are straight. The meridians are curves based on sine functions with the amplitudes increasing with the distance from the central meridian. There is no shape distortion along the central meridian and the equator, the areas are represented accurately, directions are corrected along the central meridian and the equator but distorted elsewhere and distance along all parallels and the central meridian of the projection is accurate. The projection is particularly suitable for area with a distinct north-south orientation. The central meridian of the projection needs to be specified. In case of the Nile basin 33°E is recom mended.
2. Lambert Azimuthal Equal Area. This is an azimuthal equal area projection. This projection preserves the area of individual polygons while simultaneously maintaining a true sense of direction from the center. The general pattern of distortion is radial. It is best suited for individual land masses that are symmetrically proportioned. Shape is minimally distorted, less than 2 percent, within 15° from the focal point, the area is preserved and true direction radiating from the central point and distance is true at the center and scale decreases with distance from the center along the radii and increases from the center perpendicularly to the radii. This projection is particularly suitable for square areas. The latitude of the origin and the central meridian need to be specified. In case of the Nile basin 5°S and 33°E are recommende d.
3. Albers Equal Area Conic projection. This conic projection uses two standard parallels to reduce some of the distortion of a projection with one standard parallel. Although neither shape nor linear scale is truly correct, the distortion of these properties is minimized in the region between the standard parallels. This projection is best suited for land masses extending in an east-to-west orientation rather than those lying north to south. Shape along the standard parallels is accurate and minimally distorted in the region between the standard parallels and those regions just beyond. The 90° angles between meridians and parallels are preserved, but because the scale along the lines of longitude does not match the scale along the lines of latitude, the final projection is not conformal. All areas are proportional to the same areas on the earth. Direction is locally true along the standard parallels. Distances are most accurate in the middle latitudes. Along parallels, scale is reduced between the standard parallels and increased beyond them. Along meridians, scale follows an opposite pattern. The latitude of the origin, central meridian and two standard parallels need to be specified

To prevent negative coordinates usually a false easting and northing are specified for all three projections. This is a linear shift in coordinates in both the x (easting) and y (northing) direction.

There are three fundamental rules of thumb regarding map projection and latitude:

- If the country lies in the tropics, use a cylindrical projection;
- If the country lies in the temperate latitudes, use a conical projection;

- If the country is required to show one of the polar regions, use a azimuthal projection.

Based on the latitude, the shape of the Nile basin and the criterion of acreage conservation the following projection is recommended to use in the Nile basin DSS project:

Projection type:	Sinusoidal
Central meridian:	33° E
False easting:	1,100,000
False northing:	600,000

According to the ArcGis manual the sinusoidal projection is also recommended for the Nile Basin: This projection maintains equal area despite conformal distortion. Alternative formats reduce the distortion along outer meridians by interrupting the continuity of the projection over the oceans and by centering the continents around their own central meridians, or vice versa.

Projection method
A pseudo cylindrical projection where all parallels and the central meridian are straight. The meridians are curves based on sine functions with the amplitudes increasing with the distance from the central meridian.

Linear graticules
All lines of latitude and the central meridian.

Properties

- Shape
    - No distortion along the central meridian and the equator. Smaller regions using the interrupted form exhibit less distortion than the uninterrupted sinusoidal projection of the world.
- Area
    - Areas are represented accurately.
- Direction
    - Local angles are correct along the central meridian and the equator but distorted elsewhere.
- Distance
    - The scale along all parallels and the central meridian of the projection is accurate.

Limitations
Distortion is reduced when used for a single land mass rather than the entire globe. This is especially true for regions near the equator.

Uses and applications
Used for continental maps of South America, Africa, and occasionally other land masses, where each has its own central meridian.

7 Conclusions

The Nile Basin DSS will provide the necessary knowledge base and analytical tools to support the planning of cooperative joint projects and the management of the shared Nile Basin water resources. Various work packages have been developed to built the DSS of which WP2 foresees in the bulk of the data compilation activities. To ensure a fast release of the first version of the DSS a relatively small activity was initiated to carry out a preliminary data collection. This report describes the inception phase of this activity which will be followed up by the analysis phase.

During the analysis phase the following activities will be carried out:

- Collecting and gathering data as described and recommended in this inception report.
- Store data in a common framework.
- Check data on consistency.
- Update data gaps.
- Discuss options to resolve data gaps.

8 References

Abuzeid, K.M., Hartveld, A., Riddel, P., Seidelmann, R., Tamrat, I. 2008. Needs assessment and conceptual design of the Nile basin decision support system consultancy: Final inception report. Hydrophil. Vienna. Austria

Droogers, P., D. Seckler and I. Makin. 2001. Estimating the potential of rainfed agriculture. IWMI Working Paper 20.

Droogers, P., C. Perry. 2008. Scenario based water resources model to support policy making. Report FutureWater 79.

Droogers, P., W. Immerzeel, C. Perry. 2009. Application of Remote Sensing in National Water Plans: Demonstration cases for Egypt, Saudi-Arabia and Tunisia. Report FutureWater: 80.

Droogers, P. 2008. MRC-IKMP Modelling Component: Final Report. Mekong River Commission.

FAO, 1974 Soil Map of the World, 1:5.000.000, Volume I, Legend; Volume VI, Africa. Unesco, Paris.

FAO/IIASA/ISRIC/ISSCAS/JRC, 2008. Harmonized World Soil Database (version 1.0). FAO, Rome, Italy and IIASA, Laxenburg, Austria.

Fedra, K. 2008. Needs assessment and conceptual design of the Nile basin decision support system consultancy: Final analysis and DSS design report. Hydrophil. Vienna. Austria

Immerzeel, W.W. and P. Droogers, 2008. Calibration of a distributed hydrological model based on satellite evapotranspiration, J. of Hydr. 349: 411-429

Loveland, T.R., Zhu, Z., Ohlen, D.O., Brown, J.F., Reed, B.C., and Yang, L., 1999. An Analysis of the IGBP Global Land-Cover Characterization Process. Photogrammetric Engineering and Remote Sensing , v. 65, no. 9, p. 1021-1032.

Mayaux, P., E. Bartholomé, S. Fritz, A. Belward. 2004. A new land-cover map of Africa for the year 2000. Journal of Biogeography (J. Biogeogr.) (2004) 31, 861–877

Mitchell T.D., T.R. Carter, P.D. Jones, M. Hulme, M. New. 2004. A comprehensive set of high resolution grids of monthly climate for Europe and the globe: the observed records (1901 - 2000) and 16 scenarios (2001-20016), Climate Research Unit, School of Environmental Sciences, University of East Anglia, UK

Siebert, S., Hoogeveen, J., Frenken, K. 2006. Irrigation in Africa, Europe and Latin America. Update of the Digital Global Map of Irrigation Areas to Version 4. Frankfurt Hydrology Paper 05, University of Frankfurt (Main), Germany and FAO, Rome, Italy.

Tapley, B.D., Bettadpur, S., Ries, J.C., Thompson, P.F. and Watkins, M.M., 2004. GRACE measurements of mass variability in the earth system, Science, pp. 503-505.

Van Engelen, V.W.P., T.T. Wen. 1995. Global and National Soils and Terrain Digital Databases (SOTER). Procedures Manual. International Soil Reference and Information Centre, Wageningen, Netherlands.