Preliminary Data Compilation for the Nile Basin Decision Support System

ANALYSIS PHASE REPORT

March 2010

FutureWater Report 92

Commissioned by
Nile Basin initiative

Authors
Dr. P. Droogers
Dr. 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. These 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 second phase of the project (Analysis Phase) and is a result of the following activities:

- Global data collection
- Local data collection
- Developing PostGreSQL database
- Data processing
- Demonstrating relevant hydrological parameters derived from data

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. 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 (about one-third of the area of the Basin) yield negligible flows; whereas, the Highland of Ethiopia, comprising 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 riparian countries 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.

The initial first phase of the DSS development is 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 is a substantially revised version of report Version 4 where a straightforward database structure was developed. However, given the advanced interface that will be developed it was decided to create a more sophisticated PostGreSQL database.

2 Methodology

2.1 Introduction

The methodology to develop a preliminary data base to be used by the Nile Basin Initiative Decision Support System (NBI-DSS) under development has been divided into four main activities:

- compilation and structuring of relevant data,
- processing the compiled data,
- quality control of the compiled and processed data,
- organizing and archiving compiled data.

These four main steps will be further explained in the following sections, while the next Chapters will provide details per data type.

2.2 Compilation and structuring

Data compilation and structuring can be considered to cover four main domains: (i) assessment of data requirements, (ii) inventory of data availability, (iii) identifying potential data gaps, and (iv) define a structure to store data. All these issues have been covered in the Inception Report (Droogers and Immerzeel, 2009). Main conclusion of this Inception Report is summarized in Table 1.

Table 1. Data gap identification. Numbers in “source” refer to the sections in this Inception Report, letters refer to additional comments.

	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

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

Additional comments on letters in the data gap identification table (Table 1):

- *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*

2.3 Processing

Global and local data as defined during the Inception Phase have been collected and processed. Data were in all different formats and a lot of emphasis was put on processing data into a uniform format. Since the final decision on formats and data base systems will be made during WP2 of the NBI-DSS project, the most transparent formats have been used to ensure easy transformation to the final formats.

For point data the following tables have been used (see Appendix 3):

- tbl_countries
- tbl_data tbl_parameters
- tbl_series
- tbl_session
- tbl_source
- tbl_station_param_types
- tbl_stations
- tbl_time_steps
- tbl_units
- tbl_value_types

All spatial data have been converted to ArcGis where point, line and polygon data have been stored as Shape files, and raster data in ArcGis GRD format.

2.4 Quality control

Data quality control is a broad term reflecting many aspects of the entire process of data collection, data reporting, data storage, error checking etc. In Google the search terms “data quality control” returned about 90,000,000 pages! There are also many definitions of data quality control. Some are more restrictive and relate only to how the data reflects reality, like: “data quality refers to the degree of excellence exhibited by the data in relation to the portrayal of the actual phenomena”. Other definitions take into account the objective of the use of the data, like “the state of completeness, validity, consistency, timeliness and accuracy that makes data appropriate for a specific use”. The quality control and validity checks as described under the specific sections of this report relate to the latter definition as the database is prepared for a specific purpose: to develop a database to support decision making with a strong link to modeling.

One of the advantages of the data as collected under this project is that all data originate from existing data bases, so rigorous quality control has been undertaken already.

2.5 Organizing and archiving

The last step in the data flow from compilation to processing to quality control is to organize and archive the data. As described in the Inception Report quite some options exist ranging from straightforward to very complex systems. It has been proven that the most simple system is often the best, as long as it is not too simple. Moreover, the system must be transparent and flexible to ensure that converting to other systems is possible if so required under WP 2 of the NBI-DSS project.

Regarding spatial data no final decision of the specific format has been made. It was therefore decided to store all data in one of the most commonly used GIS system: ESRI ArcGis. All spatial data have been converted to ArcGis where point, line and polygon data have been stored as Shape files, and raster data in ArcGis GRD format.

In summary all point sourced data are stored in a PostGreSQL data base, while the spatial data are stored in ArcGis. In this way, data can be converted easily to any other system.

2.6 Local Data Collection

2.6.1 Introduction

One remaining issue of the data compilation was the local data. Details about the global data, sources, formats etc. are all provided in the Inception Report. However, the Inception Report did not provide substantial information on the availability of local data and it was decided to include this issue in this Analysis Report.

A total of about 60 GB of local information was made available. This information ranges from photographs, reports, posters, GIS layers, and time series data. Data vary a lot and ranges from very relevant for the DSS till less relevant. A complete description and compilation of these 60 GB is provided in Appendix 1.

2.6.2 Nile-DST

Special emphasis is given to the FAO Nile-DST project, which has just been terminated. The Nile DST (Decision Support Tool) is one component of the project “Information Products for Nile Basin Water Resources Management”. This project was intended to strengthen the ability of the governments of the ten Nile countries to take informed decisions with regard to water resources policy and management in the Nile basin. A thorough understanding of the state of the Nile resource, and the current use and productivity of its waters, will enable decision makers to better assess trade-offs and implications of shared-vision development scenarios.

The project was supported by the Government of Italy and carried out under the umbrella of the Nile Basin Initiative, of which Italy is a full partner. It is implemented by the ten Nile riparians with technical and operational assistance of the Food and Agriculture Organization of the United Nations (FAO). The project had a budget of US$ 5 million, bringing the total contribution since 1996 of the Government of Italy to the Nile process to over US$ 16 million. The project became operational in November 2004 and activities were completed by May 2008.

The actual Nile-DST was developed by the Georgia Water Resources Institute developed the under a contract with FAO . It was released by the Nile Council of Ministers for Water Affaires

(Nile COM) in February 2003. The Nile-DST is a prototype software that models the entire Nile system and serves to assess the trade-offs and consequences of alternative cross-sectoral and basin-wide development scenarios. Further development of decision support systems in the Nile basin is currently being implemented by the Water Resources Planning and Management Project.

Data compiled during the Nile-DST project can be considered as resulting in the best possible data set ever given the resources put into the project. One could therefore claim that in terms of temporal data this is the best source currently available to be used for the DSS. Details of the compilation and structuring of the Nile-DST data are further explained in Appendix 1.

2.6.3 ENTRO's OSI

A project has been executed entitled: Eastern Nile Subsidiary Action Program (ENSAP) Multipurpose Development of the Eastern Nile, One-System Inventory (OSI). The entire database as developed consisted out of 206 excel files and 16 shape files. The excel files were a mixture of actual data (precipitation, evaporation, flows, sediments), calculations, analysis, data filling etc.

In terms of actual data 37 stations with precipitation records were found (file: RainFall.xls) and 15 stations with evaporation records. Given the nature of the project (preliminary data analysis), tools were developed to implement ENTRO-OSI data and were actual applied on three stations and imported in the database:

- Addis Zemen (precipitation)
- Ambo (precipitation)
- Agaro (evaporation)

The tool developed to deal with the ENTRO-OSI data is called GetENTRO. This tool, executable and source code, has been handed over to NBI-DSS staff. The execution of the tool is straightforward:

- Convert XLS file to CSV file
- Format CSV to:
    - Line 1: (ids according to PostGreSQL data base
        - station_name
        - station_lat
        - station_lon
        - time_step_id
        - source_id
        - units_id
        - value_type_id
        - parameter_typeid
        - FirstYear
        - LastYear
    - Line 2: dummy
    - Line 3: year , 12 monthly values
    - Line 4: etc.
- Run: GetENTRO [infile] [outfile]
    - e.g. GetENTRO AddisZemen.csv ENTRO.csv
- Output can be automatically imported in PostGreSQL using NBItool.exe

2.6.4 Rusumo, (Burundi/Rwanda/Tanzania: Kagera River)

In 2006 a study was completed with the title “A preliminary hydrological and sedimentological investigation of the Rusumo Falls Hydropower and Multipurpose Project in Burundi/Rwanda/Tanzania located on the Rwandan/Tanzanian border”. Data collected in the context of this study were stored in a Microsoft Access database. The following tables are relevant for the DSS and have been imported:

- Meteo: meteorological stations locations
- Precipitation: data with precipitation records (in mm per day)
- Stations: flow record station locations
- Readings: streamflow levels (in m)

Three steps have been taken to import these data into the database:

- Import the Microsoft Access database into a PostGreSQL database (using Navicat)
- Use a Query to get all data in CSV format
- Import the CSV data into the database using the NBItool.

2.6.5 SMMDSS

Sio-Malaba-Malakisi Decision Support System (SMM DSS) is a water resources planning and management software package developed by WREM International Inc. as part of the “Sio-Malaba-Malakisi Transboundary Integrated Water Resources Management and Development Project”.

Data from the SMMDSS have been converted to the PostGreSQL format. A total of 1,434,870 records were included in this database.

3 Climate Data

3.1 Overview

Climate data have been obtained from global as well as local data sources as described in the Inception Report. The following sources have been compiled and included in the database:

- GSOD
- SMMDSS
- Nile DST
- Ethiopian Master Plan
- ENTRO-OSI
- RUSUMO

3.2 Example of applications

For two stations in the vicinity of each other a typical application of data evaluation and quality control will be demonstrated here. The two stations selected are BUGARAMA_Aeronaut (id 113) and BUGARAMA_Commune (id 114) in Burundi and are about 2 km located from each other. Daily data from the two stations are available from 1971 to 1990.

One of the first issues to explore precipitation data is to plot these data. From Figure 1 it is clear that there are quite some differences in these two stations although they are only 2 kilometer apart. Especially the two rainfall events of about 120 mm of id113 are not shown at all for id114.

Second analysis is to evaluate total annual precipitation of the two stations considered. Figure 2 indicates that the stations are quite similar in annual precipitation. The year 1982 shows much lower rainfall for id114 compared Figure 3 to id113, but this can be explained by quite some missing data for station id113 in this year.

Finally, extreme analysis is especially for rainfall (and for streamflow as well) a good indicator of quality of data and similarity of stations. In Figure 3 return period of annual maxima is plotted as $Tr = (N+1)/m$, where m is the rank and N is the length of the record in years. From the figure it is obvious that station id113 can not be well fitted due to the two high values. Therefore, same figure has been plotted by ignoring the two high rainfall events.

From the analysis it might be quite sure that the two very high rainfall extremes might be contribute to errors in observations or recording. Comparing the two events with station id114 it is most likely that just an additional 1 has been added and that the real rainfall was exactly 100 mm lower for these two days.

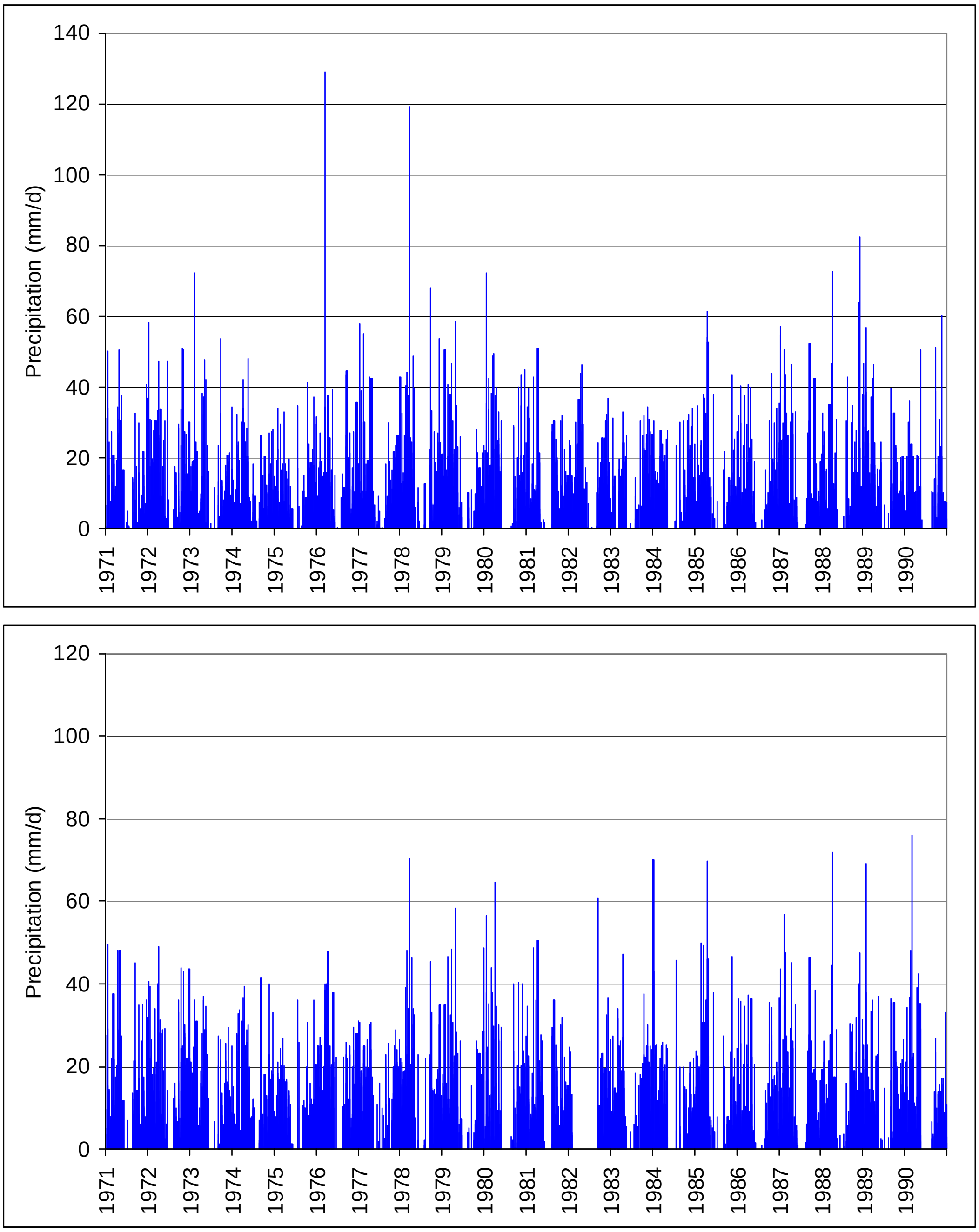


Figure 1. Daily precipitation for two stations id113 (top) and id114 (bottom) in Burundi.

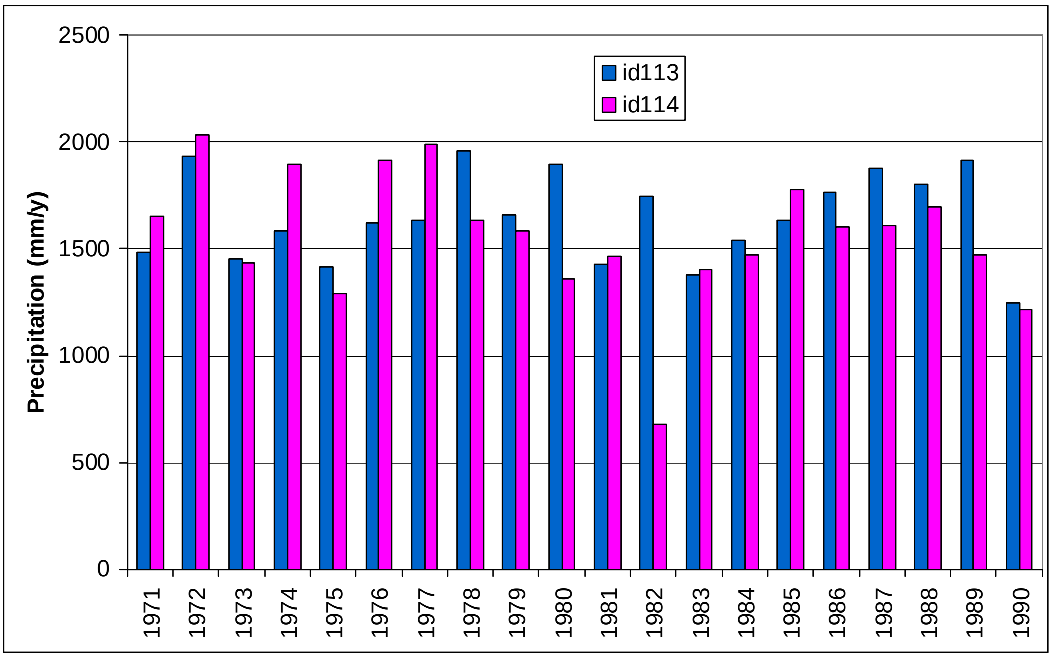


Figure 2. Annual precipitation for two stations id113 (top) and id114 (bottom) in Burundi.

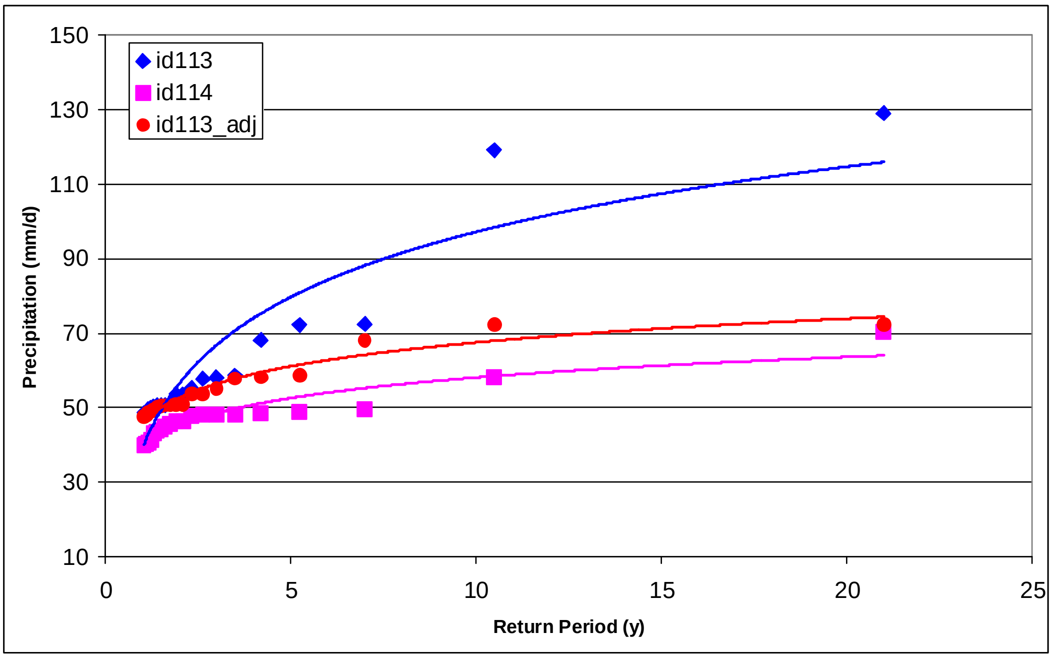


Figure 3. Extreme analysis for the two precipitation stations and the adjusted data.

4 Flow Data

4.1 Overview

Data sources for flow data have been defined in the Inception Report. Flow data (streamflow and water levels) have been obtained from global as well as local data sources as described in the Inception Report. The following sources have been compiled and included in the database:

- LEGOS, Satellite
- Nile DST
- RivDis
- GRDC

4.2 Example of applications

4.2.1 Monthly flow data

The Inception Report identified a few sources of global datasets on streamflow records. Most of these datasets included identical data and two datasets were identified as being unique:

- RivDis, River Discharge Data Base (Vörösmarty, et al. 1998) (http://www.rivdis.sr.unh.edu/)
- GRDC, Global Runoff Data Center (GRDC, 2009) (http://grdc.bafg.de)

The RivDis data base includes data of 22 stations in the Nile Basin (Map 1) and the GRDC data base has a total of 77 stations (Map 2). RivDis includes only monthly data, while from the 77 stations included in the GRDC 14 has daily data as well. As can be seen from the two Maps there is a big overlap in the two datasets. From the 22 RivDis stations 20 are identical to GRDC stations. Combining the two databases brings the total of discharge stations at 79.

Long term average streamflow from these 79 stations has been shown in Map 3. This map has to be considered cautiously since data might be from various periods and also record length might vary. From the Map it is also clear that the locations of some of the stations in Ethiopia are not correct.

Two stations are interesting to perform some preliminary data comparison and quality control. The records of Aswan and Khartoum are having a very long record, although not so much recent data are available (Aswan 1870-1984; Khartoum 1912-1982). Typical streamflow for three 10-years periods are plotted in Figure 4.

The first period, 1912-1921, shows streamflow from the earliest records of both stations and can be considered as the more or less natural flow hardly affected by human interference. Flows at Aswan were always somewhat higher than at Khartoum, due to inflow from . The second period, 1956-1965, reflects the situation just before completion of the Aswan dam in 1968. Flows at Khartoum are similar to the earlier years, but flows at Aswan show much higher peaks during the rainy season. This can be explained by the impact of changes in land cover, with less forests and other natural vegetation. Rainfall falling between Khartoum and Aswan became faster to runoff and less water evaporated. Finally the last 10-years period shows the impact of Aswan dam. Flows at Khartoum were fairly constant, although somewhat lower due to prolonged droughts. Flows at Aswan were however substantially regulated as a result of the High Aswan dam.

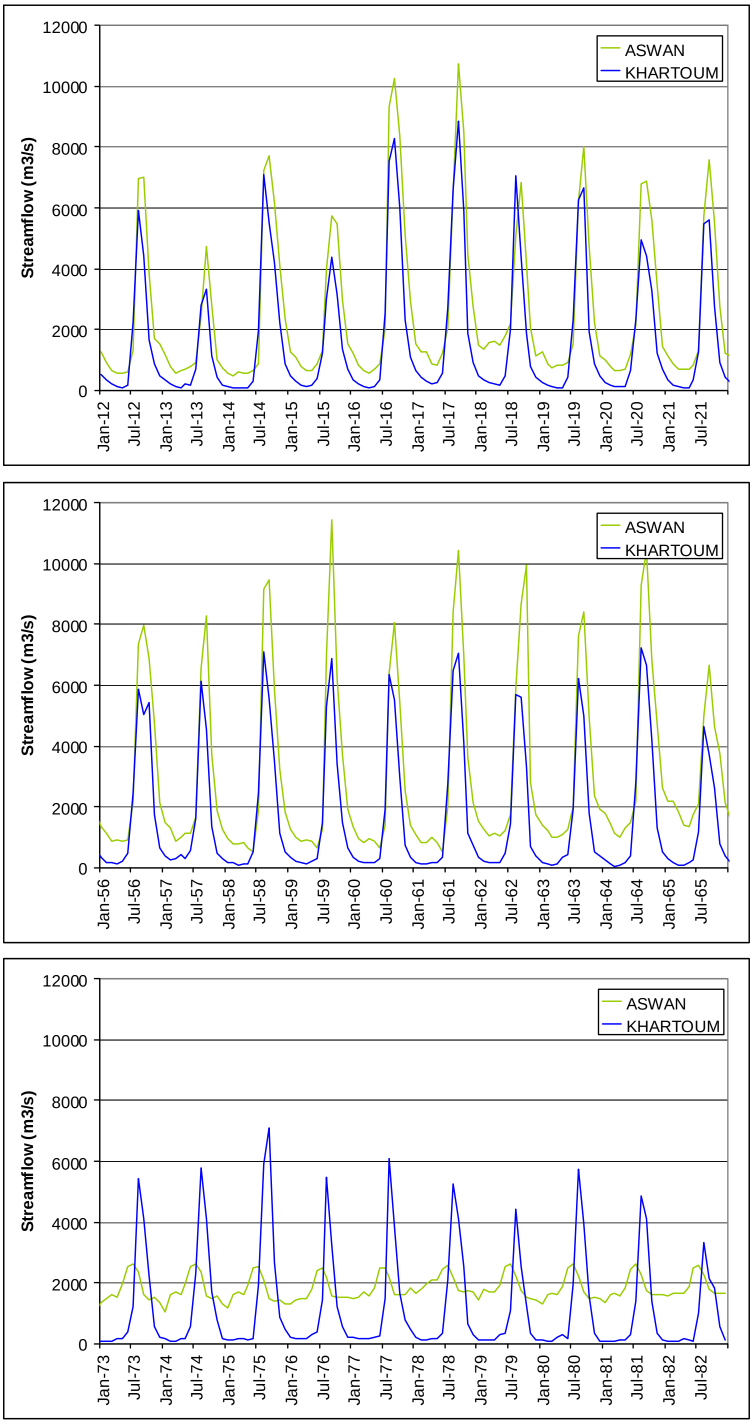


Figure 4. Monthly streamflow records for Aswan and Khartoum for three periods.

Annual discharges at Aswan show a small decreasing trend with a sudden drop after the completion of the High Aswan Dam. This drop can be explained that the location of the station changed in such a way that before 1965 total flow was recorded and after 1965 only releases from the High Aswan Dam were observed. Most interesting is to plot the difference between annual discharge at Khartoum and Aswan indicating the net inflow in this part of the Nile.

Distinct periods can be observed with relatively constant net inflows till 1950, followed by higher net-inflows as a consequence of land use changes. The prolonged droughts in the 1970's can be clearly observed in the graph and some recovering in the 1980's. The dataset is not sufficient recent to do more recent evaluations.

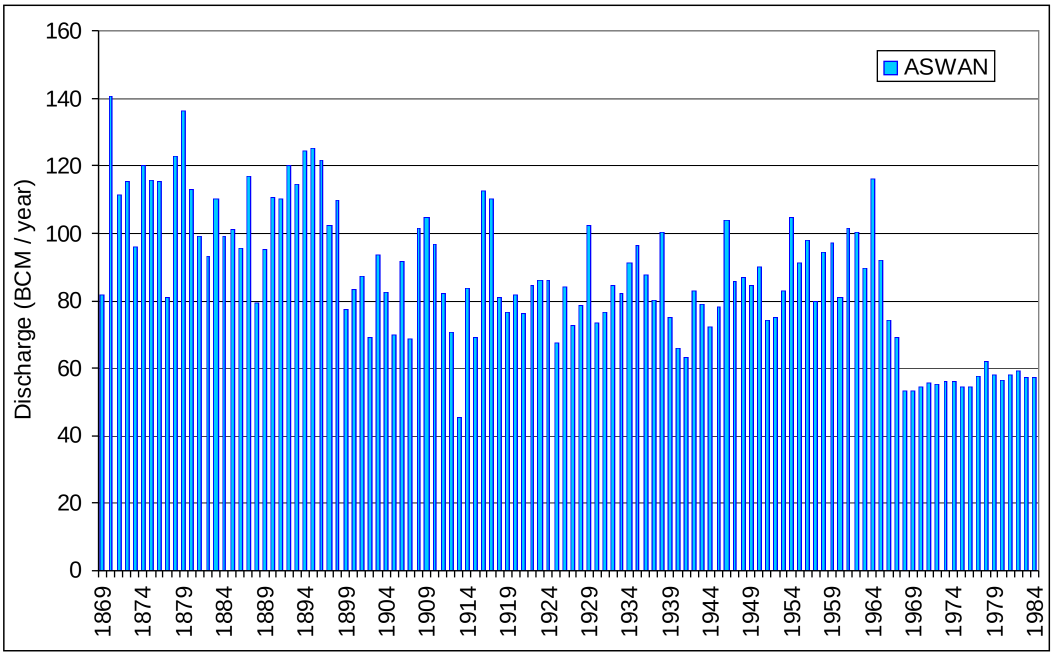


Figure 5. Annual discharge at Aswan.

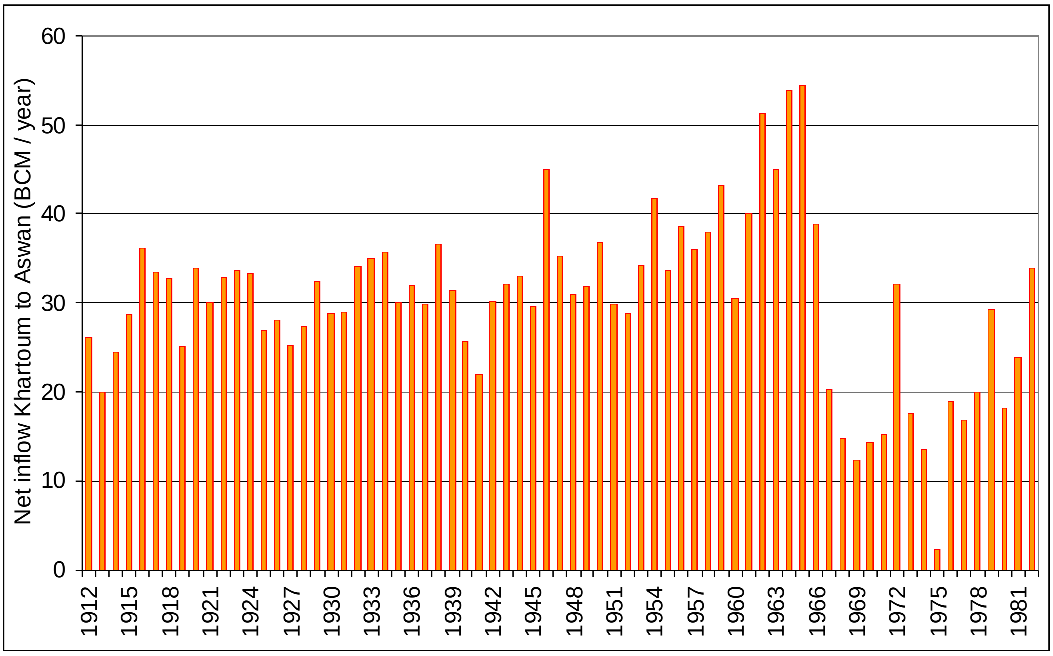


Figure 6. Annual net inflow between Khartoum (Blue Nile) and Aswan.

4.2.2 *Daily flow data*

There are 22 stations that have daily data. However for all these stations only data from 1978 to 1980 are available, making the data somewhat less suitable for the DSS. To get an impression on the data daily streamflow has been plotted in Figure 7 for a station at the Blue Nile in Ethiopia (GRDC 1563500) and for the Nyando river in Kenya (GRDC 1769100). The streamflow characteristic is very different for the two stations. The daily data from the Blue Nile stations show a clear seasonal pattern reflecting the rainfall. The station in Kenya however shows a much more erratic behavior. It is clear that the modelling part of the DSS can benefit greatly from calibration and validation on daily streamflow data in addition to the monthly data.

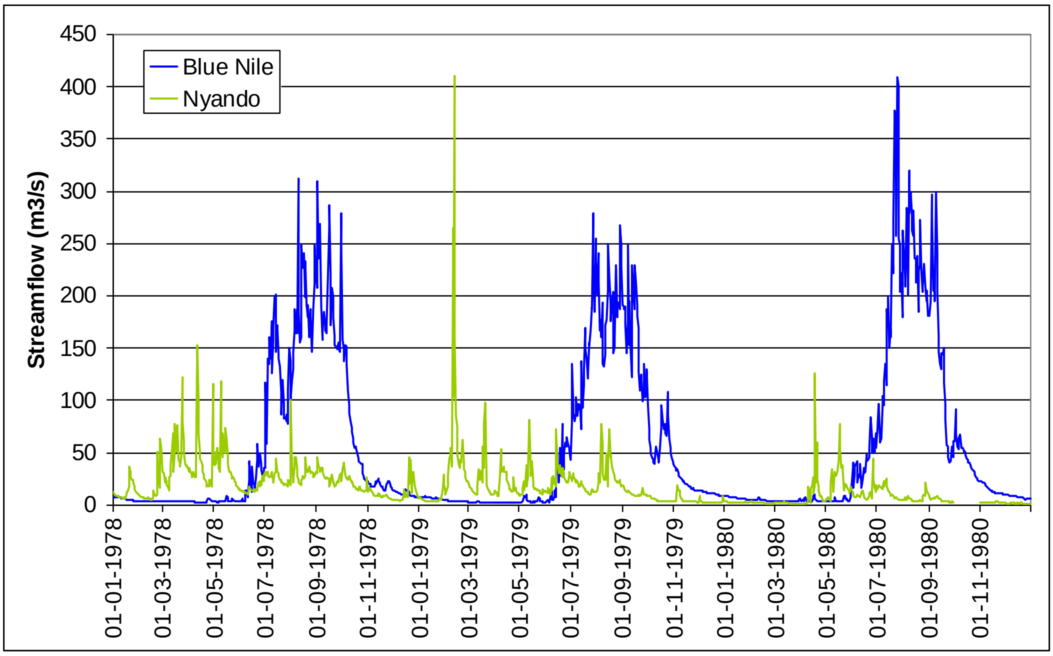


Figure 7. Example of daily streamflow data from GRDC.

4.2.3 *Satellite based water levels*

Satellite radar altimeters can potentially monitor the variation of surface water height for many large inland water bodies (lake, wetland or river) Topex/Poseidon measurement system. The Topex/Poseidon satellite altimeter was launched on 10 August 1992. The Jason-1 satellite which was launched in December, 2001. Every 10 days, the satellite orbiting at an altitude 1300 km provides a complete coverage of the ocean surface topography with unequalled precision: within 5 cm for an instantaneous measurement and within 2 cm for a monthly average.
Both satellites have been used to observe lake levels and levels or large rivers.

More background information can be found at :

- http://www.legos.obs-mip.fr/en/observations/ctoh/altimetres/topex, and
- http://www.pecad.fas.usda.gov/cropexplorer/global_reservoir/index.cfm

while actual data can be downloaded from:

- http://www.legos.obs-mip.fr/en/soa/hydrologie/hydroweb/

In total 32 water level observation points are available (Map 5). Of these 32 points 8 are reflecting lake or reservoir levels and 24 reflect Nile levels. A typical example of the data is presented in Figure 8 for two lakes. These data can be of very valuable input for the DSS and the associated modeling framework.

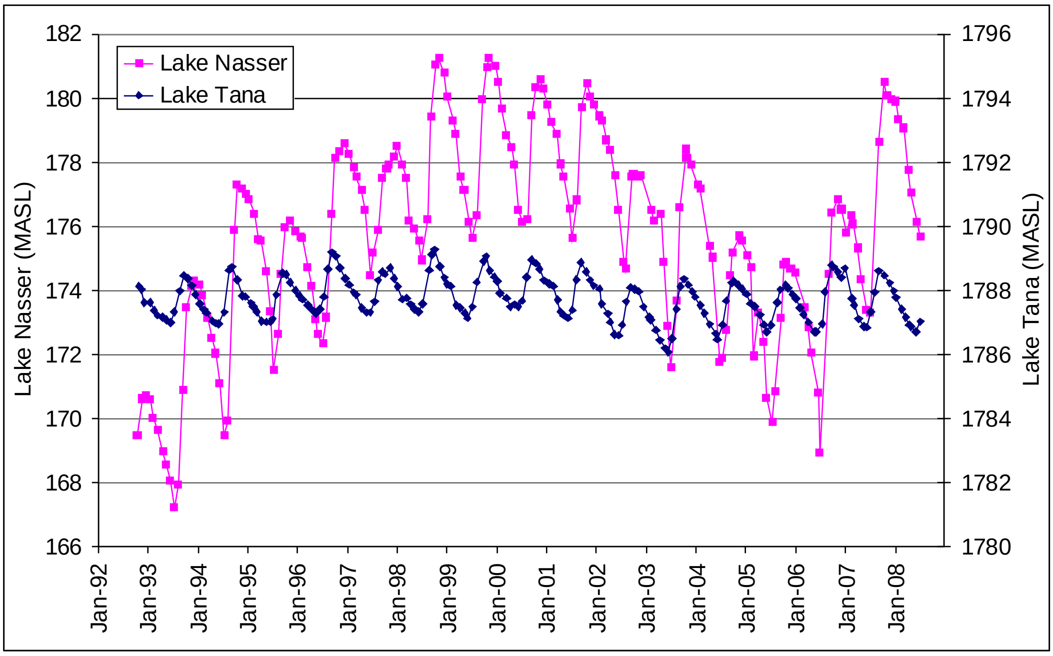


Figure 8. Example of water levels from the Topex and Jason satellites.

5 Spatial Data

5.1 Digital Elevation Models

5.1.1 Compilation and structuring

Digital elevation data are required for many DSS processing like defining watershed boundaries, streamflow delineation, streamflow directions, etc. The ultimate source for these data at high resolution is the SRTM 90m. However given the large size of this dataset of over 2 GB (32 tiles of each 72 MB) covering the Nile Basin, it would be difficult to handle these for the entire area. Therefore also the lower resolution Hydro1K dataset has been included in the data base.

Further details of compilation and structuring are already provided in the Inception Report.

5.1.2 Processing

The processing of the SRTM data included two steps. Although the provided GeoTiff could be displayed directly in ArcGis, projecting the data and other processing is not possible. Therefore, data have been converted to native ArcGis GRID format. Second step consists out of setting the projection. For the SRTM this is LatLon with WGS84.

Processing of the 1km Hydro 1k included importing the .bil data and as second step to define the projection as provided in the documentation with the HYDRO1k: Lambert Azimuthal Equal Area

5.1.3 Quality control

Specific data control has been undertake already on the two data sets. Details can be found in:

- Jarvis et al., 2008
- CIAT, 2008
- Reuter, et al., 2007.

5.1.4 Organizing and archiving

References of the two DEM datasets have been added to the Meta Data Base. No additional projection has been applied to the two data sets, but can be done easily ones the final projection for the DSS has been defined.

5.1.5 Example of applications

The main application type of a DEM in the DSS will be the delineation of the basin and sub-basins and the definition of the streamflow network. To undertake this for the entire Nile Basin is a major achievement, but here an example will be provided. Moreover, the difference in the two DEMs, HYDRO1k and SRTM will be demonstrated as well.

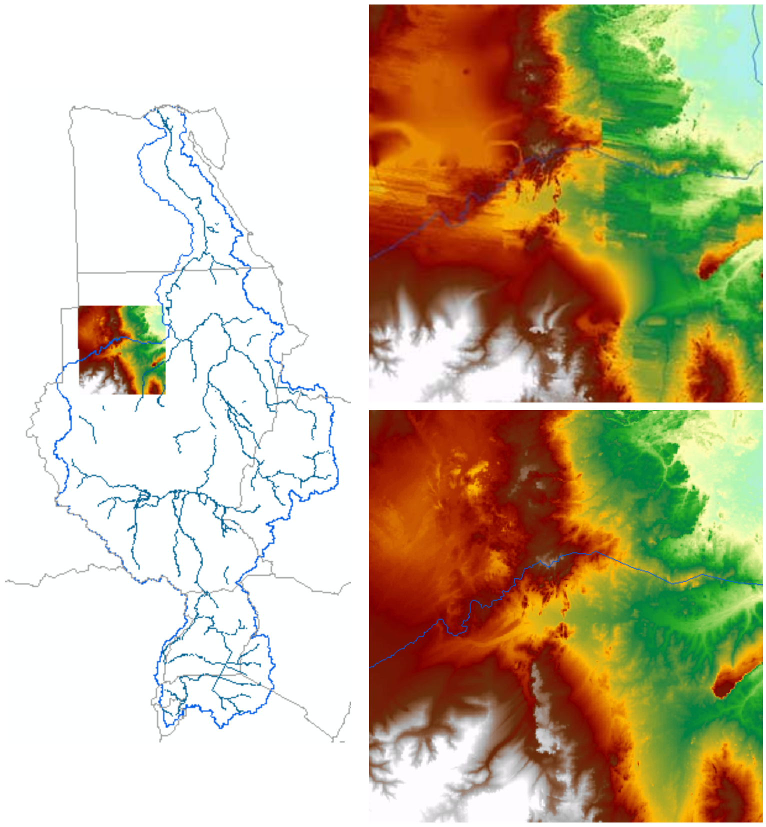


Figure 9. Example of the HYDRO1K (top right) and SRTM (bottom right) DEM for north-west Sudan.

Figure 9 demonstrates for one SRTM tile (42_09) in North-West Sudan the difference between SRTM and HYDRO1K. It is clear that the HYDRO1K has a lot of deficiencies which are not only based on the difference resolution (1000 m vs. 90 m). There are many artifacts in the DEM which will cause huge problems if these data will be used to delineate outline of the basin, sub-basins and streamflow network.

Delineation of the outline and the streamflow network using high resolution DEM require large computational time and memory resources. In many cases software might crash as memory allocation is not sufficient (ArcGis).

The effect of resolution and the validation of the existing Nile Basin Outline was demonstrated by using the pre-processing software of SWAT on the SRTM tile. The original 90 m was resampled to 300 m and 500 m and the watershed boundary was determined for this tile.

In Figure 10 the result of this processing is seen at three levels of detail. At the highest level it can be concluded that the existing outline follows the one determined using the SRTM very well.

5.2 Soils

5.2.1 Compilation and structuring

The Inception Report concluded that regarding soil data the Harmonized World Soil Database can be considered as best to be used in the Nile-DSS. This raster database consist out of over 15000 different soil mapping units that combines existing regional and national updates of soil information worldwide (SOTER, ESD, Soil Map of China, WISE). Further details of compilation and structuring are already provided in the Inception Report.

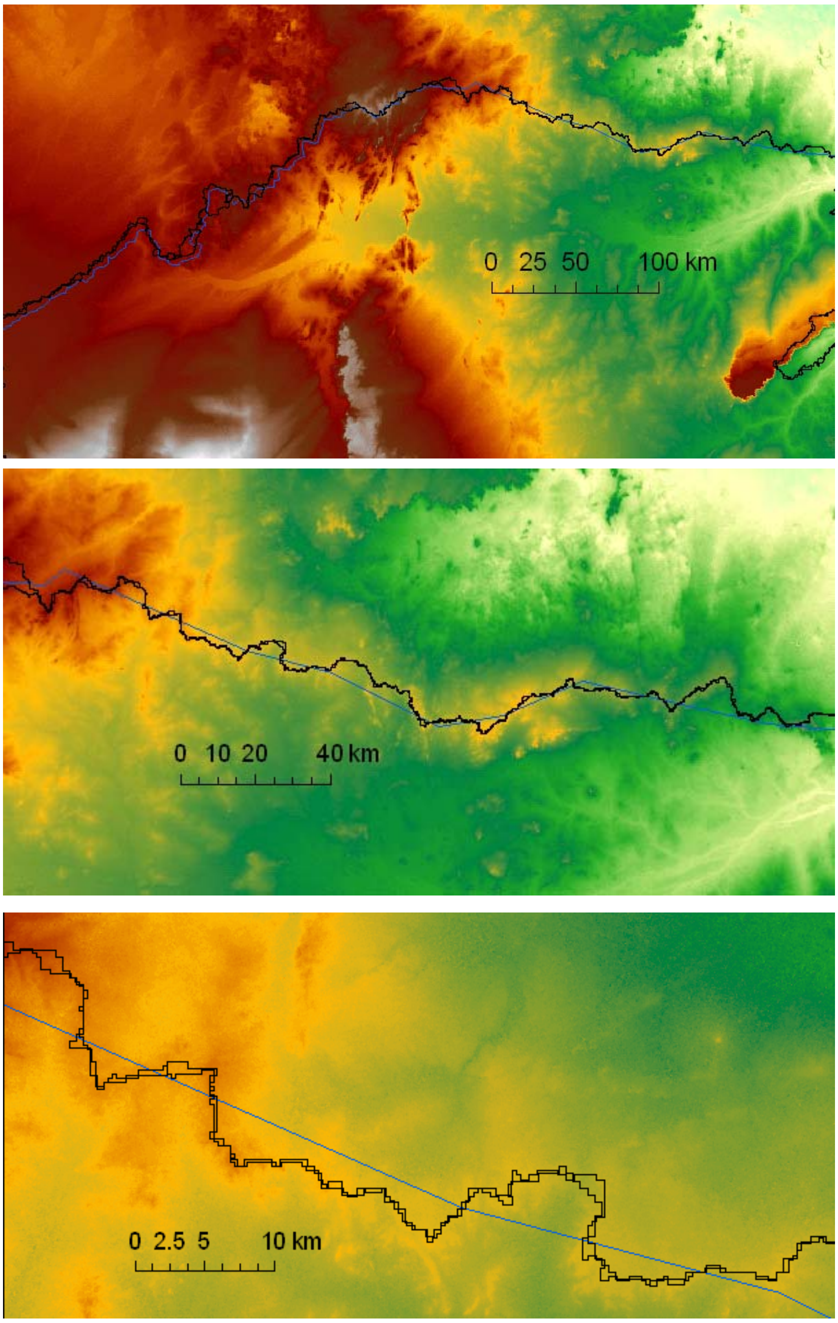


Figure 10. Outline created using resampled SRTM of 300m and 500m (in black) compared to original Nile outline (in blue).

5.2.2 Processing

The Harmonized World Soil Database is stored in a 30 arc-second raster database with over 15000 different soil mapping units. The raster database consists of 21600 rows and 43200 columns, which are linked to harmonized soil property data in MS Access. 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.

During the processing phase the global database was clipped to the Nile Basin extent (6S-33N, 22E-41E) and exported to ArcGis GRID format. This grid was joined with the actual data in MS-Access. The Value attribute in the GRID was joined with the table HWSD_DATA in the HWSD.mdb MS-Access file. Field to join is MU_GLOBAL. In summary

ArcGis Grid > HWSD_Nile > Attributes > Value
joins
MS-Access > HWSD.mdb > HWSD_DATA > MU_GLOBAL

Based on this join the entire soil attributed can be shown and used. For the DSS the most relevant information is:

- T_TEXTURE
    - Topsoil Texture
- REF_DEPTH
    - Reference Soil Depth
- DRAINAGE
    - Drainage class
- AWC_CLASS
    - AWC Range
- ROOTS
    - Obstacles to Roots
- IL
    - Impermeable Layer
- SWR
    - Soil Water Regime
- T_USDA_TEX_CLASS
    - Topsoil USDA Texture Classification

5.2.3 Quality control

Quality control for the entire global soil data base has been described by (XXX). It was concluded that reliability of the information contained in the database is variable: the parts of the database that still make use of the Soil Map of the World such as North America, Australia, West Africa and South Asia are considered less reliable, while most of the areas covered by SOTER databases are considered to have the highest reliability (Southern Africa, Latin America and the Caribbean, Central and Eastern Europe).

5.2.4 Organizing and archiving

Reference to the soil dataset has been added to the Meta Data Base. No additional change in the projection has been applied, but can be done easily ones the final projection for the DSS has been defined.

5.2.5 Example of applications

The DSS could be very helpful in future land use planning processes regarding issues like forestation, irrigation, etc. An important aspect is the options the soil offer to changes in land cover. In Figure 11 two important soil characteristics are plotted: the available water content and the drainage situation. Obviously, there is some correlation between those two where poorer drainage is correlated to higher available water contents. However, many other factors play an important role in the available water content like soil texture and soil depth.

To demonstrate the use of the soils data a simple analysis has been undertaken to obtain areas where irrigation might be potentially an option from a soils perspective, ignoring water availability, climate and socio-economic factors for the moment. Optimal soil conditions for irrigation are defined as areas with a high soil water content and a low natural drainage (assuming that water logging and salinisation can be ignored). Combining these two factors a map of potential interesting area to explore irrigation has been derived (Figure 12).

Note that this example is only provided to demonstrate the use of the Harmonized World Soil Database. To derive the real potential of irrigated areas, more components should be included like slopes, water availability, markets etc.

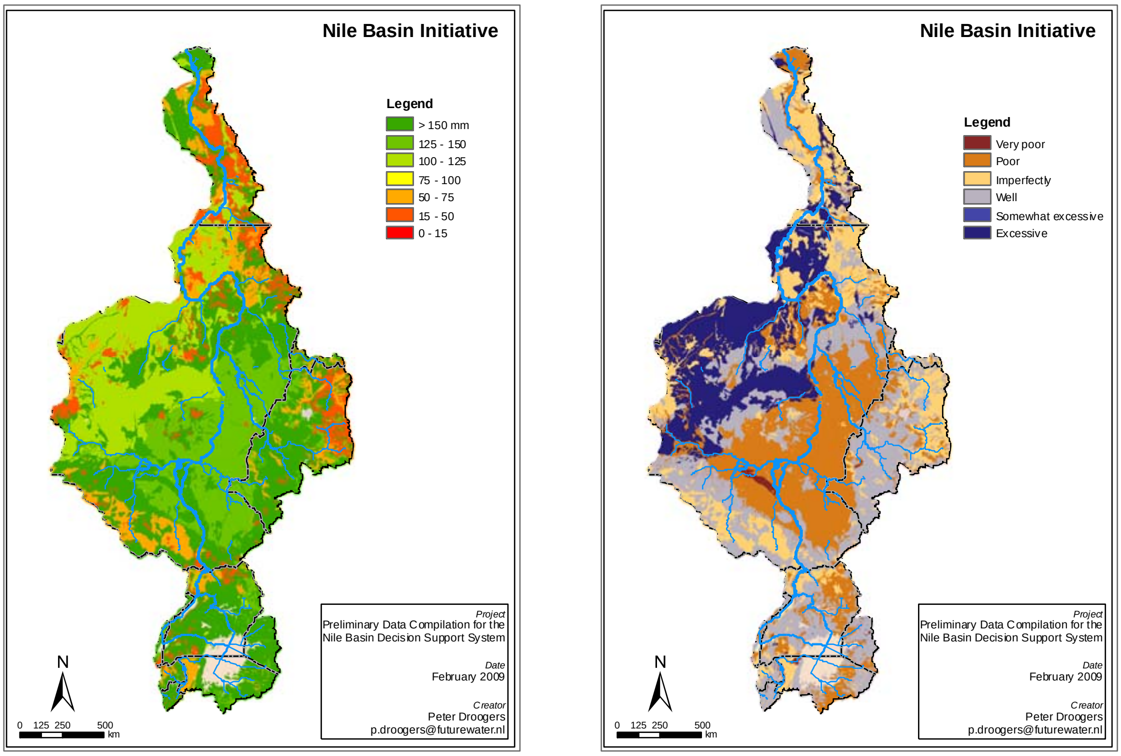


Figure 11. Two important soil properties: available water content (left) and natural drainage conditions (right).

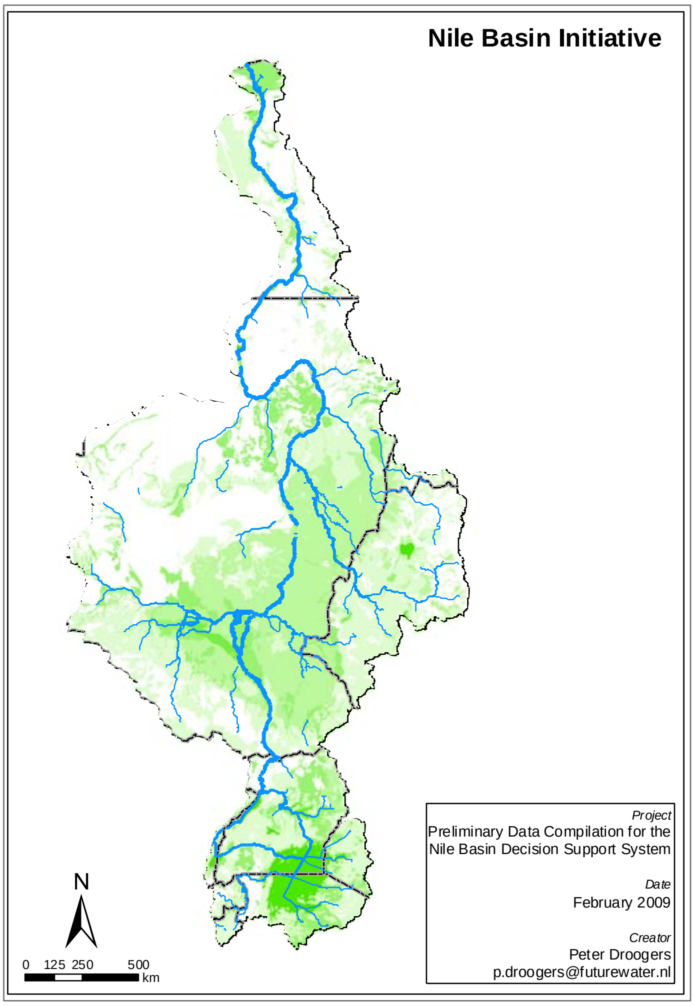


Figure 12. Areas having a high potential for irrigation from a soils perspective. Water bodies to be filtered out.

5.3 Land cover / land use

5.3.1 Compilation and structuring

In the Inception Report it was concluded 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.

Further details of compilation and structuring are already provided in the Inception Report.

5.3.2 Processing

The GLC2000 land cover data set has been obtained from the Joint Research Centre of the European Commission. The dataset is provided as a ArcGis raster with 9187 rows and 9717 columns covering the entire continent. Grid size is 0.00892857 degrees (~ 1 km). The dataset original projection is LatLon WGS1984.

5.3.3 Quality control

Quality control of large scale land cover maps is always difficult. Low scale regional maps pose problems of validation. The scale and the extent of such documents preclude any classical ground-truthing at an acceptable cost. Mayaux et al. (2004) have undertaken an initial quality control based on various techniques. The GLC2000 map was compared with different data sets in order to gain a better understanding of the thematic content and the spatial detail of the classes:

- the 18 TREES-II Landsat interpretations (Achard et al., 2002),
- the Terrestrial Ecoregions of the World Wildlife Foundation (Olson et al., 2001),
- the Vegetation Continuous Fields (VCF) derived from Moderate Resolution Imaging Spectroradiometer (MODIS) data (Hansen et al., 2003) and
- 20 Landsat scenes were randomly selected in a range of Africa's ecoregions for a visual comparison. This complies with the map validation method adopted by IGBP for the global probability-based sample (Scepan, 1999).

It was concluded that the quality of the GLC2000 data set is acceptable to represent large scale features, but for small scale features local information should be included.

5.3.4 Organizing and archiving

The GLC2000 data set has been added to the Meta Data Base. No additional change in the projection has been applied, but can be done easily ones the final projection for the DSS has been defined.

5.3.5 Example of applications

To indicate the level of detail that can be obtained by the GLC2000 data set, a detailed map of the Nile Delta and the area around Lake Victoria has been plotted in Figure 14. Around Lake Victoria dominant land cover is croplands with some patches of forests and shrublands. The data set for the Nile Delta shows a completely different picture. Obviously, irrigated agriculture is the main land cover in the delta. However, around the irrigated agriculture the GLC2000 indicates the present of large scale tree crops. In reality these are areas where irrigations is partly practiced. It is clear that this is an error in the classification technique where most likely this partly green area has been considered to be trees.

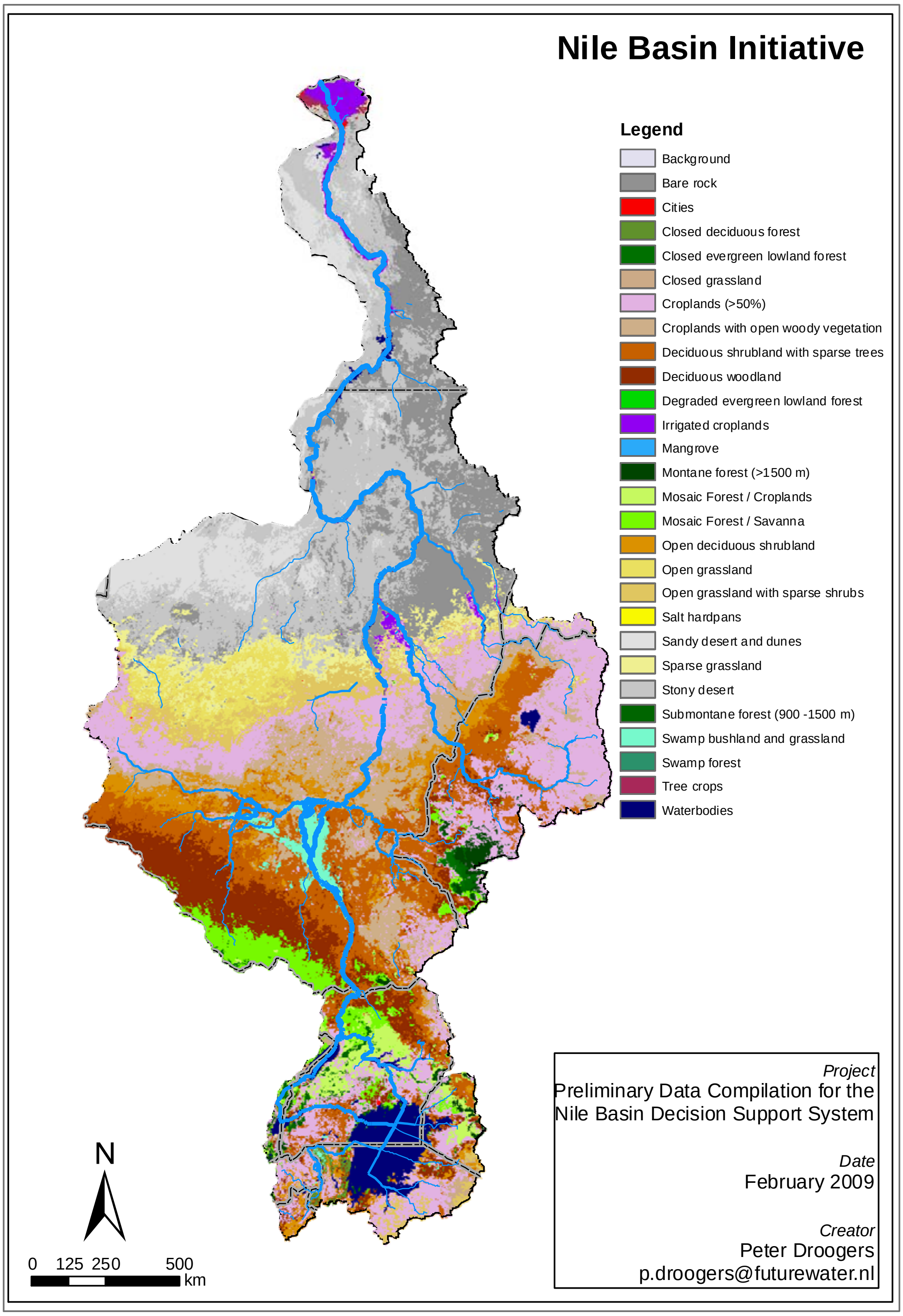


Figure 13. Land cover map.

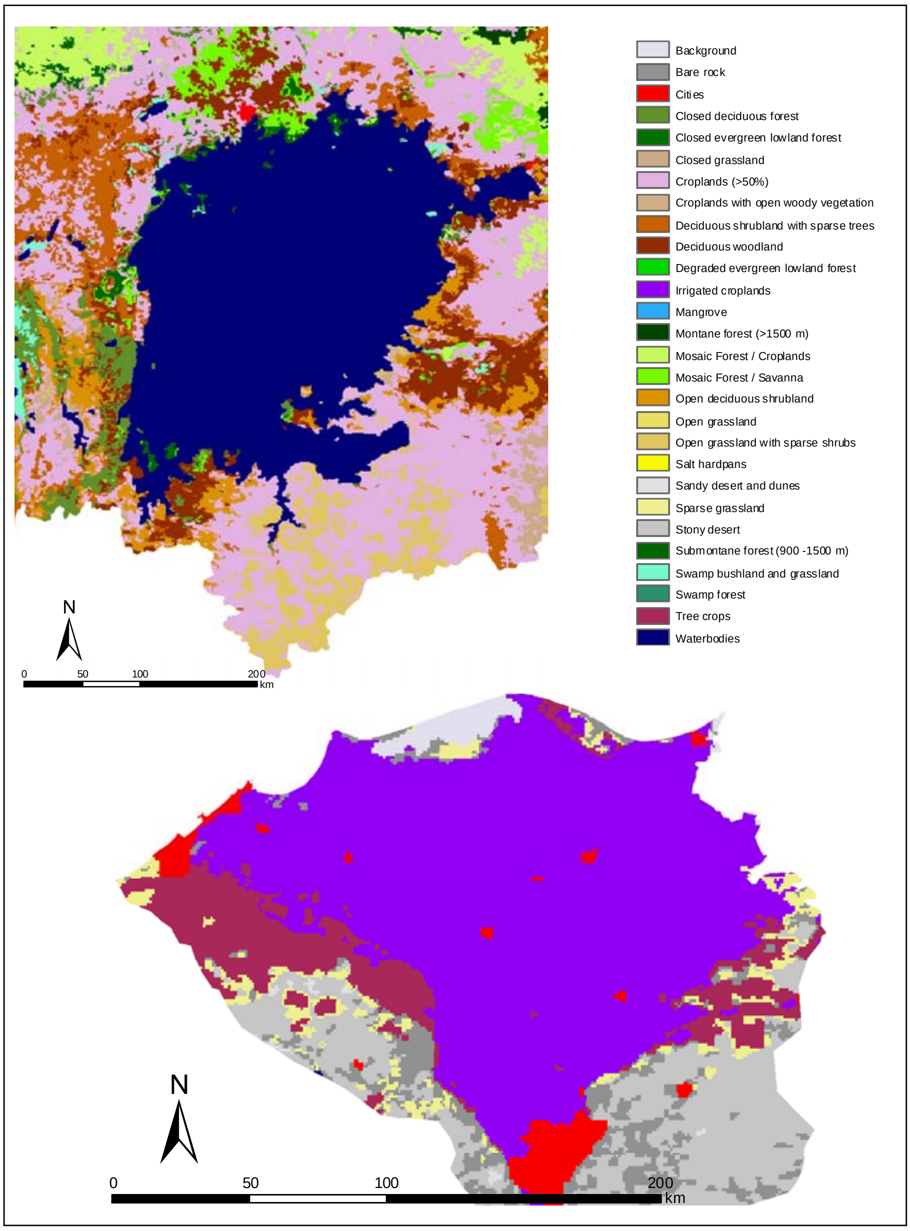


Figure 14. Details of the LCC2000 land cover map for Lake Victoria Area (top) and Nile Delta (top).

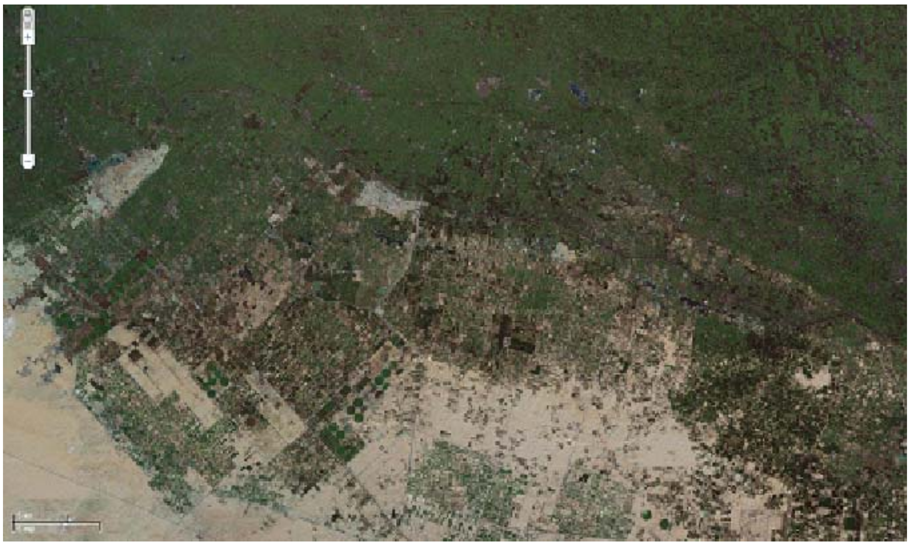


Figure 15. Nile delta showing areas fully irrigated in the North and partly irrigated in the South (source Google Maps).

5.4 Vegetation Indices

5.4.1 Compilation and structuring

In the Inception Report it was recommended to use MODIS images as vegetation indices. The advantage of MODIS is that data are free and derived products are available such as LAI (Leaf Area Index) and (Net Primary Production). Various sites are available for downloading data. The most user friendly and effective one is GloVis, the USGS Global Visualization Viewer.

5.4.2 Processing

The following three products are relevant for the DSS and can be nearly real-time downloaded:

- NDVI MODIS 13A3
- LAI MODIS 15A2
- NPP MODIS 17A3

All these files are stores in HDF format. Processing of files can be done straightforward in ArcGis version 9.3, by importing the file and selecting the appropriate layer. Older versions require a conversion tool from HDF to ArcGis raster using MGET[1].

The original files are in stored in the sinusoidal projection, which is the preferred one for the spatial data base. However, the original MODIS files use no false easting while for the Nile Basin a false easting of XXX is optimal.

[1]http://code.env.duke.edu/projects/mget

5.4.3 Quality control

Quality control of vegetation indices at larger scales is not straightforward. MODIS products are constantly monitored on quite strict quality control (Huete, 1999). The quality of all MODIS vegetation index (VI) products are assessed regularly to monitor the accuracy and integrity of the data products over long time frames that are of critical importance to studies of global change and climate. Run time and postproduction QC evaluations consists of:

- In depth QC analysis of 1 land tile per composite period requiring access to all input files (chosen from the validation site list). These QC activities will be mostly interactive.
- Continuous QC and trend analysis for 50 validation sites using subsets of up to 50 land tiles and the 16-day composited VI products. These QC activities will be mostly automated.
- QC analysis of spatially continuous data on a global scale (4 times per year) using 16-day composites.
- Ad hoc QC analysis of land tiles based on science group findings.

Details of the entire quality control are provided by Huete et al. 1999.

5.4.4 Organizing and archiving

The MODIS data set has been added to the Meta Data Base. Not all data has been actually included in the directory mentioned as this will take an enormous amount of space. Reference is therefore made to the online data base from GloVis. Bulk downloading can be done easily by using FTP:

- ftp://e4ftl01u.ecs.nasa.gov/MOLT/MOD13A3.005/

The following tiles are required to cover entire Nile Basin:

- 20/5
- 20/6
- 20/7
- 20/8
- 20/9
- 21/6
- 21/7
- 21/8
- 21/9

20v05;*20v06*;*20v07*;*20v08*;*20v09*;*21v06*;*21v07*;*21v08*;*21v09*
Each Tile is about 20MB. Total disk space required for the full dataset is 17 GB (20 MB * 9 tiles * 8 years * 12 months)

No additional change in the projection has been applied, but can be done easily ones the final projection for the DSS has been defined.

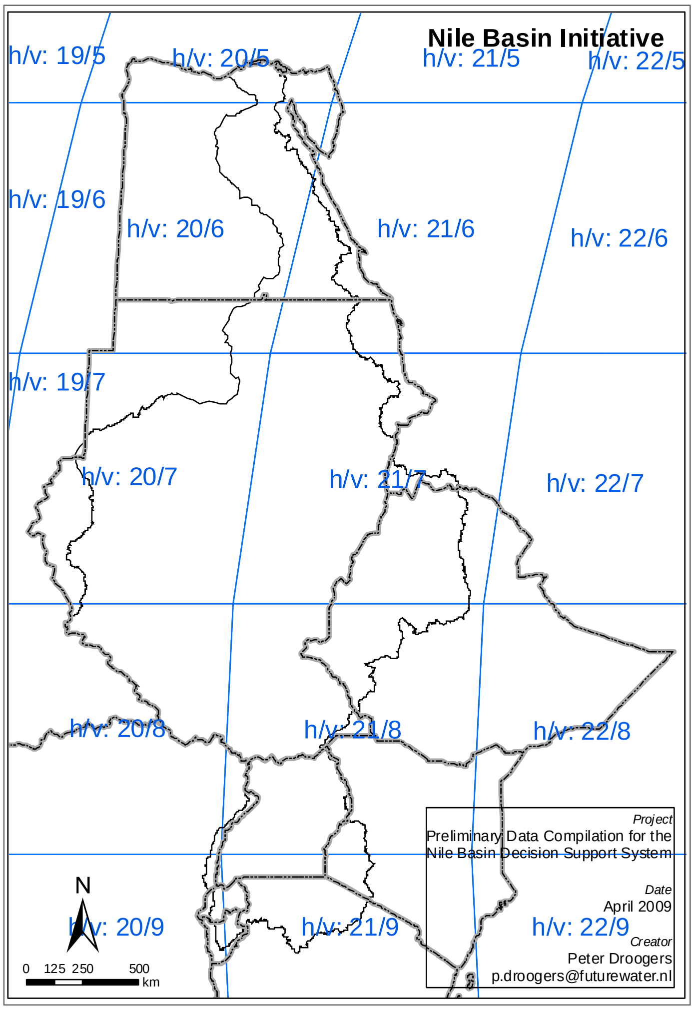


Figure 16. MODIS Tile numbers.

5.4.5 Example of applications

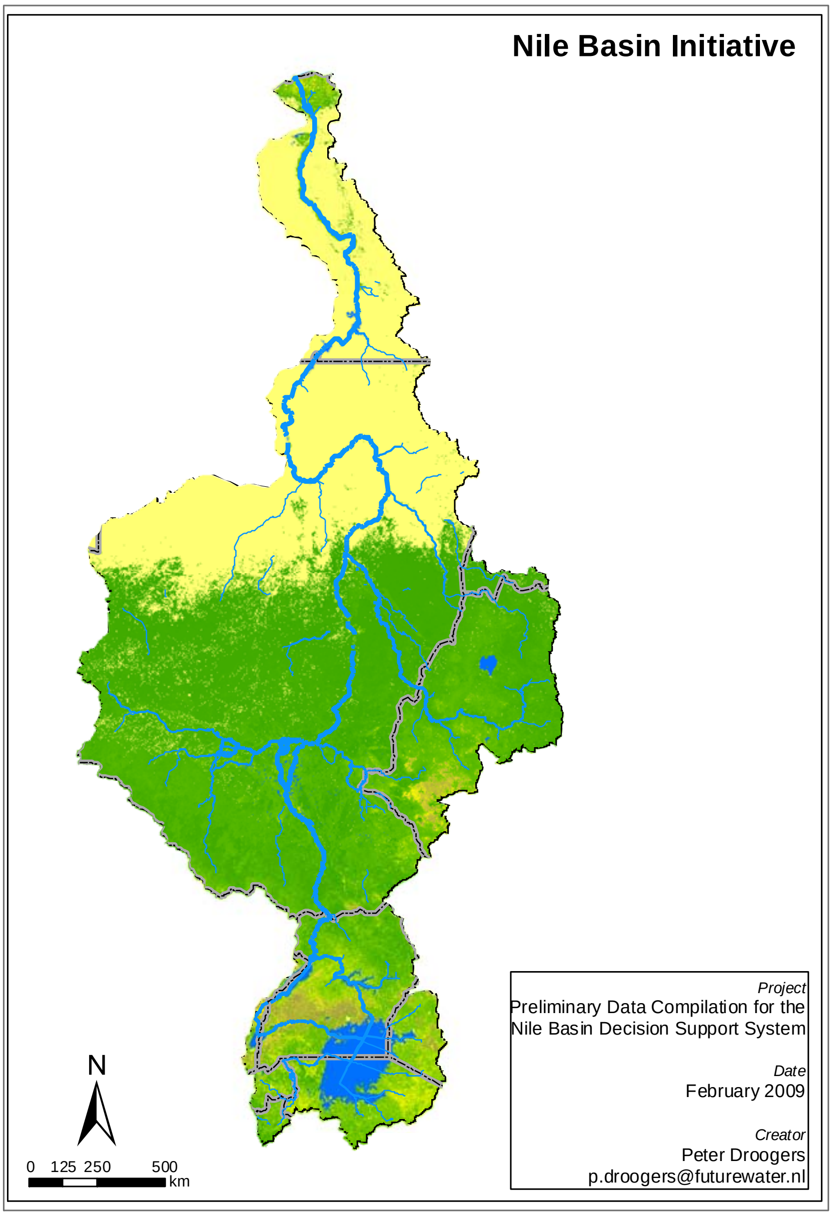


Figure 17. Example of MODIS LAI for January 2007.

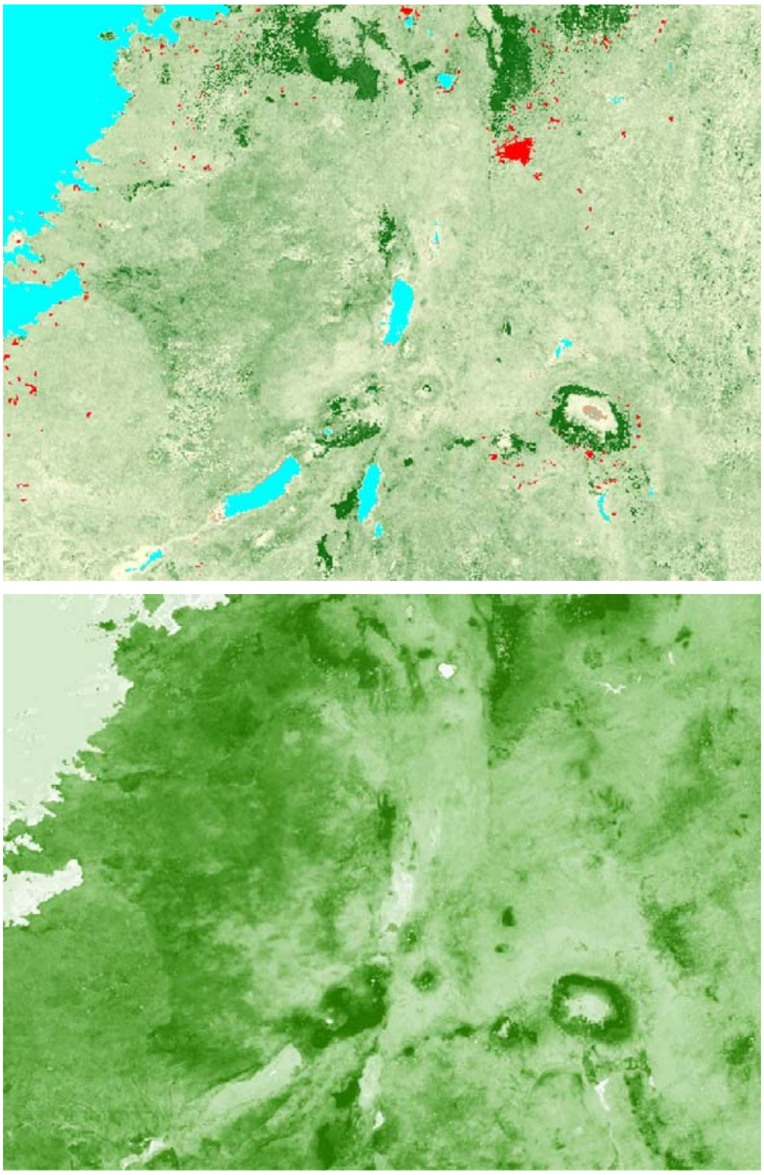


Figure 18. Detail of MODIS LAI (top) and NDVI (bottom) for the area east of Lake Victoria.

5.5 FEWS Rainfall

5.5.1 Compilation and structuring

In the Inception Report it was recommended to use TRMM data as gridded rainfall product. There is however a growing concern on the accuracy over Africa of the TRMM product, and it was therefore decided to use the FEWS-RFE dataset (Famine Early Warning Systems-Daily Rainfall Estimate). The FEWS RFE 2.0 algorithm has been implemented by NOAA's Climate Prediction Center, and uses an interpolation method to combine Meteosat and Global Telecommunication System (GTS) data, and included warm cloud information for the decadal estimates.

More information regarding FEWS can be obtained from:
http://earlywarning.usgs.gov/adds/overview.php

5.5.2 Processing

Data has been downloaded from: http://earlywarning.usgs.gov/adds/dwndailyrfe.php. Data can be downloaded as annual collections, daily collections or individual daily data.

The Year Collection of daily rainfall images are ~25 Mb per file (year). The naming convention is rain_YYYY.tar.gz where YYYY is the 4 digit year. These are UNIX .tar.gz files containing a years daily (365 or 366 in leap year) data files.

The Month Collection of daily rainfall images are ~ 2 Mb per file (month). The naming convention is rain_YYYYMM.tar.gz where YYYY is the 4 digit year and MM is the 2 digit month of year. These are UNIX .tar.gz files containing a months daily (28-31) data files.

The individual day rainfall images are ~65-80 Kb per file (day). The naming convention is rain_YYYYJDA.tar.gz where YYYY is the 4 digit year and JDA is the 1-3 digit Julian day of the year. Each daily rainfall file whether by itself or within the month or year collections is a UNIX .tar.gz file which contains the ARC/INFO created .bil image and related files (i.e. .bil, .blw, .hdr, .stx, .clr) for that day.

The individual days are stored in a ARC/INFO created .bil image and related files (i.e. .bil, .blw, .hdr, .stx, .clr) for that day. Data is in 8-km resolution, and in Conical Albert Equal Area projection.

5.5.3 Quality control

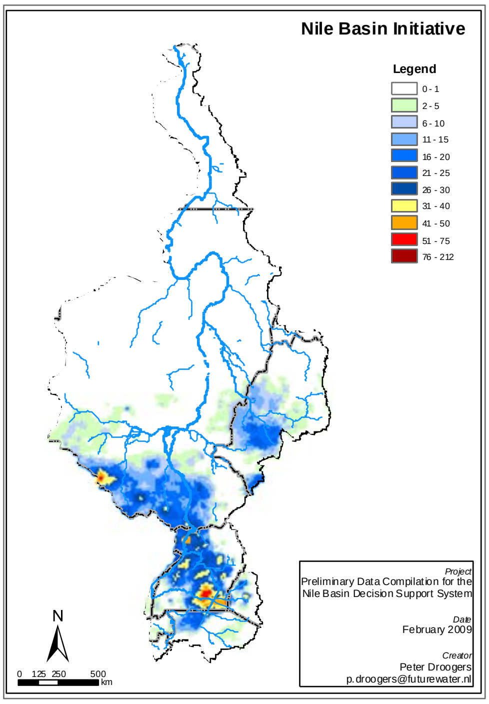
Quality control of the FEWS-RFE algorithm has been described by Herman et al. (1997) and by Xie and Arkin (1997). Although every grid based precipitation product should be always applied with care, the FEWS product has proven to provide reliable products to be included in the DSS. Validation is described by various authors such as Laws et al. (2004) and Love et al. (2004).

5.5.4 Organizing and archiving

The FEWS-RFE data sets have been added to the Supplemental Data Base. Data from 2006 to 2008 has been obtained and included in the directory mentioned.

5.5.5 Example of applications

A typical example of the use of the spatial rainfall estimates based on FEWS can be seen in Figure 19. It is clear that the high spatial resolution can never be met using conventional rainfall gauging stations.



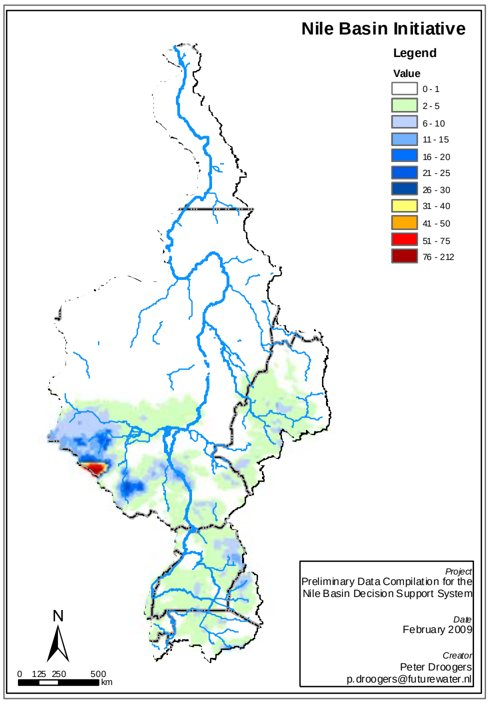


Figure 19. Precipitation (mm/d) according to FEWS for 1-Jun-2008 (left) and 2-Jun-2008 (right)

6 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 build 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 Analysis Phase of this activity which will be followed up by the Synthesis Phase.

Main activities conducted during this Analysis Phase and described in this report are:

- Global data collection
- Local data collection
- Developing PostGreSQL database
- Data processing
- Demonstrating relevant hydrological parameters derived from data

Main activities to be undertaken during the Synthesis Phase are:

- Finalize and complete the database:
    - Adding additional local data.
    - Add attributes required for modeling purposes to the soils, vegetation and land use maps.
- Complete the final report which will provide a clear description of the final database with the Analysis Report attached as background reference.

7 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

CIAT. 2008. SRTM DATA VERSION 4. CIAT-CSI SRTM website (http://srtm.csi.cgiar.org)

Droogers, P, and W. Immerzeel. 2009. Preliminary Data Compilation for the Nile Basin Decision Support System. INCEPTION REPORT. Report NBI/FutureWater.

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

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

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

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.

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

Fischer, G., F. Nachtergaele, S. Prieler, H.T. van Velthuizen, L. Verelst, D. Wiberg, 2008. Global Agro-ecological Zones Assessment for Agriculture (GAEZ 2008). IIASA, Laxenburg, Austria and FAO, Rome, Italy.

GRDC. 2009. The Global Runoff Data Centre, 5606X Koblenz, Germany.

Herman, A, Vadlamani B. Kumar, Phillip A. Arkin, Jamie V. Kousky. 1997. Objectively determined 10 day African rainfall estimates created for famine early warning systems. http://earlywarning.usgs.gov/adds/readme.php?symbol=rf

Huete, A., C. Justice. 1999. MODIS Vegetation Index (mod 13), algorithm theoretical basis document Version 3.

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

Jarvis A., H.I. Reuter, A. Nelson, E. Guevara, 2008, Hole-filled seamless SRTM data V4, International Centre for Tropical Agriculture (CIAT), available from http://srtm.csi.cgiar.org.

Kim, D., Edward Tollerud, Steve Vasiloff, Jason Caldwell. 2006. Comparison of Manual and Automated Quality Control of Operational Hourly Precipitation Data of the National Weather Service. http://ams.confex.com/ams/pdfpapers/146869.pdf

Kondragunta, C.R., K. Shrestha. 2006. Automated Real-Time Operational Rain Gauge Quality-Control Tools in NWS Hydrologic Operations.

Laws, K.B., John E. Janowiak, George J. Huffman. 2004. VERIFICATION OF RAINFALL ESTIMATES OVER AFRICA USING RFE, NASA MPA-RT, AND CMORPH. AMS Hydrology conference. http://www.cpc.ncep.noaa.gov/products/fews/rfe.shtml

Love, T.B., Vadlamani Kumar, PingPing Xie, Wassila Thiaw. 2004. A 20-YEAR DAILY AFRICA PRECIPITATION CLIMATOLOGY USING SATELLITE AND GAUGE DATA. AMS Hydrology conference. http://www.cpc.ncep.noaa.gov/products/fews/rfe.shtml

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. Bartholome, S. Fritz, A. Belward. 2004. A new land-cover map of Africa for the year 2000. Journal of Biogeography 31, 861–877.

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

Reuter H.I, A. Nelson, A. Jarvis, 2007, An evaluation of void filling interpolation methods for SRTM data, International Journal of Geographic Information Science, 21:9, 983-1008.

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.

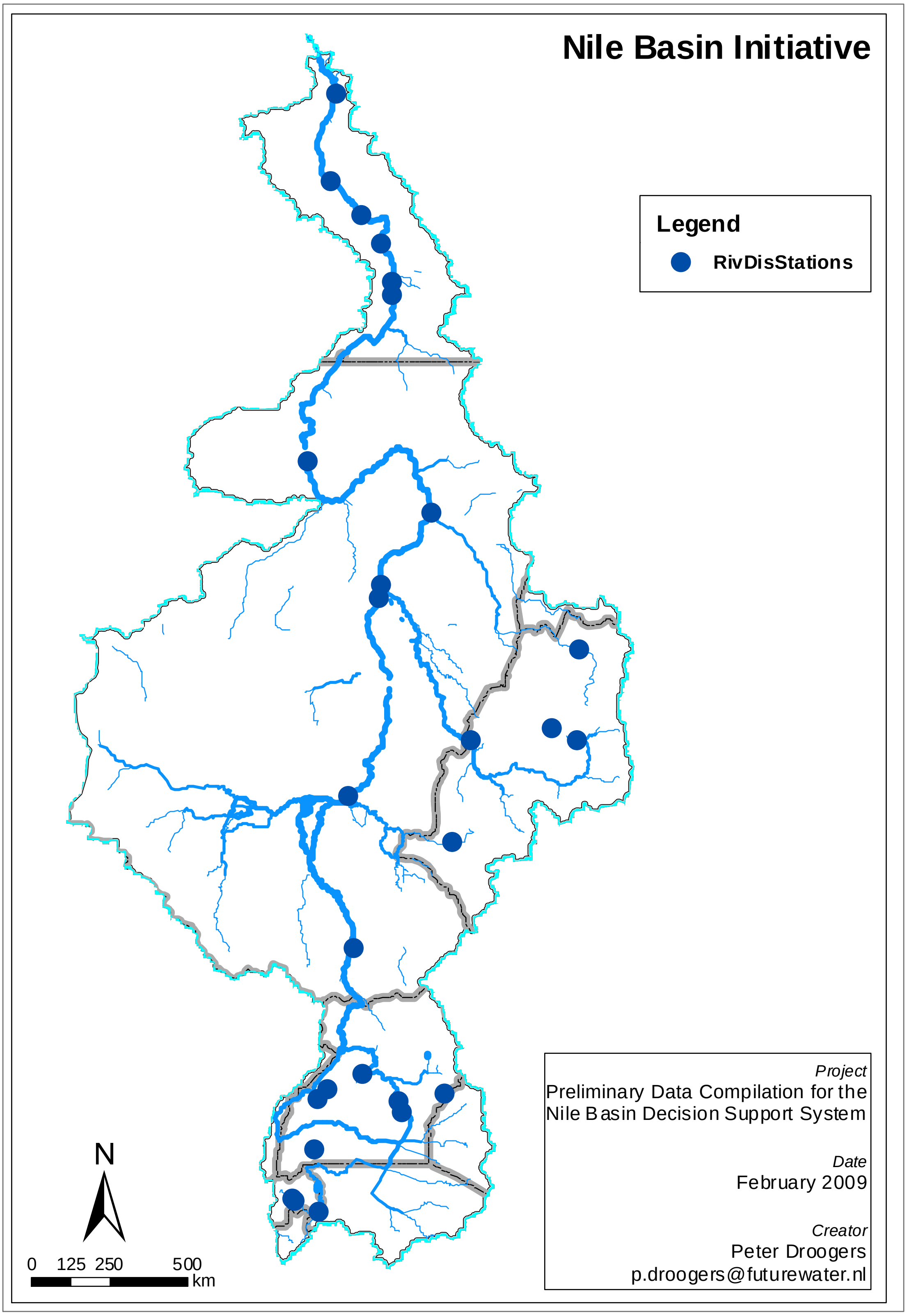
Silberschatz, A., H.F. Korth, S. Sudarshan. 2005. Database System Concepts. Fifth Edition.

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.

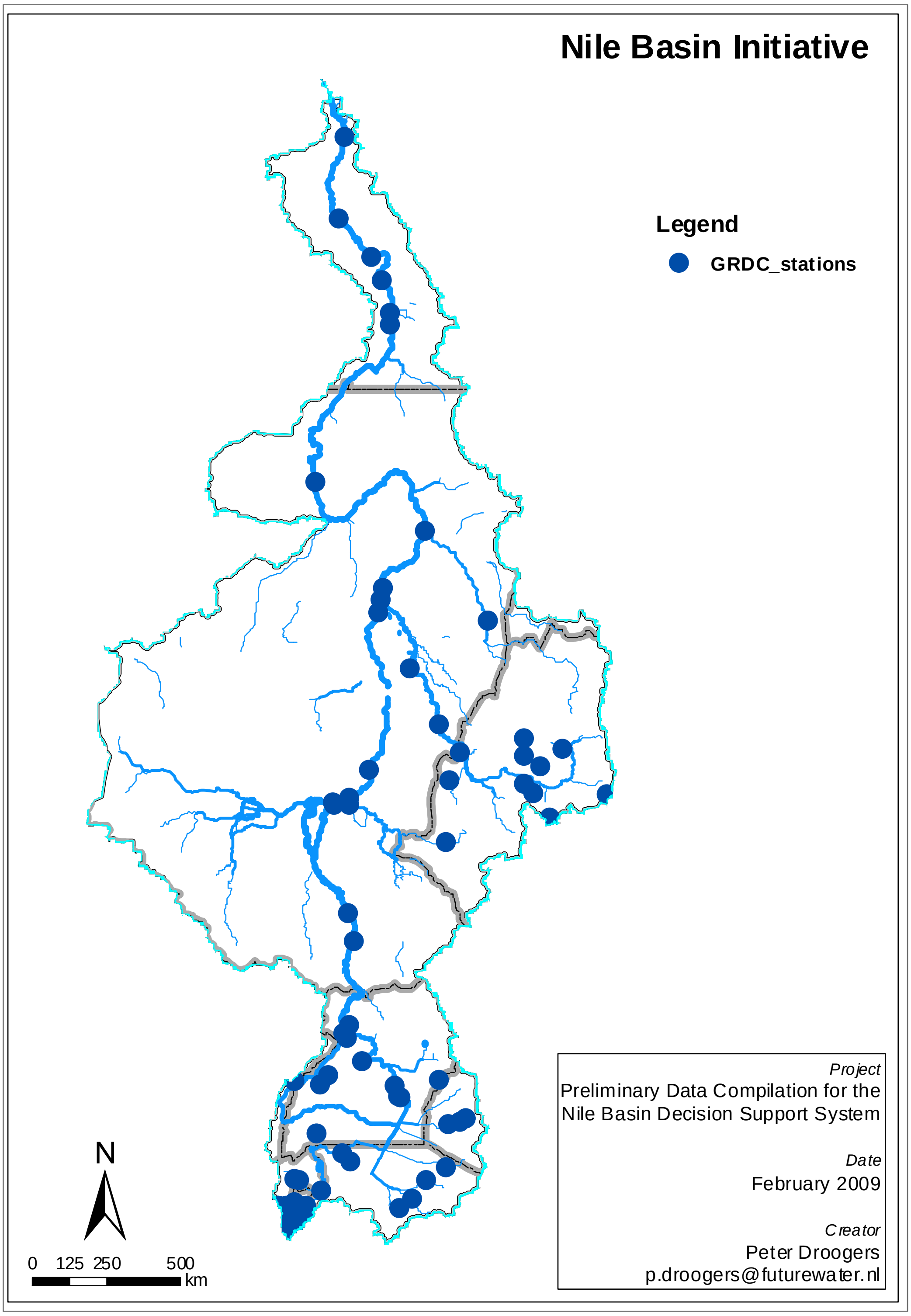
Van Engelen, V.W.P., T.T. Wen. 1995. Global and National Soils and Terrain Digital Databases (SOTER). Procedures Manual. International Soil Reference and Information Centre, Wageningen, Netherlands.

Vörösmarty, C.J., B. Fekete, and B.A. Tucker. 1998. River Discharge Database, Version 1.1 (RivDIS v1.0 supplement). Available through the Institute for the Study of Earth, Oceans, and Space / University of New Hampshire, Durham NH (USA) at http://pyramid.unh.edu/csrc/hydro/.

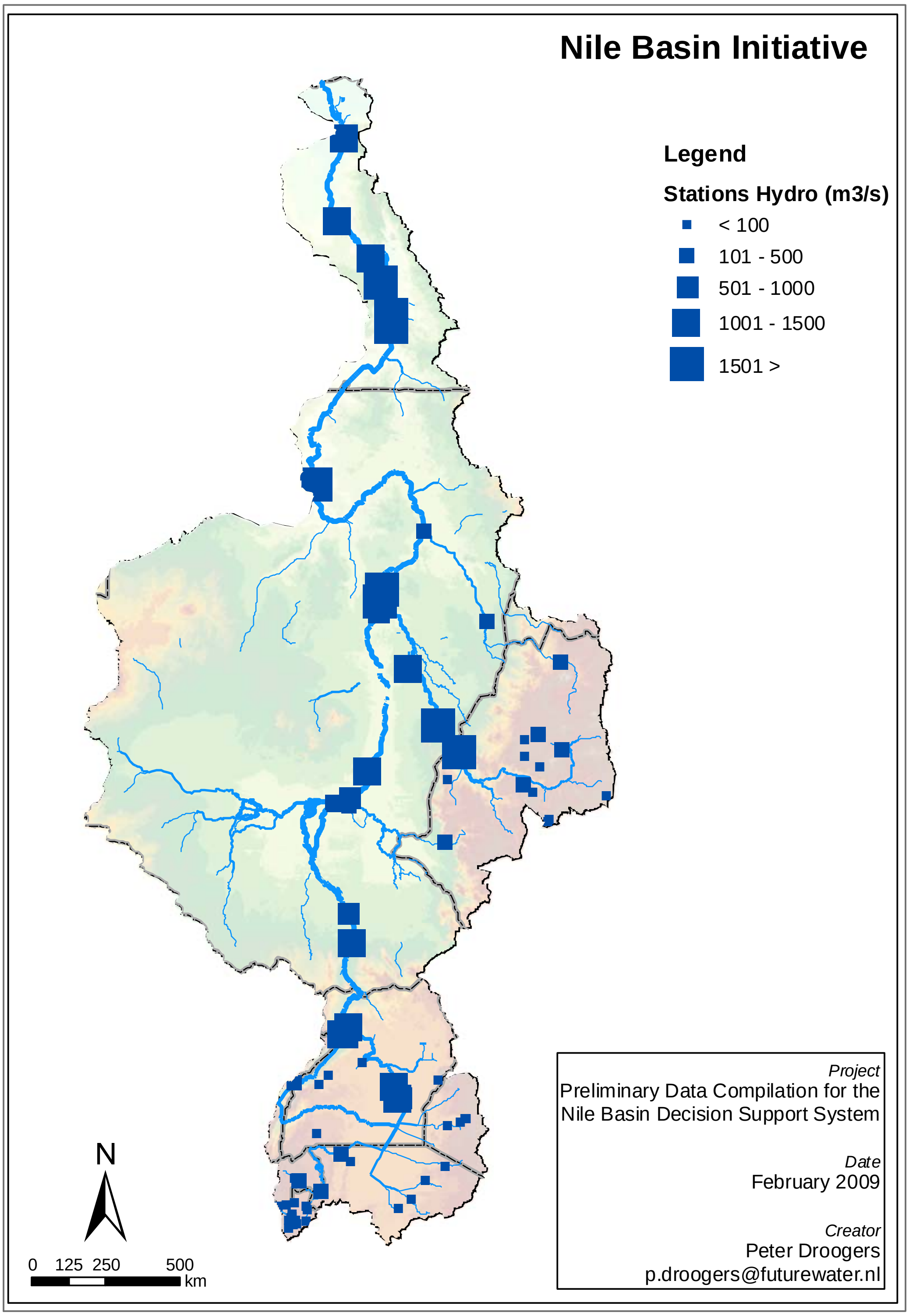
Xie, P. and P. A. Arkin, 1997. A 17-year monthly analysis based on gauge observations, satellite estimates, and numerical model outputs. Bulletin of the American Meteorological Society 78(11): 2539-58.



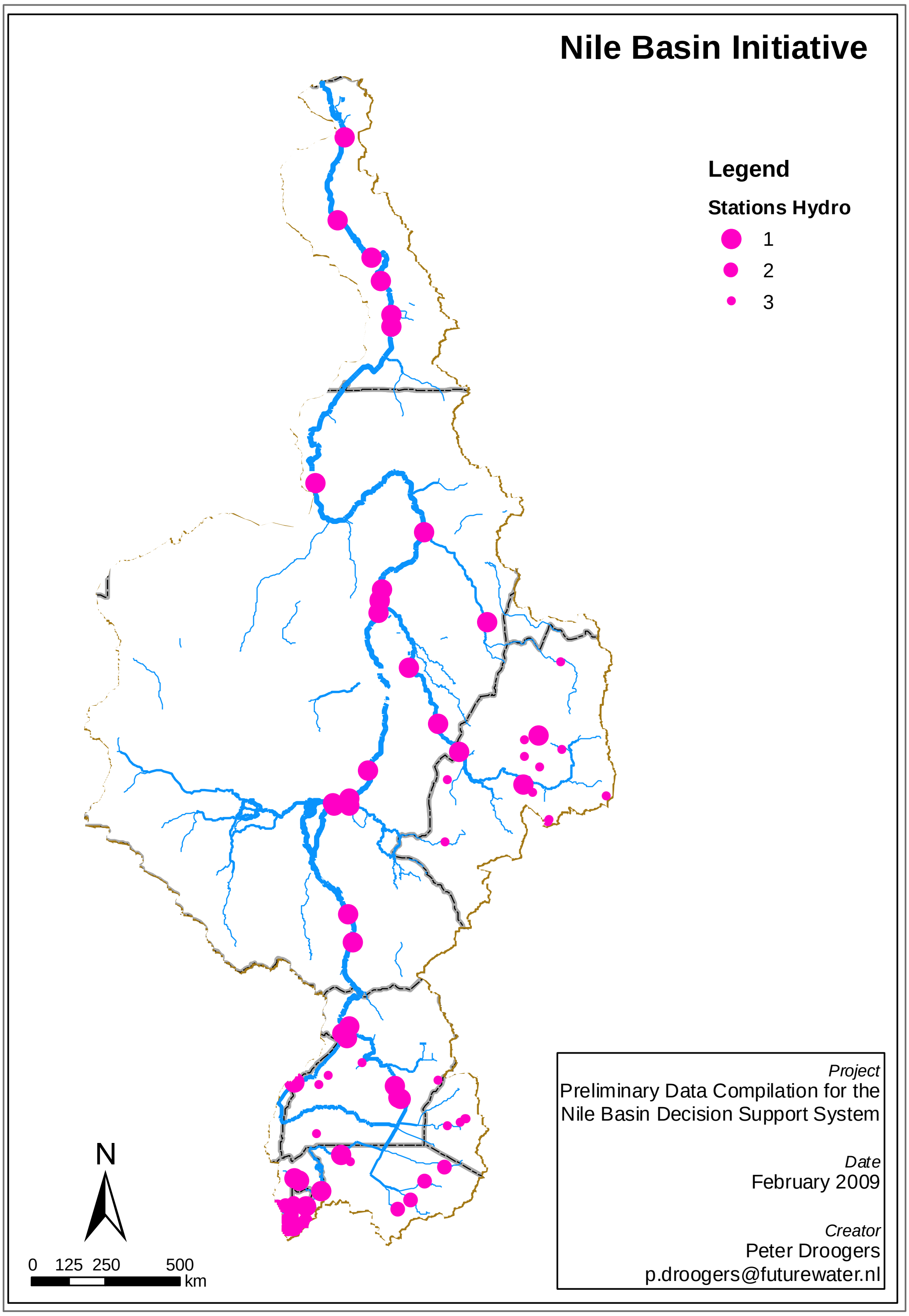
Map 1. Stations included in the RivDis database.



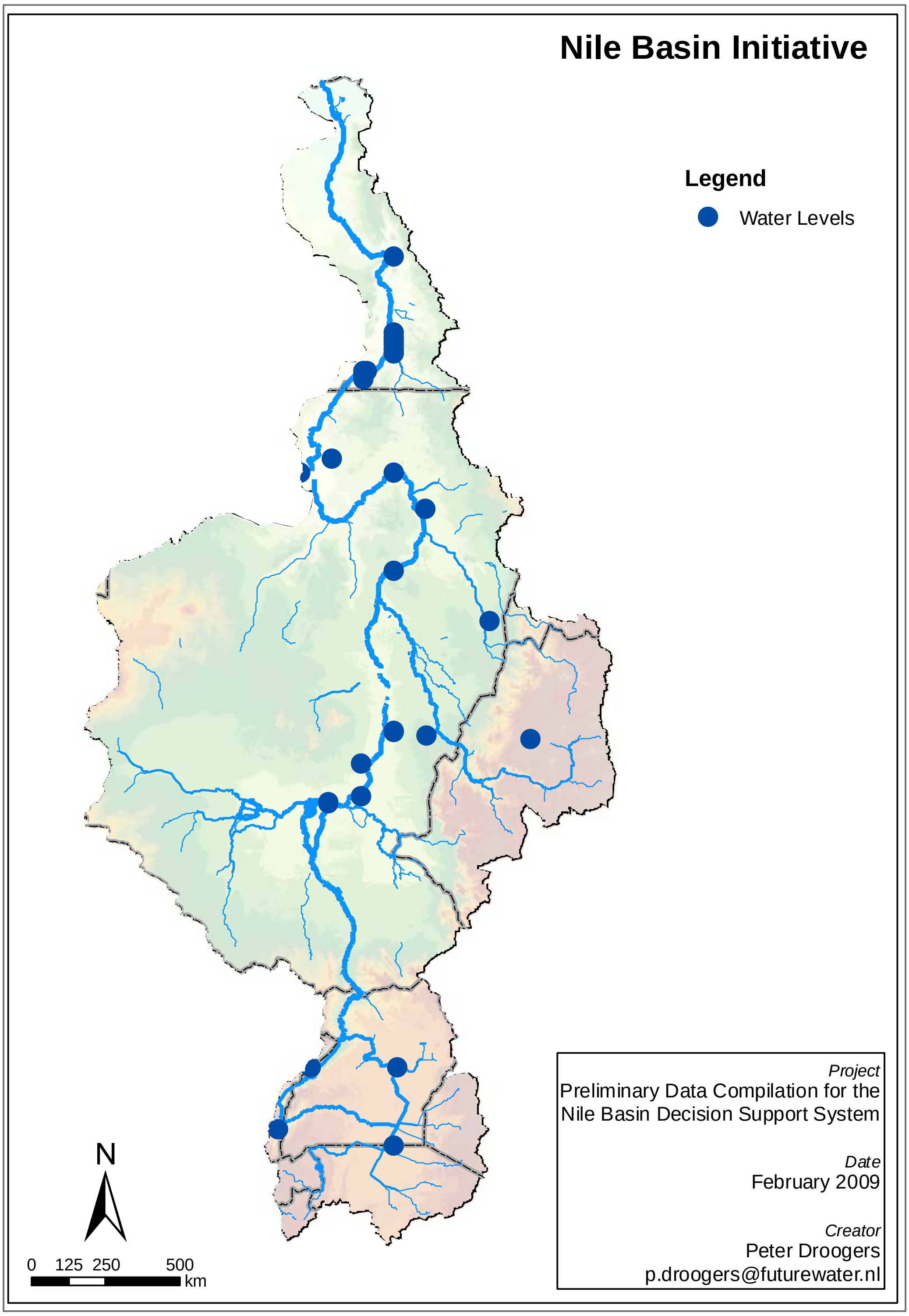
Map 2. Stations included in the GRDC database.



Map 3. Long term average streamflow.



Map 4. Streamflow stations in the public domain by category.



Map 5. Location of water levels observations based on the Poseidon and Jason satellites.

Appendix 1: Useful tools

During the development of the preliminary database a couple of tools have been proven to be very useful and effective for dealing with large data sets. Some of these tools were exclusively developed for this particular consultancy, but might be used for other projects as well. For all tools developed specifically for the project is source code handed over to NBI as well. So adjustment can be made in case required for related projects.

This appendix provides an overview of tools used. Note that this section describes only tools used for the temporal PostGreSQL database.

pgAdmin

pgAdmin is the most popular and feature rich Open Source administration and development platform for PostgreSQL, the most advanced Open Source database in the world.

pgAdmin is designed to answer the needs of all users, from writing simple SQL queries to developing complex databases. The graphical interface supports all PostgreSQL features and makes administration easy. The application also includes a syntax highlighting SQL editor, a server-side code editor, an SQL/batch/shell job scheduling agent, support for the Slony-I replication engine and much more. Server connection may be made using TCP/IP or Unix Domain Sockets (on *nix platforms), and may be SSL encrypted for security. No additional drivers are required to communicate with the database server.

pgAdmin is developed by a community of PostgreSQL experts around the world and is available in more than a dozen languages. It is Free Software released under the BSD License.

Navicat

Navicat for PostgreSQL is a powerful Database administration and development tool for PostgreSQL. It works with any PostgreSQL Database Server from version 8.0 or above, and supports most of the latest PostgreSQL features including Trigger, Function, View, and Manage User, etc. Features in Navicat are sophisticated enough to provide professional developers for all their specific needs, yet easy to learn for users who are new to PostgreSQL.

With Navicat well-designed Graphical User Interface (GUI), Navicat for PostgreSQL lets you quickly and easily create, organize, access and share information in a secure and easy way, taking PostgreSQL administration to the next level.

Navicat is developed by PremiumSoft and is a commercial product. Price varies somewhat between different versions and ranges from US$ 79 for the standard version up to US$ 179 for the enterprise version.

NBItool

The NBItool was developed specifically for the project. Most important component of the NBItool are:

- Import Stations
  - Import a CSV file with one or more stations
- Import Data
  - Import a CSV file with one or more data records
- Infill
  - Fills in missing data in the database

Some other functions were included as well, but were specifically developed for some older databases to be included in the PostGreSQL database.

A detailed manual has been written for the NBItool.

The tool is developed in Visual Studio C#. Executable as well as source code has been handed over to NBI.

SplitCSV

In some cases ascii files are too large to deal with. A typical example is in importing data into the PostGreSQL database using the NBItool. If files are exceeding 100,000 lines, performance of the NBItool are lower.

SplitCSV is specifally designed to be used in combination with the NBItool, since the header line of the original file will be copied to all other files.

The program works straightforward as a simple command line (DOS) tool:

- SplitCSV [infile] [NrLines]
- e.g. SplitCSV OLD_DATA_CLIMATE.CSV 50000

The tool is developed in Delphi. Executable as well as source code has been handed over to NBI.

Appendix 2: Local Data

Overview

During the consultancy trip of February 2009 a huge amount of data was obtained. A total of 61 GB of data were obtained consisting out of 50,000 files. Data have been divided in Public (called "professional data" in the context of the NBI-DSS) and Local (called "corporate data"). Table 2 and Table 3 shows the tree structure of these data as well as size in MB and number of files and folders.

In order to evaluate the relevance for the NBI-DSS the following description of each set of data are given here:

- `\Local Data\Database Kagera Monograph`
    - Data from the "Kagera Transboundary Integrated Water Resources Management & Development Project", located West of Lake Victoria covering parts of Uganda, Tanzania, Rwanda and Burundi.
    - All information is related to the relatively detailed level of the Kagera area and therefore somewhat less relevant for the entire NBI-DSS.
    - Much information is provided in pdf, jpg etc. Some Access databases, but many are unfinished or lacking consistency.
    - Most relevant info for the NBI-DSS can be found in `\Local Data\Database Kagera Monograph\GIS_DB\KAGERA_GIS_DB_3\1Nile_DST2006\Nile_GM`
        - FAO-DST with data for each country (this is exactly the same as the NileDST Cental Database)
- `\Local Data\Ethiopia MasterPlan`
    - Good dataset for Ethiopia. To be included in the NBI-DSS. Although not much recent data.
- `\Local Data\Karadobe`
    - Detailed study of the Abbay River Basin in Ethiopia. Includes data on streamflow to be included in the NBI-DSS.
- `\Local Data\NileDST Central Database`
    - Most relevant and complete information based on the FAO project, to be included in NBI-DSS. Details are provided hereafter
- `\Local Data\OSI Data`
    - Data mainly from Abbay, Baro and Tekeze Basins. Also some other graphs from Nile water flows.
- `\Local Data\Rusumo`
    - Report on Kagera (Rwanda/Tanzania): mainly hydropower analysis
- `\Local Data\SMMDSS Docs`
    - Small scale DSS for Sio-Malaba-Malakisi basin (Kenya/Uganda). Good documentation on general background of DSS.
- `\Public\FAO_Nile_Public_Data_20090109\`
    - The entire AfriCover
    - Results of FAO IGADD (Intergovernmental Authority on Drought and Development) crop production system zones (CPSZ) analysis.
    - Case study for Uganda on the development land use based on land cover.
    - Some MeteoSat images
    - Modis images from 2002-2004
    - Population

- \Public\IWMI-USDS\
    - Good report on Blue Nile. No actual data.

Table 2. Overview local data as obtained in February 2009 (sorted on size).

Name	Size	Allocated ▲	Files	Folders
Public	**31,290 MB**	**31,329 MB**	**13,857**	**725**
├**FAO_Nile_Public_Data_20090109**	**31,231 MB**	**31,269 MB**	**13,828**	**721**
└IWMI-USDS	60 MB	60 MB	29	2
├[Files]	57 MB	57 MB	1	2
├Working Papers	2 MB	2 MB	2	0
└Abstracts	1 MB	1 MB	26	0
Local Data	**30,023 MB**	**30,120 MB**	**36,660**	**2,470**
├**Database Kagera Monograph**	**28,160 MB**	**28,254 MB**	**34,822**	**2,242**
├NileDST Central Database	909 MB	909 MB	80	14
├SMMDSS Docs	510 MB	511 MB	523	29
├Karadobe	230 MB	230 MB	361	18
├Rusumo	137 MB	137 MB	356	61
├OSI Data	55 MB	55 MB	357	92
└Ethiopia MasterPlan	22 MB	22 MB	161	7

Table 3 **DST**

Directories:

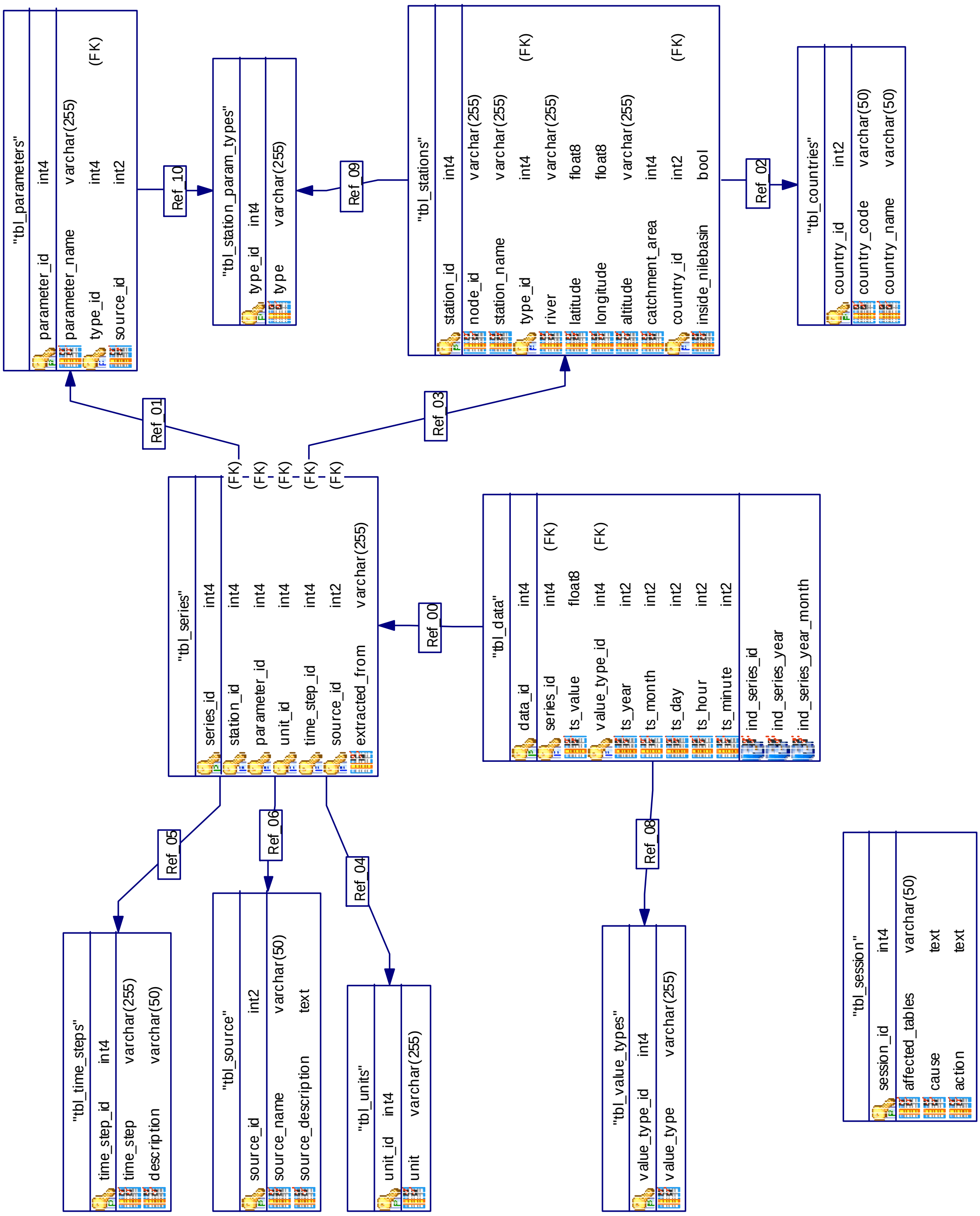
- \NileDST Central Database\Burundi\
    - 5 MS-Access files (see hereafter)
- \NileDST Central Database\DrCongo
    - 5 MS-Access files (see hereafter)
- \NileDST Central Database\Egypt
    - 5 MS-Access files (see hereafter)
- \NileDST Central Database\Eritrea
    - No data
- \NileDST Central Database\Ethiopia
    - 5 MS-Access files (see hereafter)
- \NileDST Central Database\Kenya
    - 5 MS-Access files (see hereafter)
- \NileDST Central Database\LVDSS_Data
    - Climate data from 45 stations
- \NileDST Central Database\MiscDataSets
    - Empty MS-Access files
    - Shape files of watersheds
- \NileDST Central Database\NileClimateZones
    - Shape files from countries and currencies
- \NileDST Central Database\Rwanda
    - 5 MS-Access files (see hereafter)
- \NileDST Central Database\Sudan
    - 5 MS-Access files (see hereafter)
- \NileDST Central Database\Tanzania

    - 5 MS-Access files (see hereafter)
- \NileDST Central Database\Uganda
    - 5 MS-Access files (see hereafter)

Each of the countries has 4 MS-Access files (except Eritrea which holds no data at all):

- Country Station Metabase.mdb
    - Station names, parameters, lat, lon, startdate, enddate
- Country StreamFlow Station Metabase.mdb
    - Station names, parameters, lat, lon, startdate, enddate
- CountryDB.mdb
    - All climate data
- CountryStreamFlowDB.mdb
    - All streamflow data

Appendix 3: PostGreSQL database



InfoSeries: provide info per series

```
SELECT
"public"."tbl_series"."series_id",
"public"."tbl_stations"."station_name",
"public"."tbl_parameters"."parameter_name",
Min("public"."tbl_data"."ts_year"),
Max("public"."tbl_data"."ts_year"),
Count("public"."tbl_series"."series_id")
FROM
"public"."tbl_series"
Inner Join "public"."tbl_data" ON "public"."tbl_series"."series_id" =
"public"."tbl_data"."series_id"
Inner Join "public"."tbl_stations" ON "public"."tbl_series"."station_id" =
"public"."tbl_stations"."station_id"
Inner Join "public"."tbl_parameters" ON "public"."tbl_series"."parameter_id" =
"public"."tbl_parameters"."parameter_id"
GROUP BY
"public"."tbl_series"."series_id",
"public"."tbl_stations"."station_name",
"public"."tbl_parameters"."parameter_name"
ORDER BY
"public"."tbl_stations"."station_name" ASC
```

CountSource: count data by data source name

```
SELECT
"public"."tbl_source"."source_name",
Count("public"."tbl_data"."data_id")
FROM
"public"."tbl_data"
Inner Join "public"."tbl_series" ON "public"."tbl_data"."series_id" =
"public"."tbl_series"."station_id"
Inner Join "public"."tbl_source" ON "public"."tbl_series"."source_id" =
"public"."tbl_source"."source_id"
GROUP BY
"public"."tbl_source"."source_name"
```

InfoMetaData: Obtain active meta data

```
SELECT
"public"."tbl_series"."series_id",
"public"."tbl_stations"."station_name",
"public"."tbl_countries"."country_name",
"public"."tbl_parameters"."parameter_name",
Min("public"."tbl_data"."ts_year") AS "FirstYear",
Max("public"."tbl_data"."ts_year") AS "LastYear",
Count("public"."tbl_series"."series_id") AS "NrRecords",
```

```
Min("public"."tbl_data"."ts_value") AS "Lowest",
Max("public"."tbl_data"."ts_value") AS "Highest"
FROM
"public"."tbl_series"
Inner Join "public"."tbl_data" ON "public"."tbl_series"."series_id" =
"public"."tbl_data"."series_id"
Inner Join "public"."tbl_stations" ON "public"."tbl_series"."station_id" =
"public"."tbl_stations"."station_id"
Inner Join "public"."tbl_parameters" ON "public"."tbl_series"."parameter_id" =
"public"."tbl_parameters"."parameter_id"
Inner Join "public"."tbl_countries" ON "public"."tbl_stations"."country_id" =
"public"."tbl_countries"."country_id"
GROUP BY
"public"."tbl_series"."series_id",
"public"."tbl_stations"."station_name",
"public"."tbl_countries"."country_name",
"public"."tbl_parameters"."parameter_name"
ORDER BY
"public"."tbl_stations"."station_name" ASC
```