VOLUME II

Water Resources Database Development

Agence canadienne de développement international

Canadian International Development Agency

WRC

Water Resources Commission

FINAL TECHNICAL REPORT

Hydrogeological Assessment Project of the Northern Regions of Ghana (HAP)

CIDA Ref: 7038883
CEA Ref: 604138
INRS Ref: R1327
ISBN Ref: 978-2-89146-710-0

December 2011

SNC·LAVALIN International

INRS Université d'avant-garde

SUMMARY

This report documents the work carried out to establish a sustainable groundwater resources database at the WRC. This database will contribute to the strengthening of WRC's institutional capacity and will at the same time help the WRC fulfill its central coordination role in integrated water resources management and planning in the water sector of Ghana.

The Arc Hydro database model was used to create a groundwater and surface water database for the Northern Regions of Ghana referred to as the *WRC Water Resources Database*. The Arc Hydro database model was chosen because it was designed by experienced GIS and water analysts and because it is offered freely to the GIS community. However, since it only offers a basis for the creation of a database, this model had to be modified and customized in order to meet the WRC data management needs. These modifications notably involved the creation or modification of database feature classes, domains and relationships.

Following the creation of the WRC Water Resources Database template, available groundwater data and part of the surface water data collected under the HAP or provided by the WRC were imported into this database using Arc Hydro tools. The import process required available data to be reorganized and standardized. Additional data were also generated to complement the existing information.

While the resulting database can provide a sustainable groundwater and surface water data repository for the WRC, additional work could improve the quality and quantity of water data stored in the database. Notably, some of the existing data still need to be verified against original borehole logs in order to correct remaining errors (e.g. keypunching errors) and also to add missing data, if necessary. The most important task to ensure the long term usefulness and continued existence of this database will however be related to database management. While this can be facilitated by standardizing data entry at the source (i.e. data coming from NGOs, donors, or contractors) and data integration operations at the WRC, adequate resources (i.e. human, financial, technical) will be required at the WRC in order to manage the database in an efficient and sustainable way.

TABLE OF CONTENTS

TABLES

FIGURES

APPENDICES

LIST OF ACRONYMS AND ABBREVIATIONS

AHGW	Arc Hydro Groundwater
AFD	Agence Française de Développement
CGIAR	Consultative Group for International Agriculture Research
CIAT	International Centre for Tropical Agriculture
CIDA	Canadian International Development Agency
CSI	Consortium for Spatial Information
CWSA	Community Water and Sanitation Agency (Ghana)
ESRI	Environmental Systems Reseach Incorporated
EU	European Union
GIS	Geographical Information System
GLOWA	Globalen Wasserkreislauf (Global Change in the Hydrological Cycle)
GSD	Geological Survey Department of Ghana
GWW	Ground Water for Windows software
HAP	Hydrogeological Assessment Project of the northern regions of Ghana
HGU	Hydrogeologic units
HSD	Hydrological Services Department of Ghana
MSD	Meteorological Services Department of Ghana
NGO	Non-governmental organization
RDBMS	Relational database management systems
SRTM	Shuttle Radar Topographic Mission (NASA)
SWERA	Solar and Wind Energy Resource Assessment
UN	United Nations
WATSAN	Water and Sanitation
WHO	World Health Organization
WRC	Water Resources Commission (Ghana)
WRI	Water Research Institute (Ghana)
WVI	World Vision International

1. INTRODUCTION

1.1 Background

The Hydrogeological Assessment Project (HAP) of the northern regions of Ghana was designed to contribute to the collection and analysis of scientific data on groundwater with the long term objective of improving groundwater resource management and development in the northern regions of Ghana, and thus contribute towards achieving the WATSAN targets set within the Ghana Poverty Reduction Strategy through "…and enhanced knowledge base and understanding of the hydrogeological conditions in the north of Ghana".[1]

One of the first steps towards the realisation of HAP objectives consisted in establishing the basis for current knowledge of the hydrogeological setting by assessing and consolidating the content of the electronic hydrogeological databases made available by stakeholders. Adequate and sustainable storage and management of these consolidated data required the development of a new database

1.2 Objectives and report content

This report documents the work carried out to establish a sustainable groundwater resources database at WRC (referred to as the *WRC Water Resources Database* in this report). This database will contribute to the strengthening of WRC's institutional capacity and will at the same time help the WRC fulfill its central coordination role in integrated water resources management and planning in the water sector of Ghana.

The Arc Hydro groundwater database model (Strassberg *et al.*, 2007) was selected to serve as the basis for the creation of the WRC Water Resources Database. This database model is the result of many years of work by experienced GIS and water analysts that is offered freely to the GIS community. It offers the possibility to manage and analyze both groundwater and surface water and can be adapted to suit user needs. This database model will not only help improve data access and management but it also gives the possibility to WRC to ensure data integrity and sustainability.

This report first presents an overview of the electronic water-related data collected under the HAP. This is followed by the rationale behind the selection of the Arc Hydro database model and by the description of the Arc Hydro database model itself. The modifications made to this database model to meet WRC's needs are then summarized, followed by an overview of the methodology used to import the water-related data already collected and stored in the database. Finally, the way forward for the sustainable management of this database is discussed.

[1] Memorandum of Understanding between the Government of Canada and Ghana for the Hydrogeological Assessment Project (April 2005)

2. DATA COLLECTION

2.1 Groundwater data

In the first year of the HAP (2006), electronic data were obtained from various sources. Electronic databases containing hydrogeology-related information were identified and obtained from the following stakeholders in the electronic format indicated in parentheses:

- Agence Française de Développement (AFD) (MS Excel file)
- Canadian International Development Agency (CIDA) (MS Excel file)
- European Union (EU) (MS Excel file)
- Global Change in the Hydrological Cycle Project (GLOWA) (MS Access file)
- Water Research Institute (WRI) (Ground Water for Windows (GWW) file)
- World Vision International (WVI) (Ground Water for Windows (GWW) file)

The first three databases, obtained via the Community Water Supply and Sanitation Agency (CWSA), contained only new records[2] from recent projects carried out mainly in the Northern Region. The GLOWA database contained records collected from different sources, notably private drilling companies, regional CWSA offices and the WRI. Although research conducted under the GLOWA project mostly concerns the Volta Basin, this electronic database contained records for all of Ghana. On the basis of this review of databases contents, the WRI database was considered the most complete hydrogeological database for the northern regions of Ghana at the time of data collection and comprised records from many water supply projects carried out in the northern regions. The WVI database contained records created under WVI projects and some records from the WRI database for the Northern Region that were considered as duplicates for the consolidation of collected data. A few records from various smaller water supply projects were also obtained from the CWSA.

The content of the electronic databases obtained through the above mentioned stakeholders were validated and consolidated into a unique database to serve HAP's purposes and to help for future hydrogeological projects in Northern Ghana. To facilitate the validation and consolidation process, all databases were converted into MDB files (i.e. Microsoft Access Database files). Some manual data transfer was required for databases in the GWW format as some data fields (e.g. lithology) could not be exported automatically in a convenient format. Prior to consolidation, the content of each database was assessed to identify 1) unique records among all available databases, 2) reliable records among these unique records (N.B.: reliability based on location data) and 3) resultant data gaps within the unique and reliable records. The first objective was aimed at eliminating redundant and duplicate information for data consolidation while the second and third objectives were aimed at establishing the need for additional data collection to achieve HAP's objectives.

2.1.1 Overview of data consolidation

The **identification of unique records** was not done the same way for all databases as reference information (e.g. report ID, project ID …) was not always available. The records from the AFD, EU and CIDA databases were all considered unique since they came from recent projects and were unlikely to have been entered in any other database yet (and thus duplicated). Consequently, a thorough verification of record uniqueness was not undertaken for these databases. For the GLOWA and WVI databases, it was generally possible to determine the presence of redundant records through database queries. As most of these records were originally taken from the WRI database, records from the latter were considered unique while redundant records in the GLOWA and WVI databases were identified as duplicates. The complete methodology is outlined in a preliminary database assessment presented in Appendix A (N.B.: results from this report have been updated since, as described below).

[2] A record of a specified database usually refers to a drilled well (or borehole) with all its associated data.

The **evaluation of record reliability** was carried out only for the data fields containing well coordinates, i.e. longitude and latitude. As original location data (e.g. hard copies of borehole logs) were generally not available, coordinates reliability was evaluated using spatial analysis functions with respect to administrative boundary and community GIS layers. Records identified as unreliable were flagged and kept for future and more thorough verification. It is important to mention that this was a time consuming task and that, therefore, it was not carried out with the same level of detail for all databases. An example of the process followed for the identification of reliable records is given in the preliminary database assessment report provided in Appendix A. While a complete analysis of all databases could have yielded more reliable records, the effort needed to accomplish this was considered disproportionate at the time in regards to the results that would be obtained.

For the HAP's purposes, the following data fields were considered to be the minimum data requirements[3] to carry out the necessary hydrogeological analysis: 1) Well state, 2) Well depth, 3) Lithology descriptions, 4) Static groundwater level, 5) Yield and 6) Water quality data (pH, conductivity. Fe, Mn, F). The **identification of data gaps** for these data fields is twofold: 1) identification of gaps in terms of data quantity and 2) identification of data gaps in terms of spatial distribution of data. The former was simply done through statistical analysis. Records containing information for each of these data field were compiled in order to evaluate the quantity of data available for each data field. The records resulting from this compilation were then plotted for each data field to assess their spatial distribution. This was done with reference to a 15 x 15 km cell grid that was considered the minimum requirement in terms of data distribution for HAP's purposes (i.e. min. density of 1 record per 225 km², with 507 cells in Northern Ghana) and considered suitable for the study area given the resources available.

Following the analysis of unique and reliable records, statistics were calculated for each database. Statistics showed that, for the initial assessment, the total number of unique records for the whole study area was 9 851 and that 7 594 of these records were considered reliable (as far as location data are concerned). Of the unique and reliable records, only 71 contained the minimum required data. These statistics also revealed that there were major gaps in the lithology and water quality data fields. Although further validation could have helped increase the number of reliable records, it was considered that the amount of work needed would be disproportionate in regards of the results expected. This was notably explained by the following problems:

- syntax errors in community and district names
- absence of community names in some databases
- presence of new communities in some database
- coordinate discrepancies for the same community

The verification of the spatial distribution of the 71 records identified above revealed that only 26 cells contained one or more boreholes with most of them located in the Upper East Region. Considering that 507 cells were necessary to cover the entire study area, it was obvious that additional data were needed, both in terms of quantity and in terms of spatial distribution, to carry out any significant data analysis required to meet HAP's objectives.

2.1.2 Collection and integration of additional groundwater data

Given the results of the initial assessment of the databases, additional data were thus collected manually from reports (soft and hard copies) identified earlier in the project. New data consisted mainly of reliable boreholes from water supply project reports that met at least one of the data requirements mentioned earlier. The selection was made mostly on the basis of the following:

- well data location (one well per 15 x 15 km cell)

[3] Minimum data requirements only apply to hydrogeological data found in available databases; other required data for HAP, such as climate data, are not considered in these requirements.

- well data representative of conditions encountered in the cell

This exercise, which was carried out by WRI staff yielded about 200 new reliable records, 65 of which met all the minimum data requirements.

2.1.3 Groundwater data collection summary

Following the additional data collection, the newly obtained reliable records were appended to the previously consolidated data. In addition to the new records obtained from the reports, new data generated by HAP targeted field work was also added to the consolidated data. This notably included 40 monitoring wells with their associated data (i.e. groundwater levels, groundwater sample analysis results, pumping test results …), geochemical data from three sampling campaigns targeting over a 100 boreholes, and analytical results from porewater and rainwater samples used for groundwater recharge estimation. Thus, the final number of records available at the time of writing was 10 139, of which 7 874 were considered unique and reliable records. Of these, 191 records had the minimum required data and those were spread out in 125 cells. While the validation process and the correction of errors described above were by no means an exhaustive exercise, the data collected and consolidated at that time of the project were deemed reliable enough to yield sound analysis results for HAP.

Throughout the consolidation process, data were temporarily kept in a flat file (i.e. a single database table termed HAP consolidated database) since potential software and models for this database were still being evaluated at the time. The database model evaluation and selection are discussed in section 3.

2.2 Surface water data

Surface water data were also collected during the HAP to support various analyses, validate results and prepare maps. Three types of data were mainly gathered: 1) hydrometric data (i.e. river flow measurements), 2) hydrographic data (i.e. watercourses and water bodies) and 3) rainfall data. The bulk of data were collected early in the project (2007-2008) but more recent data were also obtained towards the project end (2010).

Hydrometric datasets were provided by the Ghana Hydrological Services Department (HSD) in the form of MS Excel files for the following gauging stations in Northern Ghana:

- Black Volta River basin: Bui and Bamboi
- White Volta River basin: Daboya, Kpasenkpe, Nabogo, Nawuni, Paga, Pwalugu, Yarugu
  - Kulpawn River sub-basin: Yagaba
  - Nasia River sub-basin: Nasia
  - Red Volta River basin: Nangodi
  - Sisilli River sub-basin: Nakong and Wiase
- Oti River basin: Saboba

Datasets for these gauging stations include daily streamflow values (in cubic meters per second, m^3/s) for the 1951-2004 period, although they contain multiple data gaps of variable time length (from days to decades). For some gauging stations, it was possible to fill a number of data gaps with additional data obtained from Global Runoff Data Centre (www.grdc.bafg.de). However, considering the available streamflow data (in terms of temporal and spatial coverage) and project timeframe, it was not possible to generate representative streamflow estimates from existing data to fill all these gaps. Later in the project (2010), more recent daily hydrometric data (2004-2008 period) were also obtained for the following stations in Northern Ghana:

- White Volta River basin: Daboya, Kpasenkpe, Nabogo, Pwalugu, Yarugu
  - Kulpawn River sub-basin: Yagaba
  - Nasia River sub-basin: Nasia

- Red Volta River basin: Nangodi
- Sisilli River sub-basin: Nakong and Wiase

In addition, datasets from the following gauging stations located outside the northern regions were also obtained:

- Ankobra River basin: Beppoh, Bonsaso, Dwokwa, Prestea
- Dayi River basin: Gbefi, Hohoe
- Densu River basin: Asuboi, Akwadum, Mangoase, Pakro
- Pra River basin: Brenasi, Mfensi, Twifu Praso
- Tano River basin: Hwidiem, Sefwi Wiawso, Tanoso

While data obtained for most of these gauging stations consist of daily streamflow values, some also had monthly average streamflow values. The datasets cover variable time periods between 1962 and 2008. Table 2-1 summarizes all daily hydrometric data obtained.

Table 2-1 – Summary of available daily streamflow datasets

Station name	Region	Basin	Available years		Missing daily values (%)
			From	To	
Akwadum	Eastern	Densu	1977	2004	4307 (42 %)
Asuboi	Eastern	Densu	1964	2004	5297 (35 %)
Bamboi	Northern	Black Volta	1951	2004	7228 (37 %)
Beppoh	Western	Ankobra	2008	2008	273 (75 %)
Bonsaso	Western	Ankobra	2008	2008	273 (75 %)
Brenasi	Eastern	Pra	1991	2007	1168 (19 %)
Bui	Northern	Black Volta	1954	2004	3553 (19 %)
Daboya	Northern	White Volta	1962	2006	5152 (31 %)
Dwokwa	Western	Ankobra	2008	2008	278 (76 %)
Gbefi	Volta	Lower Volta (Dayi)	2002	2007	365 (16%)
Hohoe	Volta	Lower Volta (Dayi)	1962	2007	354 (21 %)
Hwidiem	Brong Ahafo	Tano	1991	2007	2355 (38 %)
Kpasenkpe	Northern	White Volta	2004	2006	124 (11 %)
Mangoase	Eastern	Densu	2002	2004	365 (33 %)
Mfensi	Ashanti	Pra	1991	2007	511 (8 %)
Nabogo	Northern	White Volta (Nabogo)	1962	2006	5590 (34 %)
Nakong	Upper East	White Volta (Sisilli)	1965	2006	10558 (69 %)
Nangodi	Upper East	Red Volta	1957	2006	13701 (75 %)
Nasia	Northern	White Volta (Nasia)	1951	2006	7142 (35 %)
Nawuni	Northern	White Volta	1953	2005	1636 (8 %)
Paga	Upper East	White Volta	2004	2005	115 (16 %)
Pakro	Eastern	Densu	1965	2004	8416 (58 %)
Prestea	Western	Ankobra	2008	2008	273 (75 %)
Pwalugu	Upper East	White Volta	1951	2006	9823 (48 %)
Saboba	Northern	Oti	1953	2004	3718 (20 %)
Sefwi Wiaso	Western	Tano	1991	2007	2145 (35 %)
Tanoso	Brong Ahafo	Tano	1991	2007	511 (8 %)
Twifo Praso	Central	Pra	1991	2007	1149 (18 %)
Wiase	Upper East	White Volta (Sisilli)	1961	2006	9458 (56 %)
Yagaba	Northern	White Volta (Kulpawn)	1957	2006	10636 (58 %)
Yarugu	Upper East	White Volta	1962	2006	9256 (56 %)

Hydrographic data were first obtained from the Solar and Wind Energy Resource Assessment Project in shapefile format (SWERA, 2005). The data, which covered the whole of the country, consisted of two files containing watercourse and water body features. However, close examination of these data revealed an insufficient level of detail for HAP's purposes. Additionally, data inaccuracies were noted in the northern regions, notably watershed limits were overlapping head streams in some areas. Therefore, larger scale hydrographic data were later obtained from the Ghana Survey Department along with other thematic data (i.e. cultural, forest, hypsographic, hydrographic, landform, transport and utility features). The data provided consisted of 43 files in ArcInfo interchange format (i.e. e00 files), each representing a 50k map sheet. As these files were not readily usable, data were re-organized by creating a mosaic out the 50k map sheets. This exercise however revealed several topological errors (e.g. overlapping polygons) that could not be corrected within project timeframe. The presence of these topological errors represents a limitation for spatial analysis and prevents the creation a flow network that could allow water flow modeling.

Rainfall data were obtained from the Ghana Meteorological Services Department (MSD). Daily, monthly and yearly datasets were obtained for different stations (all in millimeters, mm). Daily data covering the 2000-2010 period were first obtained for selected synoptic stations as part of meteorological datasets comprising the following data: precipitation, temperature, relative humidity, sun hours and wind speed. Because of project constraints, only rainfall data could however be included in the database (N.B.: other data can however be integrated into the database and linked to relevant stations in the future). These daily datasets were provided for the following 8 meteorological stations: Bole, Kete-Krachi, Navrongo, Sunyani, Tamale, Wa, Wenchi and Yendi. Concurrently, monthly meteorological datasets covering the 1961-2005 period were also obtained for the following stations: Ada, Bole, Kete-Krachi, Navrongo, Tamale, Wa and Yendi. Later in the project, additional daily and monthly datasets were obtained for other stations, some of them outside the northern regions. Daily rainfall datasets covered the 1950-2009 period although many of them contained multiple data gaps of variable time length (from days to decades). These datasets were provided for 34 non-synoptic stations, plus Akuse and Ho synoptic stations. As for the additional monthly rainfall datasets, they covered the 1934-2005 period (although with significant gaps also) for 10 non-synoptic stations, plus the Koforidua synoptic station. Finally, in order to improve the spatial distribution of rainfall data, yearly datasets comprising monthly average rainfall data were also obtained from 141 stations throughout Ghana for the 1971-2001 period. Information related to daily and monthly rainfall datasets acquired under HAP is summarized in tables 2-2 and 2-3.

Table 2-2 – Summary of available monthly rainfall datasets

Station name	Region	Station type	Available years		Missing monthly values (%)
			From	To	
Ada	Volta	Synoptic station	1961	2005	-
Adeiso	Eastern	Rainfall station	1939	1996	277 (40 %)
Apedua	Eastern	Rainfall station	1955	2002	184 (32 %)
Asafo Akim	Eastern	Rainfall station	1972	1995	50 (17 %)
Bole	Northern	Synoptic station	1961	2005	-
Kete-Krachi	Volta	Synoptic station	1961	2005	-
Koforidua	Eastern	Synoptic station	1965	2005	6 (1 %)
Kukurantumi	Eastern	Rainfall station	1960	2005	205 (37 %)
Mangoase	Eastern	Rainfall station	1939	2002	310 (40 %)
Nankese	Eastern	Rainfall station	1960	2002	90 (17 %)
Navrongo	Upper East	Synoptic station	1961	2005	-
Nsawam	Eastern	Climatological station	1934	2005	53 (6 %)
Pokoase	Greater Accra	Climatological station	1953	2005	23 (4 %)
Suhum	Eastern	Rainfall station	1939	2005	207 (26 %)
Tafo (CRIG)	Eastern	Agrometric station	1950	2005	53 (8 %)
Tamale	Northern	Synoptic station	1961	2005	-
Wa	Upper West	Synoptic station	1961	2005	-
Yendi	Northern	Synoptic station	1961	2005	-

Table 2-3 – Summary of available daily rainfall datasets

Station name	Region	Station type	Available years		Missing daily values (%)
			From	To	
Akaa	Volta	Climatological station	1956	2006	1223 (7 %)
Akuse	Eastern	Synoptic station	1950	2009	138 (1 %)
Amedzofe	Volta	Climatological station	1963	2009	1750 (10 %)
Anfoega Akukome	Volta	Rainfall station	1955	2002	7741 (44 %)
Baglo	Volta	Rainfall station	1955	1980	786 (8 %)
Bawku	Upper East	Rainfall station	1950	2004	8902 (44 %)
Bole	Northern	Synoptic station	2000	2010	365 (9 %)
Bolgatanga	Upper East	Agrometric station	1954	2009	9266 (45 %)
Busunu	Northern	Rainfall station	1953	1983	4282 (38 %)
Gambaga	Northern	Rainfall station	1950	2006	8377 (40 %)
Garu	Upper East	Agrometric station	1954	2009	11122 (54 %)
Helu	Volta	Rainfall station	1961	1983	3828 (46 %)
Ho	Volta	Synoptic station	1950	2009	75 (0.3 %)
Hohoe	Volta	Climatological station	1950	2009	2697 (12 %)
Karaga	Northern	Rainfall station	1954	1978	5142 (56 %)
Kayoro	Upper East	Rainfall station	1956	1983	5927 (58 %)
Kete-Krachi	Volta	Synoptic station	2000	2010	365 (9 %)
Kpandu	Volta	Climatological station	1958	2009	378 (2 %)
Kpeve	Volta	Agrometric station	1950	2009	1323 (6 %)
Kunkungu	Upper West	Rainfall station	1954	1965	575 (13 %)
Kusawgu	Northern	Rainfall station	1955	2008	7448 (38 %)
Leklebi Dafo	Volta	Rainfall station	1950	1980	1943 (17 %)
Likpe Mate	Volta	Rainfall station	1955	1985	2730 (24 %)
Nakrong	Upper East	Rainfall station	1967	2008	6420 (42 %)
Nalerigu	Northern	Rainfall station	1952	1981	4316 (39 %)
Nankpanduri	Northern	Rainfall station	1954	1979	1702 (18 %)
Nasia	Northern	Climatological station	1967	1983	881 (14 %)
Navrongo	Upper East	Synoptic station	2000	2010	365 (9 %)
New Ayoma	Volta	Rainfall station	1955	1986	1968 (17 %)
Paga	Upper East	Rainfall station	1956	1997	6291 (41 %)
Pong Tamale	Northern	Climatological station	1950	2009	1906 (9 %)
Pusiga	Upper East	Rainfall station	1955	1987	990 (8 %)
Sunyani	Brong Ahafo	Synoptic station	2000	2010	365 (9 %)
Tamale	Northern	Synoptic station	1950	2010	-
Teteman	Volta	Rainfall station	1976	1983	273 (9 %)
Tolon	Northern	Rainfall station	1961	1979	3740 (54 %)
Tsito	Volta	Rainfall station	1959	2009	6139 (33 %)
Tumu	Upper West	Rainfall station	1950	2004	8297 (41 %)
Wa	Upper West	Synoptic station	2000	2010	365 (9 %)
Walewale	Northern	Climatological station	1972	2009	2820 (20 %)
Wenchi	Brong Ahafo	Synoptic station	2000	2010	365 (9 %)
Wiaga	Upper East	Rainfall station	1954	2002	5251 (29 %)
Yapei	Northern	Rainfall station	1955	2004	11430 (63 %)
Yendi	Northern	Synoptic station	2000	2010	365 (9 %)

3. DATABASE MODEL

3.1 Evaluation and selection of database model

As data were collected and consolidated under HAP, discussions were initiated on the subject of the database model and software to be used. The selection of an adequate database model and software required investigations into database needs and capabilities at the WRC.

The first step in this process consisted in the evaluation of the database software used by the WRC at the beginning of the project. This software, named Ground Water for Windows (GWW), is a free hydrogeological database software developed in the early 1990's through a United Nations (UN) program. While this software comprises several useful database features and analysis functions, its initial assessment revealed some concerns. First, there has not been any update of the software since its release in 1994, which makes the software interface outdated. Another point of concern is that the GWW database structure and format do not allow a straightforward connection and data export to other widely used database software. Thus, any data analysis or browsing to be done with other software would involve time-consuming transfer of required data through error prone procedures. Spatial analysis and mapping capabilities of GWW are also limited in comparison to up to date GIS software. In order to address these concerns, the GWW lead developer, Jasminko Karanjac, and the GWW UN contact, Claude Sauveplane, were contacted to enquire if changes could be made to the software. As this proved unfeasible within HAP's timeframe and budget, other solutions were explored.

Therefore, the next step involved the investigation of other various database software, some specifically designed to host water-related data. These included the following: MS Access, ArcGIS, Rockworks, HydroGeo Analyst and Equis. Specifications and capabilities of these software packages were compared with those of GWW and with each other. This initiated further discussions with the WRC in relation with long-term database needs. Among the factors discussed were database analysis and management capabilities, software purchase and upgrade costs, software training and software compatibility. As it was collectively agreed that the WRC required an ArcGIS license to meet their mapping needs (whatever the database software selected), it was decided to use a database model that could be implemented within ArcGIS and its database platform, SQL Server Express. Investigations identified the Arc Hydro database model as the most relevant.

3.2 Description of Arc Hydro database model

The Arc Hydro database model started as a model for representing surface water systems within an ArcGIS geodatabase. In 2002, the Arc Hydro database model was published as an ESRI book entitled: Arc Hydro GIS for Water Resources (Maidment, 2002). Arc Hydro has been highly successful and has been widely adopted in industry. The groundwater database model has since been developed as a companion to the surface water database model.

The Arc Hydro groundwater database model is a geodatabase design used for representing multidimensional groundwater data (Strassberg *et al.*, 2007). The database model supports representations of different types of groundwater data including representation of data from aquifer maps and well databases, data from geologic maps, 3-D representations of borehole and hydrostratigraphy, temporal information, and data from simulation models. The model is based on the newly designed Arc Hydro framework, which is shared by the surface water and groundwater database models. Users can add groundwater and surface water components (GIS layers, data files…) to the framework as necessary, or develop their own components. This new componentized approach enables the tailoring of the geodatabase design to meet specific needs. More detailed information on the Arc Hydro database model can be found at:

- www.archydrogw.com
- www.crwr.utexas.edu/giswr/hydro
- www.aquaveo.com
- http://support.esri.com

3.2.1 Database file format

Arc Hydro is a conceptual and technical design implemented within a geodatabase (Strassberg *et al.*, 2007). Although the logic of the database model can be implemented in different GIS software packages, the proposed Arc Hydro design implements object classes from the ArcGIS geodatabase model. In order to query, visualize and use the content of such a geodatabase, an ArcView license level is required. To create, modify and properly manage geodatabase elements, an ArcEditor level license is required.

The geodatabase is a repository of geographic information organized into geographic data sets built on top of relational database management systems (RDBMS) such as Microsoft Access, Oracle, or Microsoft SQL Server that are customized for storing spatial data structures. Thus, the geodatabase performs as any standard RDBMS with additional capabilities related to the storage of geospatial features. Geodatabase objects used in the Arc Hydro database model include features, feature classes, relationships, rasters, and raster catalogs. **Features**, which are spatial vector objects (e.g., points, lines, polygons) with attributes (or data fields) to describe their properties, are stored within a **feature class** which represents a collection of features with the same geometry type, attributes, and relationships (N.B.: in a feature class, a feature is equivalent to a row in a table). **Relationships** are objects that define associations between feature classes based on key fields (e.g. a relationship can associate an aquifer and a well). Finallly, **raster data sets** represent imaged, sampled, or interpolated data on a uniform rectangular grid, and **raster catalogs** are used for storing, indexing, and attributing raster data sets. A detailed description of the geodatabase model and feature classes is provided by Zeiler (1999) and can also be found on the ESRI support website (http://support.esri.com).

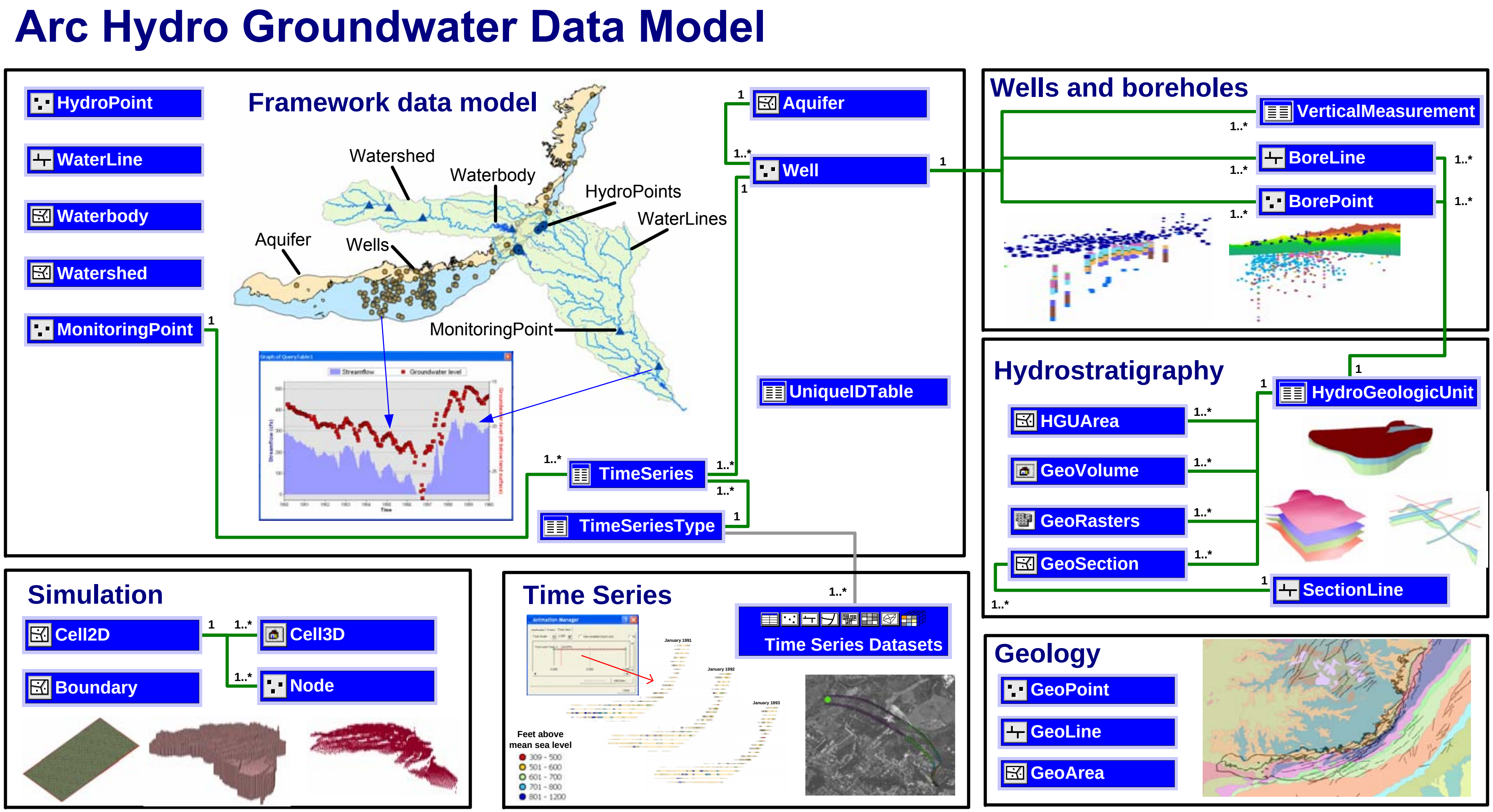
3.2.2 Overview of architecture

The Arc Hydro groundwater database model provides data structures for representing two-dimensional (2-D) and three dimensional (3-D) hydrogeologic features (e.g., aquifers, wells, faults, cross sections, and volumes), objects for describing computational grids (cells and nodes) to represent inputs and outputs from simulation models, and objects for storing tabular or gridded temporal information such as water levels and water quality measurements. The main components of the model are illustrated in Figure 3-1 and summarized below based on information given at: www.archydrogw.com/ahgw/Arc_Hydro_Groundwater_Data_Model. A complete description of the database model can be found in Strassberg, 2005.

The **Arc Hydro framework** provides a simple data structure for storing basic spatial datasets describing hydrologic systems. The framework supports basic water resources analyses such as tracing water as it flows over the terrain in watersheds, streams, and water bodies, creating groundwater levels and groundwater quality maps, and viewing time series data related to monitoring stations and wells. The Arc Hydro framework includes the following features:

- Aquifer- Polygon features representing aquifer boundaries. The features can be classified to represent different zones such as outcrop and confined sections of the aquifer.
- Well – Point features representing well locations and their attributes.
- WaterLine - Line features representing hydrographic “blue lines”, which represent mapped streams and water body center lines.
- Waterbody – Polygon features representing areas such as ponds, lakes, swamps, and estuaries.
- Watershed – Polygon features representing drainage areas contributing water flow from the land surface to the water system.
- HydroPoint – Point features representing hydrographic features such as springs, water withdrawal/discharge locations, and structures.
- MonitoringPoint – Point features representing locations where hydrologic variables are measured, such as stream-gage stations and precipitation gages.

Figure 3-1 – Main components of the Arc Hydro groundwater database model



The **Borehole component** contains classes for representing 3-D information recorded along boreholes. The data can be stored as tabular information related to well features or as 3-D point and line features that can be visualized in ArcScene.

- BoreholeLog - Table for representing vertical data along boreholes. Each row in the BoreholeLog table represents a point or interval along the borehole.
- BorePoint - 3-D point feature class for representing point data along boreholes.
- BoreLine - 3-D line feature class for representing interval data along boreholes.

The **Hydrostratigraphy component** represents hydrogeologic units (HGU) using 2-D and 3-D features. Classes in the component enable the representation of hydrogeologic models including 2-D polygons representing the extent of hydrogeologic units, cross sections, surfaces representing the top and bottom of hydrogeologic units, volume elements.

- HydrogeologicUnit – Table for defining conceptual hydrogeologic units.
- GeoArea - Polygon feature class for representing the 2-D extent of hydrogeologic units or parts of them.
- GeoSection – 3-D multipatch feature class representing vertical cross sections.
- SectionLine – Line feature class for representing 2-D section lines.
- GeoRasters - Raster catalog for storing raster surfaces. The catalog enables storing rasters within the geodatabase and adding attributes describing raster datasets.
- GeoVolume – multipatch feature class for representing 3-D volumes.

The **Geology component** consists of a set of objects for representing data from geologic maps and to integrate geologic data with other groundwater-related datasets.

- GeologyPoint – Points that represent locations such as springs, caves, sinks, and observation points.
- GeologyLine – Line features that describe objects such as faults, contacts, and dikes.
- GeologyArea – Polygon features describing areal features such as rock units.

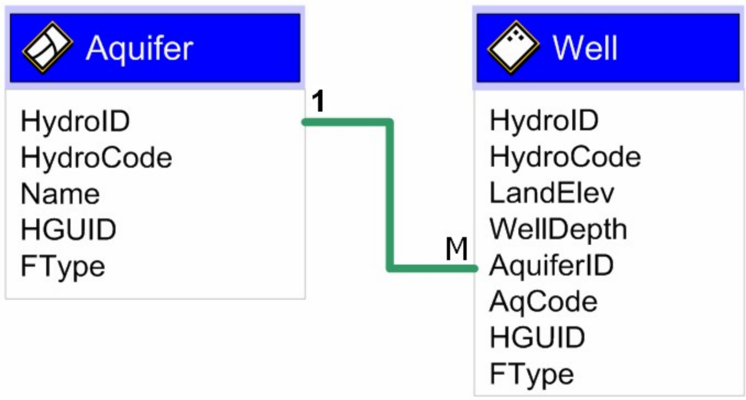
The **Time Series component** provides a design for dealing with temporal data series within Arc Hydro. This design provides better support for utilizing multiple representations of time series data and a table structure to describe time series variables.

The AHGW **Simulation feature dataset** is a set of vector feature classes that can represent common modeling objects. It is designed to allow for representation of finite element and finite difference systems in a GIS. The simulation feature dataset includes five feature classes: Boundary, Cell2D, Cell3D, Node2D, and Node3D. These feature classes enable storage and representation of model inputs and outputs related to the simulation objects.

- Boundary - Polygon feature class that represents the two dimensional extent of a model. It is not an essential part of the model representation but it can be useful to illustrate the location of the simulation model to support a simple spatial reference or database query.
- Cell2D – Polygon feature class used to represent cells or elements associated with two-dimensional simulation models or a single layer of a three-dimensional model.
- Cell3D – Multipatch feature class for representing three-dimensional cells and elements of simulation models.
- Node2D – 2-D Point feature class that represents nodes in a 2-D model grid/mesh or the nodes in a single layer of a 3-D model grid/mesh.
- Node3D – 3-D Point feature class that represents nodes in a model grid/mesh.

Internal relationships between the components of the AHGW database model are managed through the use of a unique identifier termed HydroID (usually an integer number). Every feature in the AHGW database model is assigned a unique HydroID that is used to manage relationships between features and relate features with tabular data. Figure 3-2 illustrates the relationship between features of the Well feature class and features of the Aquifer feature class. In this example, the relationship has a one-to-many cardinality[4] since an aquifer feature can be associated with one or more wells. The latter are related to aquifer features through the AquiferID data field, which is equal to the HydroID of an aquifer feature. Other database features and tabular data are related in a similar way although different types of relationships can be used (e.g. one-to-one or many-to-many).

Figure 3-2 – Example of database feature relationship using HydroID (Strassberg, 2008)



3.2.3 Arc Hydro Groundwater tools

Arc Hydro Groundwater Tools were developed by ESRI (www.esri.com) and Aquaveo (www.aquaveo.com) to improve groundwater data management capacities within ArcGIS. Based on the Arc Hydro groundwater database model, the tools enable the user to take advantage of the ArcGIS platform to archive, manage, and visualize groundwater information as well as time series and geological data. The toolkit includes Groundwater Analyst, MODFLOW Analyst and Subsurface Analyst which are described below.

Tools in the **Groundwater Analyst** allow the user to import data into the AHGW database model, manage key attributes and visualize data. With Groundwater Analyst, the user is able to import a variety of datasets (wells, time series, cross sections, volumes) into a geodatabase, manage symbology of layers in ArcMap and ArcScene, map and plot time series, and create common products such as water level, water quality, and flow direction maps. These tools are free to all licensed ArcGIS users with an ArcMap level license.

The **Subsurface Analyst** allows the user to create and visualize both 2-D & 3-D geologic models, starting with classification and visualization of borehole logs, creation and editing of cross sections, and generation of 3-D geosections and geovolumes. The complete set of tools included in the Subsurface Analyst are available for about 2000 $ CAD.

The **MODFLOW Analyst** enables the user to create, archive, and visualize modflow models within ArcGIS. The geoprocessing tools are based on the MODFLOW Database model, which supports the storage of a complete MODFLOW model (including grid structure, inputs, results) within an ArcGIS geodatabase. Tools in the toolkit enable the user to import an existing model into the geodatabase and geo-reference the model so that results can be visualized and analyzed in the context of other GIS data and that new models can be created from GIS features. The complete set of tools included in this toolkit are available for about 1500 $ CAD.

Information on software requirements and installation procedure for the Arc Hydro groundwater tools can be found on the Aquaveo website (www.aquaveo.com/ahgwinstallation).

[4] The cardinality of a relationship specifies the number of objects in the origin class that can relate to a number of objects in the destination class; a relationship can have one of three cardinalities: one-to-one, one-to-many and many-to-many.

4. DATABASE MODEL CUSTOMIZATION

To create the WRC Water Resources Database, the Arc Hydro groundwater database model described in the previous section was customized in order to adequately store all available data and to suit database needs at the WRC. Some of the feature classes and tables proposed in the original model were thus left out of the customized database model and a few new ones were created to complement the existing ones. Furthermore, additional attributes (or data fields) were created in some of the original feature classes and tables as only basic attributes are predefined in the model. To provide means of ensuring data integrity, new domains[5] and relationship classes[6] were added to those proposed in the database model. All these modifications are summarized in the following sections (N.B.: feature classes and tables are underlined and *data fields* are in italic font to clarify descriptions) and a simplified database diagram is provided in Figure 4-1.

4.1 Database structure

4.1.1 Framework component

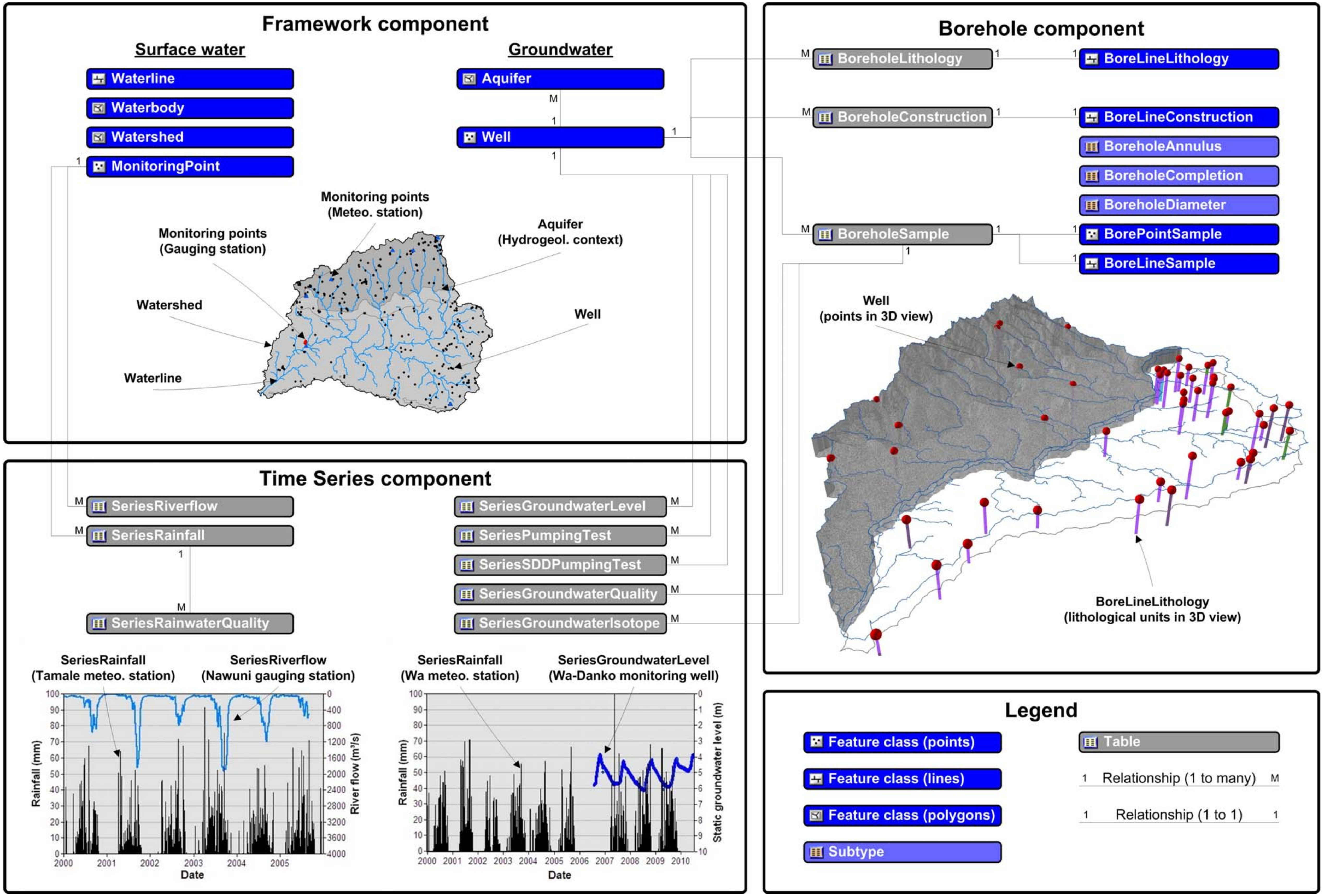
The Arc Hydro template for the Aquifer feature class was used as is, except for the addition of one data field, *GeoUnit*. In the WRC Water Resources Database, this feature class was used to store hydrogeological contexts rather than aquifers since available hydrogeological information did not allow adequate definition of aquifers. These hydrogeological contexts were primarily delineated on the basis of geology[7] and the *GeoUnit* data field thus identifies the name of the geological unit(s) comprised within each hydrogeological context. Also, the data field *HGUID* was left empty since the HGU (i.e. hydrogeologic units) defined for this project cannot be generalized laterally at regional scale with the available information.

For the Well feature class, several new data fields were added to complement the predefined data fields in the Arc Hydro template. The main data fields or groups of data fields added comprise: geographic coordinates (*Longitude*, *Latitude*), condition or state of well (*WellState*), type of well (*WellType*), use of well (*WellUse*), administrative areas (*Region*, *District*, *Community*), hydrological basin (*Basin*), companies involved in siting and drilling activities (*BhComp*, *SitingComp*), dates of siting and drilling activities (*WellDate*, *SitingDate*), drilling method (*BhMethod*), airlift yield (*AirliftYld*), pump information (*PumpType*, *PumpDepth*, *PumpDate*), comments (*CommentsGeneral*, *CommentsLocation*), record entry date (*EntryDate*), data availability or reliability identifiers (e.g. records with reliable coordinates) (*HasXY*, *HasDpt*, *HasLth*, *HasSwl*, *HasYld*, *HasGwq*) and duplicate record identifier (*HasDupl*). The content of the latter data field was generated from a non-exhaustive duplicate identification process that was carried out on data consolidated under the HAP. This process was aimed at identifying duplicates to be left out of the HAP analyses and possible duplicates that would require a more detailed verification. Aside from the creation of the above mentioned data fields, the *HGUID* field was replaced by the *HGUGeneral* field (c.f. section 4.1.3) since, for most records, more than one HGU was intercepted. The purpose of these generalized HGUs is to provide a rough indication of the productive zone intercepted by the well. It is also important to mention that space-dependent information (e.g. *LandElev*, *AquiferID*, *Basin* and *GeoUnit*) were not assigned to records for which coordinates were considered unreliable (i.e. *HasXY* = 0).

To complement the Well feature class, a BoreholeIndex table was created to store additional information that allows the identification of well in relation with other databases obtained during the HAP. Data fields in this table include well identifiers from other databases (*WellID_1*, *WellID_2*, *WellID_3*, *WellRefNo*), name and funding of the project under which the well was

[5] Domain : Defines a set of legal attribute values for a data field; it limits the possibility of data entry error because the choices for each value are limited within defined parameters.
[6] Relationship class : Defines the properties of a relationship or of an association between two objects.
[7] The geological map used for the definition of main hydrogeological contexts is the 2009 version published by the Ghana Geological Survey Department (GSD, 2009); the *GeoUnit* data field of the Aquifer feature class corresponds to a combination of the *Stratigraphy* and *TectonicDomain* data fields of the soft copy of the map.

Figure 4-1 – Simplified diagram of WRC Water Resources Database



constructed (*PrjctName*, *PrjctFund*), the licence number of the drilling contractor (*LicenceNo*) and data sources for the available information (*DSource1*, *DSource2*, *DSource3*).

The MonitoringPoint feature class was used to store and represent point locations where groundwater and surface water monitoring data were collected. For groundwater data, this consisted in a subset of the Well feature class comprising active monitoring wells. All complementary information such as lithology, construction details, groundwater levels are however linked to the features in the Well feature class. Consequently, an additional data field, *WellID*, was created in the MonitoringPoint feature class to link back to these features in the Well feature class. For surface water data, the monitoring points comprised meteorological stations (i.e. agro-metric station, climatological station, rainfall station and synoptic station) and river flow gauging stations. Data obtained from these monitoring points were respectively stored in the SeriesRainfall table and the SeriesRiverflow table (see section 4.1.4).

4.1.2 Borehole component

In the Arc Hydro database model, the BoreholeLog table of this component was designed to store and manage all vertical measurements related to features in the Well feature class. This table was however replaced by three similar tables, BoreholeLithology table, BoreholeConstruction table, BoreholeSample table, in order to organize, manage and visualize datasets of different nature in a more efficient way. The structure of these tables was based on the BoreholeLog table template to which four data fields were added: 1) a *LogID* field to provide a unique vertical measurement (or log) identifier, 2) an *FGeometry* field to specify the geometry of the feature considered (e.g. point or interval), 3) a *LogCode* field to indicate the possible values associated to the type of vertical measurement (e.g. type of screen for construction details) and 4) a *LogDescr* field to provide a free text description of the vertical measurement. In the BoreholeConstruction table, the HGUID and HGUCode fields were replaced by the HGUGeneral and HGUCodes fields since, for most vertical measurements, more than one HGU was intercepted. As stated in section 4.1.3, the information in these fields was only extracted for records with lithology data and known screened or open-hole intervals.

Three 3-D feature classes, BoreLineLithology, BoreLineConstruction and BoreLineSample, were created accordingly to represent, when possible, vertical measurements listed in the tables mentioned above. These feature classes were based on the BoreLine feature class of the Arc Hydro database model. Except fo the *FGeometry* field which would be redundant with the *SHAPE* field of the feature class, the same data fields added to the tables above were added to these feature classes.

4.1.3 Hydrostratigraphy component

Considering the extent of the project area and the available lithological data, the definition of hydrostratigraphic units (or hydrogeologic units, i.e. HGU) was largely based on levels of bedrock weathering and fracturing, which generally have a significant influence on the occurrence of groundwater in Northern Ghana. These HGU are therefore not built from sets or groups of lithological units but rather from qualitative indications given in borehole logs that are related to weathering and fracturing. These units are listed in the HydroGeologicUnit table, which is used without modification.

For some of the feature classes in the other Arc Hydro database model components, *HGUGeneral* replaced the *HGUID* data field since more than one HGU was associated with the record considered. *HGUGeneral* thus represents a generalized HGU based on all HGU intercepted (e.g. if highly and moderately weathered rock units are intercepted by the screen interval, the generalized HGU would be 'Regolith'). In some feature classes, all HGU intercepted by a well are also stored in the *HGUCodes* data field for future reference. The information in both these data fields, *HGUGeneral* and *HGUCodes*, is only available for records that have both lithology data and construction details (i.e. screen or open hole intervals).

Finally, given data availability and project constraints, the proposed spatial representation of the hydrostratigraphy within the Arc Hydro database model (Section 3.2.2) was not implemented.

4.1.4 Time series component

The storage of temporal data collected under HAP required the creation of eight additional tables:

- Groundwater data: SeriesGroundwaterLevel table, SeriesPumpingTest table, SeriesStepDrawdownPumpingTest table, SeriesGroundwaterQuality table and SeriesGroundwaterIsotope table,
- Surface water data: SeriesRainfall table, SeriesRainwaterQuality table and SeriesRiverflow table

The groundwater data-related tables were respectively created to store groundwater level data, pumping test data, step-drawdown pumping test data, geochemical data and isotopic data. The structure of these tables was based on the AttributeSeries table template. The latter template was preferred over the TimeSeries template as it allows the storage of multiple time-dependent variables in the same table (e.g. multiple physico-chemical parameters analysed at a given well on a given date). The variables included in each of these tables correspond to the most commonly measured variables in databases and in hydrogeological studies collected under the HAP. If necessary, additional variables can be appended to these tables in the future.

The surface water data-related tables were respectively created to store rainfall data[8], geochemical data on rainwater and river flow measurements. The structure of the table storing rainwater geochemical data was based on the AttributeSeries table template as it contains many time-dependent variables like the groundwater-related tables. The other two tables, storing rainfall and river flow data, were however based on the TimeSeries template as both datasets contain only a single time-dependent variable (i.e. rainfall and river flow). The time-scale of measurements for these single variables can be indicated in the *VarID* data field which links to the VariableDefinition table (see below).

In order to document variables of all time series tables in the database, the VariableDefinition table was also used. Two data fields, *Standard* and *Note*, were added to this table to respectively provide a quality standard for the variable (e.g. WHO drinking water guideline values for geochemical parameters (WHO, 2008) and a free text description or general comment related to the variable or the values entered for this variable.

Finally, the SeriesCatalog table was also incorporated into the WRC Water Resources Database in order to provide a quick overview of the times series avaiable in the database.

4.1.5 Other database components

4.1.5.1 Geology component

All feature classes of the geology component were left out of the WRC Water Resources Database. Many geological features such as caves, sinks, fault, dikes and regional alteration zones were unavailable at the time of writing and, considering that the available information (i.e. polygons features representing geological units) was used to define hydrogeological contexts, it was decided not to include this component in the database.

4.1.5.2 Simulation feature dataset

At the time of writing, all feature classes of this component were left out of the WRC Water Resources Database as there were no needs for them in the foreseeable future.

[8] Only rainfall data could be included in the database within the project timeframe; however, other meteorological data such as temperature (avg., min., max.), relative humidity (min., max.), sun hours and wind speed could also be included and linked to existing monitoring points in the future.

4.2 Database validation features

4.2.1 Attribute validation rules

The Arc Hydro groundwater database model is designed to allow the definition of relationships between database objects. This is done through the creation of relationship classes linking feature classes and tables within the geodatabase. The implementation of the modifications presented in the previous sections required the creation of additional relationship classes besides already existing ones. The final list of relationships added is presented in Table 4-1.

Table 4-1 – List of WRC Water Resources Database relationships

Name of relationship	Cardinality	Origin object class	Destination object class
AquiferHasWells	OneToMany	Aquifer	Well
BhConstructionHasBoreLineConstruction	OneToOne	BhConstruction	BoreLineConstruction
BhLithologyHasBoreLineLithology	OneToOne	BhLithology	BoreLineLithology
BhSampleHasBoreLineSample	OneToOne	BhSample	BoreLineSample
BhSampleHasSeriesGroundwaterQuality	OneToMany	BhSample	SeriesGroundwaterQuality
BhSampleHasSeriesGroundwaterIsotope	OneToMany	BhSample	SeriesGroundwaterIsotope
MonitoringPointHasSeriesRainfall	OneToMany	MonitoringPoint	SeriesRainfall
MonitoringPointHasSeriesRiverflow	OneToMany	MonitoringPoint	SeriesRiverflow
SeriesRainfallHasSeriesRainQuality	OneToMany	SeriesRainfall	SeriesRainwaterQuality
WellHasBoreholeConstruction	OneToMany	Well	BoreholeConstruction
WellHasBoreholeIndex	OneToOne	Well	BoreholeIndex
WellHasBoreholeLithology	OneToMany	Well	BoreholeLithology
WellHasBoreholeSample	OneToMany	Well	BoreholeSample
WellHasMonitoringPoint	OneToOne	Well	MonitoringPoint
WellHasSeriesGroundwaterLevel	OneToMany	Well	SeriesGroundwaterLevel
WellHasSeriesPumpingTest	OneToMany	Well	SeriesPumpingTest
WellHasSeriesSDPumpingTest	OneToMany	Well	SeriesSDPumpingTest

To improve data integrity within the WRC Water Resources Database, additional domains were also created and assigned to appropriate data fields. All domains defining the values permitted for the specified data fields are listed inTable 4-2.

4.2.2 Spatial validation rules

While there are no topology rules defined in the AHGW database model, it was decided to include at least one topology in the WRC Water Resources Database in order to facilitate spatial analysis and ensure consistent results when using the Aquifer feature class. This topology, named AquiferTopology, was defined and validated for the Aquifer feature class with two basic rules: 1) polygons must not have gaps and 2) polygons must not overlap. No topology was created for the Well feature class as too many wells were without reliable coordinates.

4.3 Database coordinate system

The coordinate system used for all database spatial features is the Ghana Meter Grid. The details of this projected coordinate system are stated below:

Projection: Transverse_Mercator
False Easting: 274319.510000
False Northing: 0.000000
Central Meridian: -1.000000
Scale Factor: 0.999750
Latitude of Origin: 4.666667
Linear Unit: Meter (1.000000)

Geographic Coordinate System: GCS_Leigon
Angular Unit: Degree (0.017453292519943299)
Prime Meridian: Greenwich (0.000000000000000000)
Datum: D_Leigon
Spheroid: Clarke_1880_RGS
Semimajor Axis: 6378249.144999999600000000
Semiminor Axis: 6356514.869549775500000000
Inverse Flattening: 293.464999999999970000

As a complement, the Well feature class also has longitude and latitude data fields to store original geographic coordinates.

Table 4-2 – List of WRC Water Resources Database domains

Name	Description
AHBoolean	Pseudo-boolean field
AquiferFType	Feature classification for aquifer contexts
AquiferName	Name of aquifer contexts
BasinName	Name of standard drainage areas named after main rivers
BhMethod	Borehole drilling method
BoreholeLogType	Classification of borehole logs
DsType	Type of time-enabled dataset
ElevationRange Domain	Range for elevation values
ElevUnits	Units for elevation values
FGeometry	Feature geometry
GPSReading	Type of locations for GPS readings
HasDpt	Classification of records based on availability of well depth values
HasDupl	Classification of records based on the existence of duplicate records
HasGwq	Classification of records based on availability of gw quality info
HasLth	Classification of records based on availability of lithology information
HasSwl	Classification of records based on availability of gw level values
HasXY	Classification of records based on coordinates availability/reliability
HasYld	Classification of records based on availability of yield values
HGUCode	Hydrogeologic unit code
HGUGeneral	Generalized hydrogeologic unit intercepted by well
Lithostratigraphy	Lithostratigraphic/lithodemic units
LogCodeBA	Borehole annulus classification
LogCodeBC	Borehole completion classification
LogCodeBD	Borehole diameter classification
LogCodeBL	Names of common lithologic units
LogCodeSM	Log sample medium
MonitoringPointFType	Feature classification for MonitoringPoint
MType	Type of groundwater level measurement
PtLoc	Location of gw level measurements associated to pumping test
PumpType	Type of pump installed
RefPt	Reference point for groundwater level measurement
RegionName	Name of administrative region
SmplMedium	Sample medium for time series
TimeUnits	Time units
TSDataType	Data type for TimeSeries
VarUnits	Units for time series variables
WellState	Classification of well state or well condition
WellType	Classification for well type
WellUse	Classification for well use

5. DATA INTEGRATION

This section summarizes the operations involved in the integration of relevant water-related data collected under the HAP into the WRC Water Resources Database. For future data integration, a complementary document entitled 'Water Resources Database Operation Manual' (SLI-INRS, 2011) was prepared. In addition to detailing data integration operations related to the database, this document provides guidelines for data analysis, extraction and representation.

5.1 Groundwater data

5.1.1 Preparation of data to be integrated

Prior to data integration, different operations were required to ensure collected groundwater data complied with the WRC Water Resources Database structure and validation features. The main operations included 1) assigning a temporary identifier to each feature to be integrated, 2) restructuring and standardizing information according to database destination tables and domains and 3) adding complementary information from external data sources.

5.1.1.1 Temporary feature identification

An arbitrary but unique numeric identifier was temporarily assigned to each feature of the consolidated data (e.g. point features such as wells or meteorological stations) in order to relate them to their associated tabular data, which will be stored in different database tables (e.g. pumping test data for wells). Once the features were integrated in the database, this temporary identifier was replaced with the HydroID identifier generated by the Arc Hydro tools (see section 5.1.2 for HydroID assignation). This approach prevented the manual assignation of the HydroID prior to data integration, which can lead to database integrity errors and gaps in the identifier sequence of the HydroID. It however required the substitution of the temporary identifier in every feature class and table of the database after data integration.

5.1.1.2 Data restructuring and standardization

Subsequently, consolidated data were reorganized in order to be integrated into the database with the 'Text Import' tool in the AHGW Tools. This step notably involved restructuring sequential data such as lithology units, hydrostratigraphy units, screen intervals and temporal data to match the structure of the WRC Water Resources Database. To facilitate this step, MS Excel templates based on the database structure were developed as data entry forms. These templates include data validation features such as drop-down lists and error messages based on database domains to ensure consistency and facilitate the standardization of values. During this operation, it was necessary to generate additional data to complement the part of the existing information. In the case of groundwater quality data, this meant that vertical intervals from which groundwater samples originated (i.e. screen or open hole sections) had to be approximated, when possible, on the basis of the construction details of the well. And, in the case of borehole construction details, it meant that all HGU intercepted by wells (i.e. at screen or open hole sections) had to extracted.

5.1.1.3 Addition of complementary information

Finally, as the data collected under the HAP lacked various complementary information (e.g. land elevation at well location), additional data had to be extracted from different external data sources. This complementary information is summarized in Table 5-1. In the eventuality that new and/or more accurate data sources are obtained, this complimentary information should be updated. This notably concerns the information related to administrative districts that was known to be outdated at the time of writing. This information could however not be updated since electronic data related to new districts were not available yet. All complimentary information was appended to the consolidated data using spatial analysis functions in ArcGIS (i.e. spatial join for vector data and extract tool for raster data).

Table 5-1 – Complementary information appended to data before integration

Complementary information	Source	Destination data field in WRC database
Administrative region & district	SWERA, 2005	Region & District
Land elevation	SRTM data(1)	LandElev
Hydrological basin	Derived from SRTM data(2)	Basin
Regional gelogical unit	GSD, 2009	GeoUnit

Note:
(1) : Shuttle Radar Topographic Mission data obtained from International Centre for Tropical Agriculture through the Consortium for Spatial Information of the Consultative Group for International Agriculture Research (CGIAR-CSI) (CIAT, 2008).
(2) : Hydrological basins were delineated from SRTM elevation data in ArcGIS software with hydrology functions of the Spatial Analyst tools.

5.1.2 Data import into the database

Following data preparation, relations between data fields of the consolidated data and the WRC Water Resources Database were identified. This exercise allowed the consolidated data to be imported to the WRC Water Resources Database using the 'Text Import' tool in the AHGW Tools. The identification of these relations is documented in detail in Appendix B and is summarized in Table 5-2 for each groundwater-related feature class and table with the exception of the Aquifer feature class. As mentioned in section 4.1.1, the latter was used to store hydrogeological contexts delineated on the basis of the geology. Features corresponding to these hydrogeological contexts were created from the most recent geological map (GSD, 2009) obtained in shapefile format. Following the creation of polygons features and of the relevant attributes, topological errors were corrected according to rules defined in section 4.2.2.

Following data import to the WRC Water Resources Database, additional operations were required to complete the data fields used to link or describe data. First and foremost, a unique database identifier was assigned to each database feature using the *HydroID* data field. As mentioned in section 3.2.2, this data field is used to manage relationships between database features and tabular data. The complete procedure used to assign the *HydroID* is detailed in Appendix C. Subsequently, unique external identifiers were assigned to each well with reliable coordinates. This new ID, to be integrated in the *HydroCode* data field of the WRC Water Resources Database, was assigned using a vector grid based on the national borehole numbering scheme (see Appendix D for complete procedure to assign *HydroCode*). This step was necessary to provide a standardized external identifier as data collected under HAP came from various databases using different identification schemes for their records (N.B.: the original database identifier was however kept and stored in the BoreholeIndex table). Finally, variables used in the time series were also defined and described in the VariableDefinition table in order to provide basic information on the content and format of the data they contain. In a similar way, HGU were also defined in the HydroGeologicUnit table based on existing HGU transferred with lithology data. Relevant time series were also described in the SeriesCatalog table.

Table 5-2 – Summary of data field correspondence between HAP consolidated data and WRC Water Resources Database

Framework component

HAP field	WRC object class	WRC field
VillageID	Well	CommID
Village	Well	Comm
DistrictID	Well	DistrictID
District	Well	District
RegionID	Well	RegionID
Region	Well	Region
Basin	Well	Basin
Z_DEM	Well	LandElev
Longitude	Well	Longitude
Latitude	Well	Latitude
GPS	Well	GPSReading
Sit_Consul	Well	SitingComp
Date_Sit	Well	SitingDate
Bh_Company	Well	BhComp
Bh_Method	Well	BhMethod
Date_Bh	Well	WellDate
Well_State	Well	WellState
Well_Type	Well	WellType
Well_Use	Well	WellUse
Well_Depth	Well	WellDepth
Airlift_Yld	Well	AirliftYld
Pump_Type	Well	PumpType
Pump_Depth	Well	PumpDepth
Date_Pump	Well	PumpDate
Date_Entry	Well	EntryDate
Comments	Well	CommentGen
Location_	Well	CommentLoc
Dtset_XY_	Well	HasXY
Dtset_Dpt_	Well	HasDpt
Dtset_Lth_	Well	HasLth
Dtset_Swl	Well	HasSwl
Dtset_Yld_	Well	HasYld
Dtset_Wq_	Well	HasGwl
Dupl_	Well	HasDupl
Group	Well	GeoUnit

HAP field	WRC object class	WRC field
WellID_1	BhIndex	WellID1
WellID_2	BhIndex	WellID2
WellID_3	BhIndex	WellID3
WellRefNo	BhIndex	WellRefNo
Prjct_Name	BhIndex	ProjctName
Prjct_Fund	BhIndex	ProjctFund
Source_1	BhIndex	DSource1
Source_2	BhIndex	DSource2
Source_3	BhIndex	DSource3
Cell_ID_	BhIndex	GridID
Order_	BhIndex	HAPID

Borehole component

HAP field	WRC object class	WRC field
Lith1Unit	BhLithology	LogCode
Lith1Bot	BhLithology	ToDepth
Lith1Desc	BhLithology	LogDescr
Lith2Unit	BhLithology	LogCode
Lith2Bot	BhLithology	ToDepth
Lith2Desc	BhLithology	LogDescr
Lith3Unit	BhLithology	LogCode
Lith3Bot	BhLithology	ToDepth
Lith3Desc	BhLithology	LogDescr
Lith4Unit	BhLithology	LogCode
Lith4Bot	BhLithology	ToDepth
Lith4Desc	BhLithology	LogDescr
Lith5Unit	BhLithology	LogCode
Lith5Bot	BhLithology	ToDepth
Lith5Desc	BhLithology	LogDescr
Lith6Unit	BhLithology	LogCode
Lith6Bot	BhLithology	ToDepth
Lith6Desc	BhLithology	LogDescr
Lith7Unit	BhLithology	LogCode
Lith7Bot	BhLithology	ToDepth
Lith7Desc	BhLithology	LogDescr
Lith8Unit	BhLithology	LogCode
Lith8Bot	BhLithology	ToDepth
Lith8Desc	BhLithology	LogDescr
Lith9Unit	BhLithology	LogCode
Lith9Bot	BhLithology	ToDepth
Lith9Desc	BhLithology	LogDescr
Hydst1Bot	BhLithology	ToDepth
Hydst1Unit	BhLithology	HGUCode
Hydst2Bot	BhLithology	ToDepth
Hydst2Unit	BhLithology	HGUCode
Hydst3Bot	BhLithology	ToDepth
Hydst3Unit	BhLithology	HGUCode
Hydst4Bot	BhLithology	ToDepth
Hydst4Unit	BhLithology	HGUCode
Hydst5Bot	BhLithology	ToDepth
Hydst5Unit	BhLithology	HGUCode
Hydst6Bot	BhLithology	ToDepth
Hydst6Unit	BhLithology	HGUCode
Hydst7Bot	BhLithology	ToDepth
Hydst7Unit	BhLithology	HGUCode
Hydst8Bot	BhLithology	ToDepth
Hydst8Unit	BhLithology	HGUCode
Hydst9Bot	BhLithology	ToDepth
Hydst9Unit	BhLithology	HGUCode

HAP field	WRC object class	WRC field
Hole1Diam	BhConstruction	LogDescr
Hole2Diam	BhConstruction	LogDescr
Hole1Bot	BhConstruction	ToDepth
Hole2Bot	BhConstruction	ToDepth
CasingType	BhConstruction	LogCode
CasingDiam	BhConstruction	LogDescr
CasingLeng	BhConstruction	ToDepth
Screen1Top	BhConstruction	FromDepth
Screen1Bot	BhConstruction	ToDepth
Screen2Top	BhConstruction	FromDepth
Screen2Bot	BhConstruction	ToDepth
Screen3Top	BhConstruction	FromDepth
Screen3Bot	BhConstruction	ToDepth
Screen4Top	BhConstruction	FromDepth
Screen4Bot	BhConstruction	ToDepth
Screen5Top	BhConstruction	FromDepth
Screen5Bot	BhConstruction	ToDepth
Annulus1De	BhConstruction	LogCode
Annulus2De	BhConstruction	LogCode
Annulus3De	BhConstruction	LogCode
Annulus4De	BhConstruction	LogCode
Annulus5De	BhConstruction	LogCode
Annulus1Bo	BhConstruction	ToDepth
Annulus2Bo	BhConstruction	ToDepth
Annulus3Bo	BhConstruction	ToDepth
Annulus4Bo	BhConstruction	ToDepth
Annulus5Bo	BhConstruction	ToDepth

Time series component

HAP field	WRC object class	WRC field
Potability	SeriesGWQuality	Potability
Sample_ID	SeriesGWQuality	SampleName
Date_Sampl	SeriesGWQuality	SampleDate
Date_Analy	SeriesGWQuality	TsTime
Laboratory	SeriesGWQuality	Laboratory
Color	SeriesGWQuality	Colour
Turbidity	SeriesGWQuality	Turbidity
Temperatur	SeriesGWQuality	Temperatur
Conductivi	SeriesGWQuality	EC
Hardness	SeriesGWQuality	TH
Alkalinity	SeriesGWQuality	TA
Tot_Colif	SeriesGWQuality	TColi
E_Colif	SeriesGWQuality	EColi
pH	SeriesGWQuality	pH
EC	SeriesGWQuality	EC
As_	SeriesGWQuality	As_
Ca	SeriesGWQuality	Ca
Cl	SeriesGWQuality	Cl
CO2	SeriesGWQuality	CO2
CO3	SeriesGWQuality	CO3
Cu	SeriesGWQuality	Cu
F	SeriesGWQuality	F
Fe	SeriesGWQuality	Fe
HCO3	SeriesGWQuality	HCO3
K	SeriesGWQuality	K
Mg	SeriesGWQuality	Mg
Mn	SeriesGWQuality	Mn
Na	SeriesGWQuality	Na
NH3_n	SeriesGWQuality	NH3_N
NH4_n	SeriesGWQuality	NH4_N
NO2_n	SeriesGWQuality	NO2_N
NO3_n	SeriesGWQuality	NO3_N
Pb	SeriesGWQuality	Pb
PO4	SeriesGWQuality	PO4
SiO2	SeriesGWQuality	SiO2
SO4	SeriesGWQuality	SO4
TDS	SeriesGWQuality	TDS
Zn	SeriesGWQuality	Zn

HAP field	WRC object class	WRC field
Swl_1	SeriesGWLevel	Swl
SwlDatum_1	SeriesGWLevel	RefPtHgt
Date_Swl_1	SeriesGWLevel	TsTime
Swl_2	SeriesGWLevel	Swl
SwlDatum_2	SeriesGWLevel	RefPtHgt
Date_Swl_2	SeriesGWLevel	TsTime

HAP field	WRC object class	WRC field
Date_Pt	SeriesPumpTest	TsTime
Pt_Dura	SeriesPumpTest	PtTime
Pt_Swl	SeriesPumpTest	PtSwl
Pt_Dwl	SeriesPumpTest	PtDwl
Pt_Yld	SeriesPumpTest	PtYld
Recov_Dura	SeriesPumpTest	PtReTime
Recov_Wl	SeriesPumpTest	PtReSwl
Spec_Cap	SeriesPumpTest	SC
Transmiss	SeriesPumpTest	T
Storage	SeriesPumpTest	S

HAP field	WRC object class	WRC field
Sd_Yield_1	SeriesSDPumpTest	SdPtYld_1
Sd_Yield_2	SeriesSDPumpTest	SdPtYld_2
Sd_Yield_3	SeriesSDPumpTest	SdPtYld_3
Sd_Dura	SeriesSDPumpTest	SdPtReTime
Sd_Recov	SeriesSDPumpTest	SdPtReSwl

5.2 Surface water data

5.2.1 Preparation of data to be integrated

As for groundwater data, different operations were required to modify and adapt surface water data before their integration to the WRC Water Resources Database. This involved: 1) combining related data from multiple files, 2) assigning a temporary identifier to each feature to be integrated and 3) restructuring and standardizing information according to database destination tables and domains. These operations apply to hydrometric and rainfall data. Hydrographic data were not integrated into the WRC Water Resources Database as limitations were still associated to them at the moment of writing (see section 2.2).

5.2.1.1 Related data combination

Surface water data to be integrated in the WRC Water Resources Database were all provided in MS Excel format. However, a dataset for a given location (e.g. a river gauging station) was often provided as a collection of files with data covering different time periods. Consequently, for each surface water monitoring point (i.e. meteorological and river gauging stations), related data from multiple files had to be combined iwithin a single file.

5.2.1.2 Temporary feature identification

Following the combination of related data, an arbitrary but unique numeric identifier was temporarily assigned to each surface water feature (i.e. monitoring point features such as river gauging stations) in order to relate them to associated tabular data that will be stored in different database tables. As for groundwater data, this temporary identifier was replaced with the HydroID identifier generated by the Arc Hydro tools (see section 5.2.2) once the features were integrated into the database. As stated previously, this prevented the manual assignation of the HydroID prior to data integration, which can lead to database integrity errors and gaps in the identifier sequence of the HydroID.

5.2.1.3 Data restructuring and standardization

As a last step before integration, surface water data were reorganized in order to match the structure of the relevant WRC Water Resources Database tables and feature classes. As for groundwater data, MS Excel templates based on the database structure were developed as data entry forms to facilitate the restructuring and standardization of data. The templates for surface water data also include data validation features such as drop-down lists and error messages based on database domains. These features notably allowed the identification of the variable type related to the data stored (e.g. monthly versus daily rainfall data).

5.2.2 Data import into the database

Following the restructuring and standardization process, surface water data contained in the templates were imported to the WRC Water Resources Database using the 'Text Import' tool of the AHGW Tools. Since the surface water data types were fewer (i.e. mainly streamflow and rainfall data), data integration was more straightforward and the relations between source data and relevant data fields of the WRC Water Resources Database were not explicitly identified. As a general reference, relevant surface water data were integrated as follows:

- Hydrometric data in the SeriesRiverflow table
- Rainfall data in the SeriesRainfall table
- Analytical results from HAP rainwater samples in the SeriesRainwaterQuality table

Even though the relations are not identified with the source data, relevant data fields of the WRC Water Resources Database used to store all surface water data are documented in details in Appendix B. Coordinates and station type of each meteorological and river gauging stations were also imported into the MonitoringPoint feature class.

As for groundwater data, additional operations were required following data import in order to complete the data fields used to link or describe data in the database. First and foremost, a unique database identifier was assigned to each monitoring point feature using the *HydroID* data field. As mentioned in section 3.2.2, this data field is used to manage relationships between database features and tabular data. The complete procedure used to assign the *HydroID* is detailed in Appendix C. Unique external identifiers of surface water monitoring points do not need to be assigned since the identifiers provided with the source data are already unique. Consequently, the *HydroCode* value for all river gauging stations corresponds to the original identifier provided by the HSD while the *HydroCode* value for meteorological stations corresponds to the identifier provided by the MSD. Finally, as for groundwater data, variables used in the time series were defined and described in the VariableDefinition table in order to provide basic information on the content and format of the data they contain. Relevant time series were also described in the SeriesCatalog table.

5.3 Database content summary

The final database content at the time of writing is detailed in an electronic report provided in Appendix E. For groundwater related data, all available information was integrated in the database so that the final number of records in the Well feature class is 10 139, the same as that stated in section 2.1.3. It is however important to emphasize that only 7 874 of these were considered unique and reliable records (i.e. boreholes or hand-dug wells). All groundwater-related temporal data collected under HAP were also integrated in the database and are summarized in the SeriesCatalog table. For surface water related data, all the information collected and stated in section 2.2 was also integrated in the database. In total, the MonitoringPoint feature class comprises 74 river gauging stations (33 of which have hydrometric data) and 684 rainfall stations (56 of which have rainfall data).

6. CONCLUSIONS AND RECOMMENDATIONS

The Arc Hydro database model was used to create a groundwater and surface water database for the Northern Regions of Ghana (referred to as the *WRC Water Resources Database*). This database will notably improve the capacity of the WRC to collect, analyze, manage and disseminate groundwater and surface water data. The Arc Hydro database model was chosen because it was designed by experienced GIS and water analysts and because it is offered freely to the GIS community. Since it only offers a basis for the creation of a database, this model had to be modified and customized in order to meet the WRC data management needs. These modifications notably involved the creation or modification of database feature classes, domains and relationships.

Following the creation of the WRC Water Resources Database template, available groundwater data and part of the surface water data collected under the HAP or provided by the WRC were imported into this database using Arc Hydro tools. The import required available data to be reorganized and standardized. Additional data were also generated to complement the existing information.

While the resulting database provides a sustainable groundwater and surface water data repository for WRC, additional work remains to be done in order to improve the quality and quantity water data stored in the database. The sections below summarize additional database-related work identified to date in the course of the HAP as well as general guidelines proposed to manage the database in a sustainable way.

6.1 Additional work

6.1.1 Remaining data gaps and errors

6.1.1.1 Groundwater data

As mentioned in section 2.1, the data validation carried out on the hydrogeological data collected under the HAP was not exhaustive. While the resulting database content was deemed reliable enough to yield sound analysis results for the HAP, a significant number of records still have unreliable information (e.g. static groundwater level deeper than borehole depth) or lack essential information (e.g. coordinates). Data validation helped identify the main data gaps and obvious errors but work could still be done to complete the process. Remaining work could be divided in two: 1) completion of the validation process for existing data and 2) integration of additional data to existing records.

While the completion of the validation process is a time-consuming task, it can however be carried out at intervals by different individuals that have relevant experience if the progress is well documented. The main complementary data validation activities that could be carried out include: 1) correction of obvious attribute errors (notably keypunching errors in borehole coordinates or completion date), 2) verification of incoherent attribute information (e.g. borehole completion date later than borehole sampling date) and 3) identification of duplicates. Access to original reports and/or borehole logs will be required to carry out these activities and validate the information in the database. The identification of attribute errors or incoherent attribute information has to be carried out one record at a time. For the identification of duplicates, original borehole logs will be required in order to validate the information for records identified as possible duplicates[9]. Records identified as unique or duplicate[10] records do not have to be verified again.

[9] Records identified as possible duplicates in the *HasDupl* data field are records that have the same coordinates and the same borehole depth as one of the unique records.
[10] Records identified as duplicates in the *HasDupl* data field are records that have the same coordinates, the same borehole depth, the same static groundwater level and the same borehole yield as one of the unique records.

During the data validation process described in section 2.1, essential information was found to be missing for a number the records thus rendering them unusable for some analyses and creating data gaps in terms of spatial distribution. Integration of additional information for these existing records could convert a significant number of unreliable records into reliable records. The following data fields identify which records are missing essential hydrogeological information: *HasXY* (coordinates), *HasDpt* (depth), *HasLth* (lithology), *HasSwl* (static gw level), *HasYld* (yield) and *HasGwq* (gw quality). Again, access to original reports and/or borehole logs will be required to find the missing information (if it exists). While the data fields listed above are to be considered in priority (especially coordinates), other missing information such as borehole construction details could also be integrated in the database when available (notably useful to identify the depth of the screen and the nature of the material intercepted by the screen).

6.1.1.2 Surface water data

The surface water data collection revealed that many hydrometric and meteorological datasets contain significant data gaps. While this is not uncommon for long duration datasets in Northern Ghana, it limits the potential use of the data as well as the reliability and representative nature of analyses that can be carried out with those data. Consequently, it would greatly benefit future analyses related to surface water development and management if missing values from these datasets could be estimated from other available data. Such an activity could be outsourced to consultants or could be carried out as part of a graduate student project with careful supervision.

6.1.2 Additional modifications to existing data

Some of the existing data integrated into the WRC Water Resources Database could benefit from additional modifications in order to improve analysis capacity, efficiency and/or reliability. This notably concerns hydrogeological contexts (polygons) and hydrographical features (lines and polygons).

6.1.2.1 Groundwater data

As stated in section 4.1.1, the hydrogeological contexts of the <u>Aquifer feature class</u> were delineated on the basis of geology. A more representative regional aquifer definition should be carried out in the future as additional data become available. Aside from geology, such an exercise could notably be based on hydraulic parameters such as transmissivity, yield and specific capacity, and be complemented by hydrogeochemistry data.

6.1.2.2 Surface water data

The hydrographical features, i.e. rivers, streams and water bodies, which were provided by the Ghana Survey Department (GSD), were not integrated in the WRC Water Resources Database. As stated in section 2.2, these hydrographic data contain several topological errors that could not be corrected within the project timeframe. Consequently, additional work would be required to remove these topological errors (e.g. overlapping of polygons or lines) and to create a flow network to model water flow out of the available data. The creation of such a network is supported by the Arc Hydro database model. While it requires a significant effort to build, a flow network would notably improve spatial analysis reliability, facilitate representation of hydrographical features and delineation of drainage areas, allow upstream/downstream tracing and help in the construction of surface water models.

6.1.3 Integration of complementary data

One of advantages of using the geodatabase as database format is that it has the additional capability to store geospatial features. Consequently, any complementary information that is relevant to the analysis and management of groundwater and surface water data for the WRC can be stored in the database.

While project constraints did not allow such work to be carried out under the HAP, future work by the WRC or subcontracted by the WRC could include the integration of the following thematic data: administrative limits, communities, infrastructures (e.g. roads, railroads, power lines), land use, vegetation, geology and topography. In the case of geology, the Arc Hydro database model used to develop the WRC Water Resources Database provides templates for the storage and management of geological features (point, lines or polygons). These templates notably include a data field that allows a link between geological features and defined hydrogeologic units (HGU) for analysis purposes.

Other complementary data such as climate data could also be integrated in the WRC Water Resources Database for analysis and representation purposes. While rainfall has already been integrated under the HAP, other climates variables such as temperature, relative humidity, sun hours and wind speed could also be integrated in order to allow the periodic update of the water balance calculation. These climate variables could be integrated into the database as time series (e.g. in a table based on the AttributeSeries table template) and linked to meteorological stations already integrated in the MonitoringPoint feature class.

6.2 Guidelines for database management

The lifetime and effective usefulness of the WRC Water Resources Database will mainly depend on the quality of the data it contains and on the resources allotted for efficient database management. Resources dedicated to the database (i.e human, financial, technical) are not thoroughly discussed here as they are the object of other HAP activities. As for the quality of the database content, previous sections proposed additional work to improve the quality of existing data. In complement, this section provides general guidelines for future data input and database management in order to ensure data quality and integrity.

Database management can be greatly facilitated by standardizing data entry at the source (i.e. data provided by NGOs, donors, contractors …). Standardization of data entry could be done through the distribution and use of the WRC Water Resources Database template. This template (in ArcGIS or Microsoft Access format) could be distributed through the WRC website or along with drilling permits issued by WRC. The use of the template by the different stakeholders in future projects (e.g. water supply, research, monitoring …) would notably reduce pre-formatting work required of WRC personnel for data integration in the WRC Water Resources Database. This template is provided in Appendix E along with an electronic report documenting the database schema. If another database structure is used by a stakeholder, adequate documentation of the alien database content and structure should ideally accompany the data. For projects where a formal database is not used or required, Microsoft Excel templates derived from the WRC Water Resources Database could be used. These templates, which are provided in Appendix E, could also be distributed through the WRC website or along with drilling permits issued by WRC. They would provide a data structure compatible with the database as well as ensure minimal data validation and standardization which would also help the data integration process at the WRC.

Standardization of data integration operations at the WRC can also help achieve a more efficient database management. To support the WRC in this task, basic guidelines for the integration of new data were defined in a complementary document titled 'Water Resources Database Operation Manual' (SLI-INRS, 2011). This document provides an insight of data integration operations for WRC personnel and should contribute to a more efficient database management. In addition to the guidelines provided in this document, a few general recommendations are provided below in order to facilitate data integration and subsequent database management:

- Assign HydroCode before data entry
- Verify units for numeric data fields before data entry

- Verify that the type of data being imported (e.g. string, integer, double, date …) match the data type and the domain (if there is one) of the data field that will receive the data
- Modify data to be entered in numeric fields if required (e.g. assign -9999 as the NoData value, use negative numbers for groundwater quality values under detection limit …)
- Use the 'Text Import' from Arc Hydro tools to input data when possible
- Follow the database workflow when creating/importing new data to ensure data integrity (e.g. do not enter groundwater quality data without entering the well from which the groundwater quality sample was taken)
- Update the HydroID for all the database after each data entry
- Assign complementary info to newly imported data (e.g. aquifer ID, Basin, GeoUnits …)
- Create 3D features such as BoreLines after importing new data (if applicable)

7. REFERENCES

Ghana Geological Survey Department (2009) Geological Map of Ghana – Scale 1:1 000 000. Published and edited the Ghana Geological Survey Department (Accra, Republic of Ghana) in cooperation with the Bundesanstalt für Geowissenschaften und Rohstoffe (Hannover, Federal Republic of Germany).

International Centre for Tropical Agriculture (CIAT) (2008) Hole-filled seamless SRTM data v4. Consultative Group for International Agriculture Research - Consortium for Spatial Information (CGIAR-CSI), data from http://srtm.csi.cgiar.org.

Maidment, D.R. (2002) Arc Hydro: GIS for Water Resources. Redlands, California: ESRI Press.

SNC-Lavalin International (SLI) and INRS-ETE (SLI-INRS) (2011) Hydrogeological Assessment of the Northern Regions of Ghana - Water Resources Database Operation Manual. March 2011: 95 p.

Solar and Wind Energy Resource Assessment (2005) Ghana Geospatial Toolkit Data. UNEP & Global Environment Facility Project, data from http://swera.unep.net.

Strassberg, G., Maidment, D.R., Jones, N.L. (2007) A Geographic Data Model for Representing Ground Water Systems. Ground Water 45 (4): 515-518.

Strassberg, G. (2008) Arc Hydro Groundwater Data Model. ESRI 2008 User Conference, Pre-conference seminars, August 2-3, 2008. Slideshow, 46 p.

World Health Organization (2008) Guidelines for drinking-water quality: incorporating 1st and 2nd addenda, volume 1, Recommendations. WHO publication, 3rd Edition, 668 p.

Zeiler, M. (1999) Modeling Our World: The ESRI Guide to Geodatabase Design. Redlands, California: ESRI Press.

Prepared by:

Marc-André Carrier
Hydrogeologist, GIS/Mapping Advisor
Date: July 2011

Reviewed by:

Rene Lefebvre
Technical Capacity Building Team Leader
Date: November 2011

Enoch Asare
WRC Chief Hydrogeologist
Date: December 2011

Appendix A

HYDROGEOLOGICAL ASSESSMENT OF THE NORTHERN REGIONS OF GHANA

PRELIMINARY ASSESSMENT OF AVAILABLE ELECTRONIC DATABASES

August 2006

SNC·LAVALIN International In joint venture with Université du Québec Institut national de la recherche scientifique Eau, Terre et Environnement

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

1 Introduction

Among the main objectives of the Hydrogeological Assessment Project of the Northern Regions of Ghana (HAP) is the establishment of the basis for current knowledge of the hydrogeological setting. One of the first steps to achieve this consists of assessing the content of the electronic hydrogeological databases made available by stakeholders and to eventually merge them into a consolidated database that will serve future hydrogeological projects in Northern Ghana as well as HAP downstream activities.

The main objectives of this preliminary assessment are to identify 1) the unique records among all available databases, 2) the reliable records among these unique records (N.B.: reliability based on the location data) and 3) the resultant data gaps within the unique and reliable records. The first objective is aimed at eliminating redundant information in view of a database consolidation while the second and the third are aimed at establishing the need (if any) for additional data acquisition. It is important to mention that this assessment is preliminary and is not to be held as thorough. In some cases, only parts of electronic databases were available at the time of the assessment and in all cases, a lot of the available data was left unverified. Only specific verifications that yielded results considered critical for the data consolidation process to go on were carried out. A more complete assessment will be done near the end of the database consolidation process in order to re-assess the data situation and confirm that minimum data requirements are met.

2 Available electronic databases

Through the key stakeholders, the six following hydrogeological electronic databases were obtained:

- Agence Française de Développement (AFD) database (MS Excel file)
- Canadian International Development Agency (CIDA) database (MS Excel file)
- European Union (EU) database (MS Excel file)
- Global Change in the Hydrological Cycle Project (GLOWA) database (MS Access file)
- Water Research Institute (WRI) database (Ground Water for Windows (GWW) file)
- World Vision (WV) database (Ground Water for Windows (GWW) file)

The first three databases, obtained from the Community Water Supply and Sanitation Agency (CWSA), contain only new records[1] created during their respective projects. All of these projects were carried out mainly in the Northern Region. As of the time of this assessment, only part of the expected data was available for these three databases as they originated from active projects that were not yet completed. It is expected that the additional data will be forthcoming by the end of 2006.

The GLOWA database contains records collected from different sources, notably contractors, regional CWSA offices and the WRI. Although research conducted under the GLOWA project mostly concerns the Volta Basin, this electronic database contains records for all of Ghana.

The WRI database is considered the official hydrogeological electronic database for the Northern Regions of Ghana. It comprises records from many projects (e.g.

[1] In this document, unless other wise specified, a record of a specified database refers to a well with all its associated data (i.e. descriptive attributes).

Community Water Project (COWAP)) carried out in the Upper East, Upper West and Northern Regions.

Finally, the WV database contains available records created for World Vision projects. Some of the WRI records for the Northern Region were also appended to this WV database.

3 Methodology

The format and size of most of the available databases made it possible to carry out the analysis and queries in ArcGIS 9.0. For the GWW format files (i.e. WRI and WV databases), some manual editing was required before transfer into ArcGIS as the output format of GWW is an ASCII text file. The use of a Geographical Information System (GIS) such as ArcGIS was necessary for this preliminary assessment since spatial analysis functions were required.

All available databases were first examined to determine the total number of records. The assessment then began with the identification of the unique records in each database. The evaluation of location data reliability (i.e. the coordinates) followed. Finally, statistics were calculated for each of the selected data fields. Although the electronic databases available were analysed differently with respect to their content, the general procedure is described below for each objective.

3.1 Unique records

The identification of unique records did not require the same method for all databases as information was sometimes available concerning the origin of data. First, the records from the AFD, EU and CIDA databases were all considered unique since they came from recent projects and were unlikely to have been entered (and thus duplicated) in any other database yet. Consequently, verification of record uniqueness was not undertaken for these databases. For the GLOWA and WV databases, it was possible to determine the presence of redundant records through queries. It was determine that such records were all originally taken from the WRI database. Consequently, all records in the WRI were considered unique while redundant records in the GLOWA and WV databases were ignored. Different methods were used in identifying redundant records between GLOWA & WRI and WV & WRI.

3.1.1 GLOWA database

The identification of unique records in the GLOWA database was done by relating records of the GLOWA database with records of the WRI database. The creation of reliable link between the databases required the use of two key data fields (Well ID and Project ID) and the correction of syntax errors in the Project ID field. The use the Project ID field was required since many records present in both GLOWA and WRI databases had different Well IDs although they were clearly the same wells. On the other hand, the use of the Well ID field was also required because some records present in both databases had the same Project ID, which is to be expected since many wells could have drilled during the same project. Unfortunately, the use of both data fields was not always sufficient for identification and visual inspection based on other data fields (e.g. community name and well completion date) had to be done to identify some common records. The majority of the redundant records were however identified following these steps:

1) Creation of a new temporary data field with the corrected project number
2) Determination of common records between the two databases based on Project IDs and Well IDs

3) Creation of a new data field to store a unique record identifier (unique record = 1 and common record = 0 (N.B.: data from WRI database was kept for common records))

3.1.2 World Vision database

The identification of unique records in the WV database was simpler since the Project IDs of the World Vision projects were known (personal communication with Enoch Asare from WRC, 2006). Therefore, records with the following project suffixes were identified as unique records: UNICEF, OIC, and WV.

3.2 Reliable records

Prior to the verification of the location data, regions and districts names in the databases had to be corrected for syntax errors. The table 1 presents the region and district names used. The official and updated districts names for the Northern Regions (available at the following URL http://ghanadistricts.com/home) were not used in this assessment because the associated file giving the location of each official district was not available at the moment of this assessment.

Table 1: Region and district names used

Region	District
Northern	Bole
Northern	East Gonja
Northern	East Mamprusi
Northern	Gushiegu Karaga
Northern	Nanumba
Northern	Saboba Chereponi
Northern	Savelugu Nanton
Northern	Tamale
Northern	Tolon Kumbungu
Northern	West Gonja
Northern	West Mamprusi
Northern	Yendi
Northern	Zabzugu Tatale
Upper East	Bawku East
Upper East	Bawku West
Upper East	Bolgatanga
Upper East	Bongo
Upper East	Builsa
Upper East	Kassena Nankana
Upper West	Jirapa Lambussie
Upper West	Lawra
Upper West	Nadowli
Upper West	Sissala
Upper West	Wa

For this preliminary assessment, the evaluation of reliability was carried out only for the data field containing the well coordinates (i.e. longitude and latitude). Because the original location data (e.g. paper logs or GPS datasheet) was not yet available, coordinates reliability was mainly evaluated using spatial analysis functions with respect to administrative boundaries from an independent data source[2]. The presence of syntax errors in coordinates and the inaccuracy of both coordinates and

[2] At the moment of writing, the most reliable data source for regions and district boundaries was considered to be the Solar and Wind Energy Resources Assessment (SWERA) Project.

administrative boundaries made the use of decision trees appropriate for this task. Records identified as unreliable were kept for future and more thorough verifications (with original location data if possible). It is important to mention upfront that this was a time consuming task and that, therefore, it was not carried out with the same level of detail for all databases. Priority was given to the WRI database since it was considered to hold the larger number of unique records. The GLOWA and WV databases were also verified for location data reliability but to a lesser extent. As for the AFD, EU and CIDA databases, only minor verifications were made since many records were missing coordinates (N.B.: updated versions of these databases with coordinates for all records are expected by the end of 2006). While a complete analysis of all databases could have yielded a greater number of reliable records, the effort needed to accomplish this was considered disproportionate at that time in regards to the results that would be obtained.

3.3 Data gaps for selected data fields

For the HAP purposes, the following data fields were considered to be the minimum data requirements[3] to carry out the necessary analysis: 1) Well state, 2) Well depth, 3) Weathered layer thickness, 4) Lithology, 5) Groundwater level. 6) Yield, 7) Water quality.

The identification of data gaps for these data fields is twofold: 1) identification of gaps in terms of data quantity and 2) identification of data gaps in terms of spatial distribution of data. The first part was simply done with the help of statistics. Records containing information for each of these data field were compiled in order to evaluate the quantity of data available for each data field. The records resulting from this compilation were then plotted for each data field to assess their spatial distribution. This was done with regards to a 15kmx15km cell grid that was considered the minimum requirement in terms of data distribution for HAP purposes (i.e. at least one borehole must be present in each cell – minimal density of 1 borehole per 225 km²). The dimensions selected for the grid cell size is comparable to the size used in similar regional studies.

4 Results

Table 2 presents the total number of records for each database. The sum of these total records (15,092 records) does not give a representative idea of the amount of data that can actually be used for hydrogeological analysis. It is also important to mention that the total appearing in this table for the GLOWA database represents only the number of records relevant to the Northern Regions. A subset of records had to be selected[4] since the GLOWA database includes records for all of Ghana.

4.1Unique records

4.1.1 AFD, EU, CIDA databases

As mentioned previously, records in theses databases were all considered unique as they come from ongoing projects and were therefore considered to not be duplicated among the database examined. Consequently, the numbers of unique records, which also correspond to the total number of records for each database, are 231 for AFD, 483 for EU and 859 for CIDA.

[3] These minimum data requirements only apply to hydrogeological data found in the available databases; other required data for HAP, such as meteorological data, are not included in these requirements.

[4] The selection was based on the Region data field; the actual total number of records that are available for all of Ghana in the GLOWA database is 15212.

Table 2: Total number of records for available databases

Database	Number of records
Agence Française de Développement (AFD)	231
Canadian International Developement Agency (CIDA)	859
European Union (EU)	483
Global Change in the Hydrological Cycle Project (GLOWA)	6571
Water Research Institute (WRI)	5984
World Vision (WV)	964
Total	15092

4.1.2 GLOWA database

The identification of unique records for the GLOWA database was done with respect to the WRI database, for which all records are considered unique as mentioned earlier. The analysis yielded 1,406 unique records out of a total 6,571 records for the GLOWA database. The 5,165 other records in the GLOWA database are thus common to the WRI database. A quick examination of these 5,165 records revealed that, although they contain data common to both GLOWA and WRI databases, there is also specific data unique to each databases for some records (e.g. for one particular common record, the thickness of the weathered layer might be available in the GLOWA database while nonexistent in the WRI database). Although records from the WRI database were considered over the ones of GLOWA, a closer examination of redundant records in the GLOWA database should be carried out to extract the additional information in view of the consolidation process. The details of the analysis are presented below in table 3.

Table 3: Unique and common records for the GLOWA database

GLOWA (compared to WRI)	Records	Status
Common Well ID and Project ID	3881	Common
Common Project ID only (1)	1220	Common
Common Well ID only (2)	64	Common
No data fields in common (3)	37	Unique
No data fields in common	1369	Unique
Total	6571	

Notes:

(1) : the Project ID (and other data fields) were common but Well ID was different (N.B.: it was assumed that Well IDs were changed for specific project purposes)

(2) : the Project ID was missing for these records

(3) : all data fields were different but the Project ID of these records existed in both database (N.B.: this situation may arise if different wells of a same project were entered in the two databases analysed)

4.1.3 WRI database

Records from the WRI database were all considered unique as it is the reference database. It is however important to mention that no verification was done to identify record duplicates during this assessment. The number of unique records, which in this case also corresponds to the total number of records, is 5,984.

4.1.4 WV database

Unique records from the World Vision database were identified on the basis of the Project ID. The query made for the Project IDs identified as unique WV projects returned the following number of unique records: UNICEF Project (98), OIC Project (39) and WV Project (240). The total number of unique records is thus 238 out of a total 964 records in this database.

4.2 Reliable records with respect to location data

4.2.1 AFD, EU, CIDA databases

Location data (i.e. latitude and longitude) of the EU, AFD, CIDA databases were generally considered reliable if they fell within the Northern Regions. Minor verifications were done for these databases, notably to find and correct syntax errors and to identify records with coordinates falling slightly outside Northern Regions but still relatively near to their corresponding district[5]. A complete and more thorough assessment of location data should be carried out when all coordinates will be available for these databases. Meanwhile, the following records were found to have reliable coordinates: 82 out 231 for AFD, 397 out of 483 for EU and 435 out of 859 for CIDA.

4.2.2 GLOWA database

The decision tree used to assess location data reliability is illustrated in figure 1. Results, which are also shown on that figure, reveal that 6,136 records apparently had reliable coordinates while 435 records were flagged as unreliable. In order to make all coordinates reliable, access to original data sheets or additional field work (i.e. GPS survey) will be necessary (N.B.: 433 out of these 435 records don't have any coordinates).

4.2.3 WRI database

The location data reliability analysis carried out on this database is presented in figure 2 along with the results. From this preliminary assessment, there are 4,498 records with apparently reliable coordinates and 1,486 records that either had unreliable coordinates or were missing coordinates. Although further analysis could help reduce the latter number, it is considered that the amount of work needed would be disproportionate in regards of the results expected. This is notably explained by the problems arising from the use of community names to conduct further analysis on coordinates reliability. Such problems include:

- syntax errors in community names (manual corrections);
- absence of communities in one of the database (manual update of database);
- presence of new communities in one of the database;
- coordinate discrepancies for the same community.

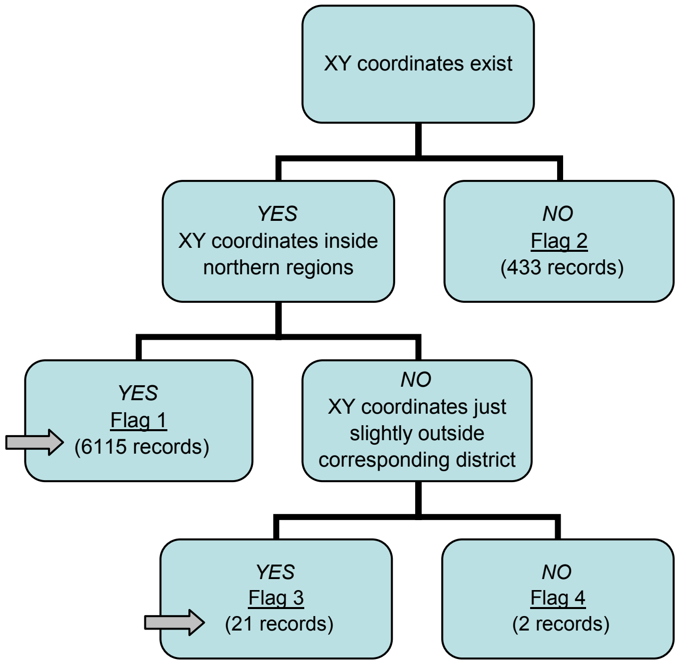
Consequently, and depending on HAP needs, further coordinate verification could be done on a limited number of these 1,486 records. In any case, access to original data sheets or additional field work (i.e. GPS survey) will probably be necessary to make all coordinates reliable.

4.2.4 WV database

Results show 898 records with apparently reliable coordinates and 66 records that either had unreliable coordinates or were missing coordinates (N.B.: 21 out of these 66 records don't have any coordinates). Obvious longitude errors (i.e. East vs West direction) were corrected and 'flagged' reliable with a short description of the correction. Figure 3 illustrates the results.

[5] The problem of coordinates falling outside the Northern Regions but near their corresponding district can be attributed to inaccuracy of administrative boundaries or of coordinates themselves.

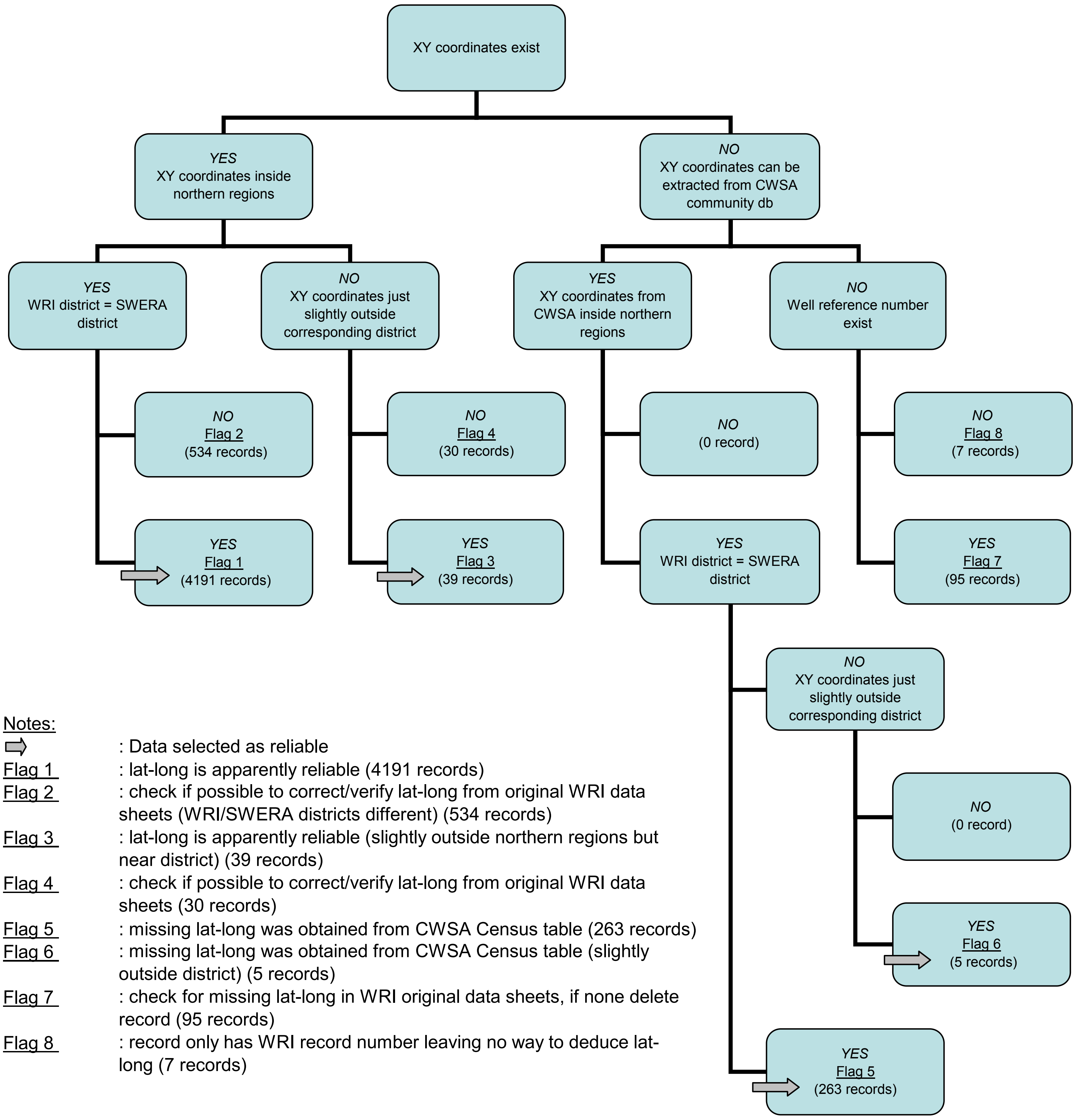
Figure 1: Location data reliability assessment (GLOWA database)



Notes:

	: Data selected as reliable
Flag 1	: lat-long is apparently reliable (6115 records)
Flag 2	: missing lat-long - check if possible to obtain from other data sources (433 records)
Flag 3	: lat-long is apparently reliable (slightly outside Northern Regions but near district) (21 records)
Flag 4	: check if possible to correct/verify lat-long from other data sources (2 records)

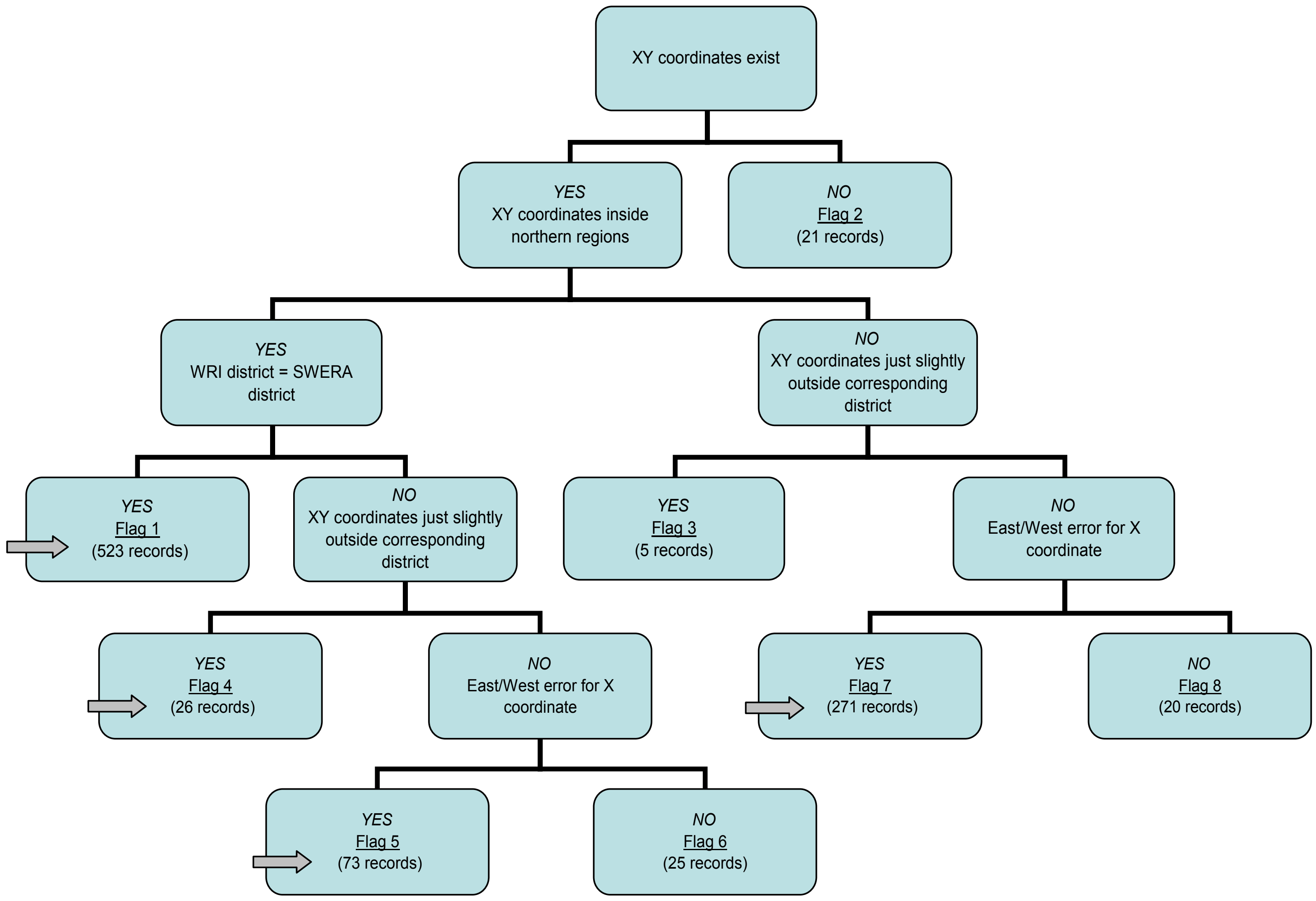
Figure 2: Location data reliability assessment (WRI database)



Notes:

⇨	: Data selected as reliable
Flag 1	: lat-long is apparently reliable (4191 records)
Flag 2	: check if possible to correct/verify lat-long from original WRI data sheets (WRI/SWERA districts different) (534 records)
Flag 3	: lat-long is apparently reliable (slightly outside northern regions but near district) (39 records)
Flag 4	: check if possible to correct/verify lat-long from original WRI data sheets (30 records)
Flag 5	: missing lat-long was obtained from CWSA Census table (263 records)
Flag 6	: missing lat-long was obtained from CWSA Census table (slightly outside district) (5 records)
Flag 7	: check for missing lat-long in WRI original data sheets, if none delete record (95 records)
Flag 8	: record only has WRI record number leaving no way to deduce lat-long (7 records)

Figure 3: Location data reliability assessment (WV database)



Notes:

⇨ : Data selected as reliable

Flag 1 : lat-long is apparently reliable (523 records)

Flag 2 : missing lat-long - check if possible to obtain from other data sources (21 records)

Flag 3 : lat-long is apparently reliable (slightly outside northern regions but near district) (5 records)

Flag 4 : lat-long is apparently reliable (slightly outside district but near district) (26 records)

Flag 5 : lat-long is apparently reliable (after longE correction) (73 records)

Flag 6 : check if possible to correct/verify lat-long from other data sources (25 records)

Flag 7 : lat-long is apparently reliable (after longE correction) (271 records)

Flag 8 : check if possible to correct/verify lat-long

4.3 Data gaps for selected data fields

Following the analysis for unique and reliable records (reliability of location data only), statistics were calculated for each database. Table 4 shows that the total number of unique records is 9,340 and that 7,056 of these records are considered reliable (as far as location data is concerned). The last column reveals that there are major gaps in the lithology and weathered thickness[6] data fields for all databases except the WV database. For most databases, other major gaps are also found in the following data fields: groundwater level, yield and water quality.

Table 5 presents a summary of the previous table. From these results, it is plain that the number of records that can actually be used for hydrogeological analysis (i.e. last row of the table) is largely insufficient for HAP purposes.

While the tables revealed the obvious need for additional data, it could not tell where data is most needed geographically. Therefore, a verification of the spatial distribution of the 71 records identified above revealed that only 26 cells have one or more boreholes in them (figure 4). Considering that 507 cells (15x15km) are necessary to cover the entire study area, it is obvious that additional data is needed. The major spatial data gaps are easily spotted on figure 4 (i.e. basically every cell without a yellow dot, so all of the Northern and Upper West Regions and parts of the Upper East Region).

5 Conclusion

The results of this assessment indicate that additional data is required, both in terms of quantity and in terms of spatial distribution, over and above what is currently available in electronic databases in the sector. In the context of the HAP, it was proposed that a subset of 450-500 wells (~1/225 km^2) with more reliable information be created. In order to build this subset, the access to original documents is crucial. The most efficient and reliable way to select these reliable wells is to go through the hard copies of available documents. The selection, which would have to be carried out or supervised by a local hydrogeologist, could be based of the following criteria:

- location (one well per 15km by 15 km cells);
- information available (more than one well per cell could be selected if data requirements can not be met with only one well);
- contractor/consulting engineers (local knowledge of data reliability with respect to contractors/consulting engineers will help in the selection);
- representativeness (the selected well(s) would have to represent the average conditions encountered in the cell - this can be based on borehole logs inspection);
- well status (selected wells would have to be active/usable to allow for possible water level measurement for example).

It is without any doubts, a long process to go through but the resulting subset would represent a significant contribution to the hydrogeological database of the Northern Regions.

[6] Although weathered thickness is considered a required data field, it is not as critical as others since access to reliable and detailed lithological information can generally be used to define the limits of the weathered layer.

Table 4: Statistics for selected data fields in all databases

Database / Field	Records				
	Total	XY reliable	Unique	Unique/XY reliable	Unique/XY reliable (%)
WRI	5984	4498	5984	4498	100%
Well state	5919	4453	5919	4453	99%
Well depth	2890	1747	2890	1747	39%
Weath. layer thickness	0	0	0	0	0%
Lithology	374	333	374	333	7%
Groundwater level	478	388	478	388	9%
Yield [(2)]	2421	2117	2421	2117	47%
Water quality [(3)]	175	145	175	145	3%
GLOWA	6571	6136	1406	1296	100%
Well state	6560	3959	1366	1261	97%
Well depth	6183	5822	1192	1122	87%
Weath. layer thickness	1112	1084	291	277	21%
Lithology [(1)]	3943	3911	511	494	38%
Groundwater level	4117	4076	655	628	48%
Yield [(2)]	3985	3959	417	406	31%
Water quality [(3)]	0	0	0	0	0%
CIDA	859	435	859	435	100%
Well state	416	294	416	294	68%
Well depth	383	273	383	273	63%
Weath. layer thickness	0	0	0	0	0%
Lithology	0	0	0	0	0%
Groundwater level	153	149	153	149	34%
Yield [(2)]	256	166	256	166	38%
Water quality [(3)]	110	106	110	106	24%
EU	483	397	483	397	100%
Well state	1	1	1	1	0%
Well depth	365	362	365	362	91%
Weath. layer thickness	0	0	0	0	0%
Lithology	0	0	0	0	0%
Groundwater level	105	104	105	104	26%
Yield [(2)]	117	116	117	116	29%
Water quality [(3)]	2	2	2	2	1%
AFD	231	82	231	82	100%
Well state	11	11	11	11	13%
Well depth	82	82	82	82	100%
Weath. layer thickness	0	0	0	0	0%
Lithology	0	0	0	0	0%
Groundwater level	81	81	81	81	99%
Yield [(2)]	82	82	82	82	100%
Water quality [(3)]	71	71	71	71	87%
WV	964	898	377	348	100%
Well state	964	898	377	348	100%
Well depth	963	897	377	348	100%
Weath. layer thickness	0	0	0	0	0%
Lithology	964	898	348	348	100%
Groundwater level	320	302	152	141	41%
Yield [(2)]	199	189	56	51	15%
Water quality [(3)]	113	108	46	46	13%
Total	15092	12446	9340	7056	-

Notes:

(1) : lithology in the GLOWA database is limited to one column (no stratigraphic unit descrip
(2) : records for which yield data from airlift or pumping tests were available
(3) : records for which water quality data is available for at least: pH, EC, F, Fe and Mn

Table 5: Summary statistics for unique and reliable records in all databases

	Records	Area covered (1)	Required for HAP
Total records (in all databases)	15092	-	-
Unique records	9340	389	-
Unique & reliable records	7056	381	-
Unique & reliable records with well state	6368	352	-
Unique & reliable records with well depth	3934	352	-
Unique & reliable records with weathered layer	277	86	-
Unique & reliable records with lithology	681	99	-
Unique & reliable records with water level	1491	289	-
Unique & reliable records with yield	2938	297	-
Unique & reliable records with water quality (2)	370	163	-
Unique & reliable records with all required fields (3)	71	26	507 (4)

Notes:

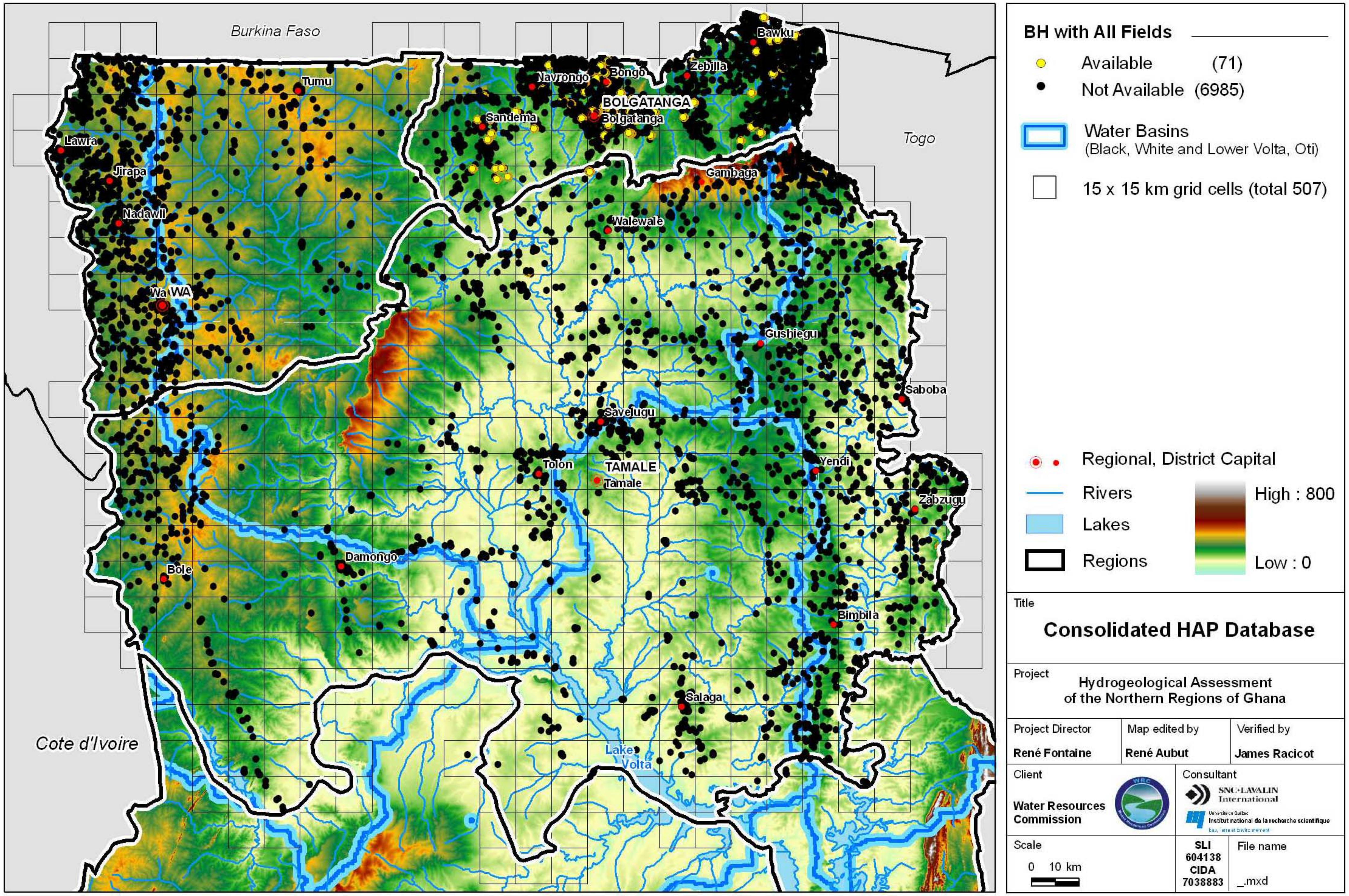
(1) : the area covered by each category is expressed by the number of 15x15km cells with one or more record in it

(2) : records for which water quality data is available for at least the following: pH, EC, F, Fe and Mn (N.B.: zeros exlcuded)

(3) : the field containing the weathered layer thickness was excluded from this calculation since it can be derived from lithological information

(4) : this target represents the total number of 15x15km cells in the northern regions (97721 km^2); for each cells, we need at least one borehole with reliable information for all required fields (N.B.: cell size is arbitrary but comparable to size used in similar regional studies; uniform data distribution resulting from this cell grid is necessary to carry out many of the required analysis)

Figure 4: Spatial distribution of unique & reliable records with all required data fields



Prepared by:	______________________ Marc-André Carrier GIS/Mapping advisor	Date: August 2006
Reviewed by:	______________________ James Racicot Ghana-based Canadian advisor	Date: August 2006
	______________________ René Lefebvre Technical Capacity Building Team Leader	Date: August 2006

Appendix B

Data field correspondence between HAP consolidated data and WRC Water Resources database

WRC Water Resources database - Dataset list		
Dataset ID	**Dataset Description**	**Dataset Type**
1	Well	FeatureClass
5	MonitoringPoint	FeatureClass
7	Aquifer	FeatureClass
12	HydroGeologicUnit	Table
37	SeriesCatalog	Table
39	VariableDefinition	Table
40	BoreholeLithology	Table
41	BoreholeConstruction	Table
42	BoreholeSample	Table
43	BoreholeIndex	Table
44	BoreLineLithology	FeatureClass
45	BoreLineConstruction	FeatureClass
46	BoreLineSample	FeatureClass
48	SeriesGroundwaterLevel	Table
49	SeriesGroundwaterQuality	Table
50	SeriesPumpingTest	Table
51	SeriesStepDrawdownPumpingTest	Table
52	SeriesRiverFlow	Table
53	SeriesRainfall	Table
54	SeriesRainwaterQuality	Table

Data field correspondence between HAP consolidated data and WRC Water Resources database (cont'd)

FeatureClassName	**Aquifer**
DatasetType	FeatureClass
Description	Describes aquifer boundaries and zones within aquifers such as confined and unconfined areas
FeatureDataset	Framework
DataTheme	ArcHydroFramework
ShapeType	Polygon
FeatureType	Simple
AliasName	Aquifer
HasM	false
HasZ	false
SubtypeFieldName	null
DefaultSubtype	null
DSID	7

WRC Water Resources database					HAP consolidated data			
Field	**Type**	**Description**	**Unique ID**	**Domain**	**Field**	**Type**	**Basic field**	**Comments**
HydroID	Integer 4	Unique feature identifier in the Geodatabase	yes	no	none		no	Internal ID: will be linked to AquiferID in other tables
HydroCode	String 30	Permanent public identifier of the feature	yes	no	none		no	ID for external link
Name	String 30	Name of the Aquifer	no	yes	stratigraphy	String 50	no	Main geological context which acts as regional hydrogeological context
GeoUnit	String 50	Geological unit (based on lithostratigraphic/lithodemic classification)	no	yes	stratigraphy	String 50	no	Geological unit (based on lithostratigraphic/lithodemic classification)
HGUID	Integer 4	Identifier of the hydrogeologic unit	no	no	none		no	Internal ID from the HydroGeologicUnit table
FType	String 30	Classification of the Feature Type for mapping and analytical purposes	no	yes	none		no	Values defined by AquiferFType domain

FeatureClassName	**MonitoringPoint**
DatasetType	FeatureClass
Description	A location where water properties are measured, such as a stream gage or a monitoring well
FeatureDataset	Framework
DataTheme	ArcHydroFramework
ShapeType	Point
FeatureType	Simple
AliasName	MonitoringPoint
HasM	false
HasZ	false
SubtypeFieldName	null
DefaultSubtype	null
DSID	5

WRC Water Resources database					HAP consolidated data			
Field	**Type**	**Description**	**Unique ID**	**Domain**	**Field**	**Type**	**Basic field**	**Comments**
HydroID	Integer 4	Unique feature identifier in the Geodatabase	yes	no	none		no	
HydroCode	String 30	Permanent public identifier of the feature	no	no	sheet_id	String 30	yes	ID for external link
FType	String 30	Classification of the Feature Type for mapping and analytical purposes	no	yes	none		yes	Type of monitoring point
Name	String 100	Name of Monitoring Point	no	no	wellid_1	String 20	no	Main well ID given under the project that funded the wells
WellID	Integer 4	Identifier that points to a well feature	no	no	none		no	HydroID found in Well Feature Class
JunctionID	Integer 4	HydroID of the related HydroJunction	no	no	none		no	HydroID of the related HydroJunction

Data field correspondence between HAP consolidated data and WRC Water Resources database (cont'd)

FeatureClassName	**Well**
DatasetType	FeatureClass
Description	A point that represents the location of a well and associated attributes
FeatureDataset	Framework
DataTheme	ArcHydroFramework
ShapeType	Point
FeatureType	Simple
AliasName	Well
HasM	false
HasZ	false
SubtypeFieldName	null
DefaultSubtype	null
DSID	1

WRC Water Resources database					HAP consolidated data			
Field	**Type**	**Description**	**Unique ID**	**Domain**	**Field**	**Type**	**Basic field**	**Comments**
HydroID	Integer 4	Unique feature identifier in the Geodatabase	yes	no	none		no	Internal ID: will be linked to WellID in other tables
HydroCode	String 30	Permanent public identifier of the feature	yes	no	sheet_id	String 30	yes	ID for external link
Longitude	Double 20	Longitude (in decimal degrees)	no	no	longitude	Double	yes	Longitude (in decimal degrees)
Latitude	Double 20	Latitude (in decimal degrees)	no	no	latitude	Double	yes	Latitude (in decimal degrees)
LandElev	Double 8	Land surface or reference elevation (in meters above mean sea level)	no	yes	z_dem	Double	yes	Extracted from SRTM data, with ~ 10m vertical accuracy
WellDate	Date 8	Well completion date	no	no	date_bh	Date 8	yes	Well completion date
WellDepth	Double 8	Depth of feature (in meters below ground level)	no	no	well_depth	Double	yes	Well depth (in meters)
WellState	String 30	Well state (e.g. dry, active …)	no	yes	well_state	String 25	yes	Well state (e.g. dry, active …)
WellType	String 30	Well type (e.g. borehole, hand-dug well, …)	no	yes	well_type	String 20	yes	Values defined by WellType domain
WellUse	String 30	Well use (e.g. monitoring, irrigation, …)	no	yes	well_use	String 20	yes	Values defined by WellUse domain
AquiferID	Integer 4	HydroID of related Aquifer	no	no	none		no	HydroID found in Aquifer Feature Class
AqCode	String 30	Text description for the aquifer related to the well	no	no	none		no	HydroCode of Aquifer Feature Class
HGUGeneral	String 30	Generalized hydrogeologic unit intercepted by well	no	yes	none		no	Generalized hydrogeologic unit intercepted by well
GeoUnit	String 50	Geological unit (based on lithostratigraphic/lithodemic classification)	no	yes	stratigraphy	String 50	no	Geological unit (based on lithostratigraphic/lithodemic classification)
CommID	String 30	Community identifier	no	no	villageid	String 20	no	Community identifier
Comm	String 50	Community name	no	no	village	String 35	yes	Community name
DistrictID	String 10	District identifier	no	no	districtid	String 20	no	District identifier
District	String 30	District name	no	no	district	String 25	yes	District name
RegionID	String 10	Region identifier	no	no	regionid	String 20	no	Region identifier
Region	String 30	Region name	no	yes	region	String 20	yes	Region name
Basin	String 30	Hydrological basin name	no	yes	basin	String 25	no	Hydrological basin name
GPSReading	String 30	Location of GPS reading (e.g. well or community)	no	yes	gps	String 20	no	Location of GPS reading (e.g. well or community)
SitingComp	String 30	Name of company in charge of borehole siting	no	no	sit_consul	String 30	no	Name of company in charge of borehole siting
SitingDate	Date 8	Date of borehole siting	no	no	date_sit	Date 8	no	Date of borehole siting
BhComp	String 30	Name of borehole drilling company	no	no	bh_company	String 20	no	Name of borehole drilling company
BhMethod	String 30	Borehole drilling method	no	yes	bh_method	String 25	yes	Borehole drilling method
AirliftYld	Double 20	Airlift yield (in L/min)	no	no	airlift_yd	Double	yes	Airlift yield (in L/min)
PumpType	String 30	Type of pump used for water supply (e.g. footpump, handpump, other …)	no	yes	pump_type	String 25	yes	Pump type (e.g. footpump, handpump, other …)
PumpDepth	Double 20	Pump installation depth (in meters)	no	no	pump_depth	Double	yes	Pump installation depth (in meters)
PumpDate	Date 8	Pump installation date	no	no	date_pump	Date 8	yes	Pump installation date
CommentGen	String 120	Comments related to data input, quality, modification, update or other	no	no	comments	String 30	no	General comments
CommentLoc	String 120	Comments related to coordinates reliability or modifications	no	no	location_	String 110	no	Comments from reliability analysis of coordinates
EntryDate	Date 8	Date of last data entry, modification or update	no	no	date_entry	Date 8	no	Date of last data entry, modification or update
HasXY	Integer 4	Classification of records based on coordinates reliability	no	yes	dtset_xy_	Integer 4	no	Classification of records based on coordinates reliability
HasDpt	Integer 4	Classification of records based on well depth information availability	no	yes	dtset_dpt_	Integer 4	no	Classification of records based on well depth information availability
HasLth	Integer 4	Classification of records based on lithology information availability	no	yes	dtset_lth_	Integer 4	no	Classification of records based on lithology information availability
HasSwl	Integer 4	Classification of records based on static gw level information availability	no	yes	dtset_swl_	Integer 4	no	Classification of records based on static gw level information availability
HasYld	Integer 4	Classification of records based on yield information availability	no	yes	dtset_yld_	Integer 4	no	Classification of records based on yield information availability
HasGwl	Integer 4	Classification of records based on gw quality information availability	no	yes	dtset_wq_	Integer 4	no	Classification of records based on gw quality information availability
HasDupl	Integer 4	Classification of records based on the existence of duplicates records	no	yes	dupl_	Integer 4	no	Classification of records based on the existence of duplicates records
DuplInfo	String 120	Supplement information on duplicate records	no	no	dupl__	String 85	no	Supplement information on duplicate records
HAPORDER	Integer 4	Unique numeric value identifying original HAP record order	no	no	order_	Integer 4	no	Unique numeric value identifying original HAP record order

Data field correspondence between HAP consolidated data and WRC Water Resources database (cont'd)

FeatureClassName **BoreLineLithology**
DatasetType FeatureClass
Description Three-dimensional line that represents lithologic units along a borehole
FeatureDataset Borehole
DataTheme WellsAndBoreholes
ShapeType Polyline
FeatureType Simple
AliasName BoreLineLithology
HasM false
HasZ true
SubtypeFieldName null
DefaultSubtype null
DSID 44

WRC Water Resources database					HAP consolidated data			
Field	**Type**	**Description**	**Unique ID**	**Domain**	**Field**	**Type**	**Basic field**	**Comments**
HydroID	Integer 4	Unique feature identifier in this FC and in all the database	yes	no	none		no	Internal ID: will be calculated in ArcHydro Data Model
HydroCode	String 30	Permanent public identifier of the feature	yes	no	none		no	ID for external link
WellID	Integer 4	Identifier that points to a well feature	no	no	none		no	HydroID found in Well Feature Class
LogLithID	Integer 4	Unique record identifier of lithologic unit in the table	yes	no	none		no	LogID found in BoreholeLithology Table
TopElev *	Double	Top Elevation of the feature (in meters)	no	no	none		no	Data comes from the BoreholeLithology Table
BotElev *	Double	Bottom Elevation of the feature (in meters)	no	no	none		no	Data comes from the BoreholeLithology Table
HGUID	Integer 4	Identifier of the hydrogeologic unit	no	no	none		no	Internal ID from the HydroGeoologicUnit table
HGUCode	String 30	Public identifier of the hydrogeologic unit	no	yes	none		no	ID for external link - abbreviation/short name for the HGUName field
LogType	Integer 4	LogType (subtype, e.g. lithology)	no	yes	none		no	Required to specify the log type (i.e. lithology)
LogCode	String 30	Publich identifier of log (domain related to LogType subtype)	no	yes	none		no	Type of lithologic units

FeatureClassName **BoreLineConstruction**
DatasetType FeatureClass
Description Three-dimensional line that represents construction details along a borehole
FeatureDataset Borehole
DataTheme WellsAndBoreholes
ShapeType Polyline
FeatureType Simple
AliasName BoreLineConstruction
HasM false
HasZ true
SubtypeFieldName null
DefaultSubtype null
DSID 45

WRC Water Resources database					HAP consolidated data			
Field	**Type**	**Description**	**Unique ID**	**Domain**	**Field**	**Type**	**Basic field**	**Comments**
HydroID	Integer 4	Unique feature identifier in this FC and in all the database	yes	no	none		no	Internal ID: will be calculated in ArcHydro Data Model
HydroCode	String 30	Permanent public identifier of the feature	yes	no	none		no	ID for external link
WellID	Integer 4	Identifier that points to a well feature	no	no	none		no	HydroID found in Well Feature Class
LogConstID	Integer 4	Unique record identifier of borehole construction feature in the table	yes	no	none		no	LogID found in BoreholeConstruction Table
TopElev *	Double	Top Elevation of the feature (in meters)	no	no	none		no	Data comes from the BoreholeConstruction Table
BotElev *	Double	Bottom Elevation of the feature (in meters)	no	no	none		no	Data comes from the BoreholeConstruction Table
HGUCodes	String 30	Public identifier of the hydrogeologic units intercepted by screened interval	no	no	none		no	Data extracted for bh with lithology data
HGUGeneral	String 30	Generalized hydrogeologic units intercepted by screened interval	no	yes	none		no	Data extracted for bh with lithology data
LogType	Integer 4	LogType (subtype, e.g. bh completion, bh annulus, ...)	no	yes	none		no	Required to specify the log type (i.e. bh completion, bh annulus, ...)
LogCode	String 30	Publich identifier of log (domain related to LogType subtype)	no	yes	none		no	Type of bh completion feature
LogDescr	String 120	Description of log (e.g. description of bh construction feature)	no	no	none		no	Description of log (e.g. description of bh construction feature)

Data field correspondence between HAP consolidated data and WRC Water Resources database (cont'd)

FeatureClassName **BoreLineSample**
DatasetType FeatureClass
Description Three-dimensional line that represents a sample along a borehole
FeatureDataset Borehole
DataTheme WellsAndBoreholes
ShapeType Polyline
FeatureType Simple
AliasName BoreLineSample
HasM false
HasZ true
SubtypeFieldName null
DefaultSubtype null
DSID 46

WRC Water Resources database					HAP consolidated data			
Field	**Type**	**Description**	**Unique ID**	**Domain**	**Field**	**Type**	**Basic field**	**Comments**
HydroID	Integer 4	Unique feature identifier in this FC and in all the database	yes	no	none		no	Internal ID: will be calculated in ArcHydro Data Model
HydroCode	String 30	Permanent public identifier of the feature	yes	no	none		no	ID for external link
WellID	Integer 4	Identifier that points to a well feature	no	no	none		no	HydroID found in Well Feature Class
LogSampID	Integer 4	Unique record identifier of sample in the table	yes	no	none		no	LogID found in BoreholeSample Table
TopElev *	Double	Top Elevation of the feature (in meters)	no	no	none		no	Data comes from the BoreholeSample Table
BotElev *	Double	Bottom Elevation of the feature (in meters)	no	no	none		no	Data comes from the BoreholeSample Table
LogType	Integer 4	LogType (subtype, e.g. interval or point samples)	no	yes	none		no	Required to specify the log type (i.e. point/interval samples)
LogCode	String 30	Publich identifier of log (domain related to LogType subtype)	no	yes	none		no	Type of sample

TableName **BoreholeIndex**
DatasetType Table
Description Table to store the different borehole IDs collected under HAP
DataTheme WellsAndBoreholes
AliasName BoreholeIndex
SubtypeFieldName null
DefaultSubtype null
DSID 43

WRC Water Resources database					HAP consolidated data			
Field	**Type**	**Description**	**Unique ID**	**Domain**	**Field**	**Type**	**Basic field**	**Comments**
WellID	Integer 4	Identifier that points to a well feature	yes	no	none		no	HydroID found in Well Feature Class
HAPID	Integer 9	Identifier given to HAP consolidated data	yes	no	order_	Integer 9	no	Data field ORDER_ from HAP consolidated data
WellID_1	String 30	One of the well ID from the many different sources	no	no	wellid_1	String 20	no	One of the well ID from the many different sources
WellID_2	String 30	One of the well ID from the many different sources	no	no	wellid_2	String 20	no	One of the well ID from the many different sources
WellID_3	String 30	One of the well ID from the many different sources	no	no	wellid_3	String 20	no	One of the well ID from the many different sources
WellRefNo	String 30	One of the well ID from the many different sources	no	no	wellrefno	String 25	no	One of the well ID from the many different sources
PrjctName	String 30	Project name under which the well was installed	no	no	prjctname	String 30	no	Project name under which the well was installed
PrjctFund	String 30	Organisation which funded the project	no	no	prjctfund	String 30	no	Organisation which funded the project
LicenseNo	String 30	Water drilling license number of drilling contractor	no	no	none		no	Licence number granted by WRC
DSource1	String 50	Data source 1	no	no	source_1	String 35	no	Data source 1
DSource2	String 50	Data source 2	no	no	source_2	String 35	no	Data source 2
DSource3	String 50	Data source 3	no	no	source_3	String 35	no	Data source 3
GridId	String 10	ID of 15x15km grid square over northern region	no	no	gridId	String 10	no	ID of 15x15km grid cells covering northern regions

Data field correspondence between HAP consolidated data and WRC Water Resources database (cont'd)

TableName **BoreholeLithology**
DatasetType Table
Description Table of vertical measurements relative to borehole lithology
DataTheme WellsAndBoreholes
AliasName BoreholeLithology
SubtypeFieldName null
DefaultSubtype null
DSID 40

WRC Water Resources database					HAP consolidated data			
Field	**Type**	**Description**	**Unique ID**	**Domain**	**Field**	**Type**	**Basic field**	**Comments**
LogLithID	Integer 4	Unique record identifier of lithologic unit in the table	yes	no	none		no	Sequential number
WellID	Integer 4	Identifier that points to a well feature	no	no	none		no	HydroID found in Well Feature Class
RefElev	Double	RefElev (same field as LandElev from FC Well) (in meters)	no	yes	z_dem	Double	yes	Extracted from SRTM data, with ~ 10m vertical accuracy
FromDepth	Double	FromDepth (in meters)	no	no	lith(X-1)bot	Double	yes	lith(X-1)bot represents the top value of unit lithX
ToDepth	Double	ToDepth (in meters)	no	no	lithXbot	Double	yes	lithXbot is the bottom depth of lithX
TopElev	Double	Top Elevation of the feature (Elevation for point features) (in meters)	no	no	none		no	Calculated from the FromDepth and RefElev
BotElev	Double	Bottom Elevation of the feature (in meters)	no	no	none		no	Calculated from the ToDepth and RefElev
ElevUnits	String 30	ElevUnits	no	yes	none		no	All elevation units from HAP consolidated data are in meters
HGUID	Integer 4	Identifier of the hydrogeologic unit	no	no	none		no	Internal ID from the HydroGeologicUnit table
HGUCode	String 30	Public identifier of the hydrogeologic unit	no	yes	none		no	ID for external link - abbreviation/short name for the HGUName field
LogType	Integer 4	LogType (subtype, e.g. lithology)	no	yes	none		no	Required to specify the log type (i.e. lithology)
LogCode	String 30	Public identifier of log (domain related to LogType subtype)	no	yes	lithXunit	String 20	no	lithXunit is the lithology code for unit X
LogDescr	String 120	Description of log (e.g. detailed description of lithologic unit)	no	no	lithXdesc	String 100	no	lithXdesc is the lithology description for unit X

TableName **BoreholeConstruction**
DatasetType Table
Description Table of vertical measurements relative to borehole construction
DataTheme WellsAndBoreholes
AliasName BoreholeConstruction
SubtypeFieldName null
DefaultSubtype null
DSID 41

WRC Water Resources database					HAP consolidated data			
Field	**Type**	**Description**	**Unique ID**	**Domain**	**Field**	**Type**	**Basic field**	**Comments**
LogConstID	Integer 4	Unique record identifier of borehole construction feature in the table	yes	no	none		no	Sequential number
WellID	Integer 4	Identifier that points to a well feature	no	no	none		no	HydroID found in Well Feature Class
RefElev	Double	RefElev (same field as LandElev from FC Well) (in meters)	no	yes	z_dem	Double	yes	Extracted from SRTM data, with ~ 10m vertical accuracy
FromDepth	Double	FromDepth (in meters)	no	no	(multiple)	Double	yes	Multiple (e.g. screenXtop is the top depth of screenX)
ToDepth	Double	ToDepth (in meters)	no	no	(multiple)	Double	yes	Multiple (e.g. screenXbot is the bottom depth of screenX)
TopElev	Double	Top Elevation of the feature (Elevation for point features) (in meters)	no	no	none		no	Calculated from the FromDepth and RefElev
BotElev	Double	Bottom Elevation of the feature (in meters)	no	no	none		no	Calculated from the ToDepth and RefElev
ElevUnits	String 30	ElevUnits	no	yes	none		no	All elevation units from HAP consolidated data are in meters
HGUCodes	String 30	Public identifier of the hydrogeologic units intercepted by screened interval	no	no	none		no	Data extracted for bh with lithology data
HGUGeneral	String 30	Generalized hydrogeologic units intercepted by screened interval	no	yes	none		no	Data extracted for bh with lithology data
FGeometry	Integer 4	Feature geometry (e.g. punctual, interval)	no	yes	none		no	Required to specify the data or feature type
LogType	Integer 4	LogType (subtype, e.g. bh completion, bh annulus, ...)	no	yes	none		no	Required to specify the log type (i.e. bh completion, bh annulus, ...)
LogCode	String 30	Public identifier of log (domain related to LogType subtype)	no	yes	(multiple)	String	no	Multiple (e.g. annulusXde is the code for annulus interval X)
LogDescr	String 120	Description of log (e.g. description of bh construction feature)	no	no	none		no	Description of log (e.g. description of bh construction feature)

Data field correspondence between HAP consolidated data and WRC Water Resources database (cont'd)

TableName **BoreholeSample**
DatasetType Table
Description Table of vertical measurements relative to borehole samples
DataTheme WellsAndBoreholes
AliasName BoreholeSample
SubtypeFieldName null
DefaultSubtype null
DSID 42

WRC Water Resources database					HAP consolidated data			
Field	**Type**	**Description**	**Unique ID**	**Domain**	**Field**	**Type**	**Basic field**	**Comments**
LogSampID	Integer 4	Unique record identifier of sample in the table	yes	no	none		no	Sequential number
WellID	Integer 4	Identifier that points to a well feature	no	no	none		no	HydroID found in Well Feature Class
RefElev	Double	RefElev (same field as LandElev from FC Well) (in meters)	no	yes	z_dem	Double	yes	Extracted from SRTM data, with ~ 10m vertical accuracy
FromDepth	Double	FromDepth (in meters)	no	no	none		yes	To be calculated based on borehole construction details
ToDepth	Double	ToDepth (in meters)	no	no	none		yes	To be calculated based on borehole construction details
TopElev	Double	Top Elevation of the feature (Elevation for point features) (in meters)	no	no	none		no	Calculated from the FromDepth and RefElev
BotElev	Double	Bottom Elevation of the feature (in meters)	no	no	none		no	Calculated from the ToDepth and RefElev
ElevUnits	String 30	ElevUnits	no	yes	none		no	All elevation units from HAP consolidated data are in meters
FGeometry	Integer 4	Feature geometry (e.g. punctual, interval)	no	yes	none		no	Required to specify the data or feature type
LogType	Integer 4	LogType (subtype, e.g. interval or point samples)	no	yes	none		no	Required to specify the log type (i.e. point/interval samples)
LogCode	String 30	Public identifier of log (domain related to LogType subtype)	no	yes	none		no	Type of sample (e.g. groundwater samples)
LogDescr	String 120	Description of log (e.g. detailed description of sample)	no	no	none		no	Description of log (e.g. detailed description of sample)

TableName **HydroGeologicUnit**
DatasetType Table
Description Table that represents properties of Hydrogeologic units
DataTheme Hydrostratigraphy
AliasName HydroGeologicUnit
SubtypeFieldName null
DefaultSubtype null
DSID 12

WRC Water Resources database					HAP consolidated data			
Field	**Type**	**Description**	**Unique ID**	**Domain**	**Field**	**Type**	**Basic field**	**Comments**
HGUID	Integer 4	Identifier of the hydrogeologic unit	yes	no	none		no	internal ID: will be linked to HGUID in other tables
HGUCode	String 30	Public identifier of the hydrogeologic unit	yes	yes	none		no	ID for external link - abbreviation/short name for the HGUName field
HGUName	String 30	Name of the hydrogeologic unit	no	no	none		no	ID for external link - name for the unit based on other data sources
AquiferID	Integer 4	Identifier for Aquifer	no	no	none		no	HydroID found in Aquifer Feature Class
AqCode	String 30	Text description of the aquifer	no	no	none		no	Same as HydroCode in the Aquifer Feature Class
Description	String 100	Description of hydrogeologic unit	no	no	none		no	Description of the HGU based on lithologic information

Data field correspondence between HAP consolidated data and WRC Water Resources database (cont'd)

TableName	**VariableDefinition**
DatasetType	Table
Description	Definition of Variables for Temporal datasets
DataTheme	Temporal
AliasName	Variable Definition
SubtypeFieldName	null
DefaultSubtype	null
DSID	39

WRC Water Resources database					HAP consolidated data			
Field	**Type**	**Description**	**Unique ID**	**Domain**	**Field**	**Type**	**Basic field**	**Comments**
VarID	Integer 4	Unique Variable ID	yes	no	none		no	Internal ID: will be linked to VarID in TimeSeries tables
VarKey	String 50	Key or field name of the variable in the time series described	yes	no	none		yes	ID for multiple attribute time series only
VarName	String 255	Variable Name	no	no	none		no	
VarDesc	String 255	Variable Description	no	no	none		no	
VarUnits	String 50	Variable Units	no	yes	none		yes	Units based on data collected under HAP
SmplMedium	String 50	Sample Medium	no	yes	none		no	
VarCode	String 50	Variable Code according to a defined naming convention	no	no	none		no	
Vocabulary	String 50	Name of vocabulary used to link Variable Names and Codes	no	no	none		no	
Standard	String 50	Quality standard for variable		no	none		no	
TimeUnits	String 50	Time Units (i.e. hours, days, weeks, months, years, …)	no	yes	none		no	
TimeStep	Double 8	Time Step	no	no	none		no	
DataType	String 50	Data Type (i.e. interval, instantaneous, average, ...)	no	yes	none		no	
NoDataVal	Double 8	No Data Value	no	no	none		no	A value of -9999 could be used as NoData
IsRegular	Integer 4	Is the time series regular	no	yes	none		no	
Note	String 255	Note on variable	no	no	none		no	Negative numbers could be used for gw quality values under detection limit

TableName	**SeriesCatalog**
DatasetType	Table
Description	Catalog of time series values
DataTheme	Temporal
AliasName	Series Catalog
SubtypeFieldName	null
DefaultSubtype	null
DSID	37

WRC Water Resources database					HAP consolidated data			
Field	**Type**	**Description**	**Unique ID**	**Domain**	**Field**	**Type**	**Basic field**	**Comments**
SeriesID	Integer 4	Time Series ID	yes	no	none		no	Internal ID to identify Time Series in ArcHydro Data Model
FeatureID	Integer 4	HydroID of related feature	no	no	none		no	HydroID found in Well Feature Class
FeatClass	String 255	Feature Class name	no	no	none		no	
VarID	Integer 4	Variable Identifier from Variable table	no	no	none		no	
TsTable	String 255	Time Series Table Name	no	no	none		no	
DSType	String 30	Type of series	no	yes	none		no	
StartTime	Date 8	Start Time of Series	no	no	none		no	
EndTime	Date 8	End Time of Series	no	no	none		no	
ValueCount	Integer 4	Count of Time Series records in the time series	no	no	none		no	

Data field correspondence between HAP consolidated data and WRC Water Resources database (cont'd)

TableName **SeriesGroundwaterLevel**
DatasetType Table
Description Table that contains time series values for groundwater level
DataTheme Temporal
AliasName SeriesGroundwaterLevel
SubtypeFieldName null
DefaultSubtype null
DSID 48

WRC Water Resources database					HAP consolidated data			
Field	**Type**	**Description**	**Unique ID**	**Domain**	**Field**	**Type**	**Basic field**	**Comments**
FeatureID	Integer 4	HydroID of related Feature	no	no	none		yes	HydroID found in Well feature class
TsTime	Date 8	Time Stamp for static groundwater level measurement	no	no	date_swl_x	Date 8	yes	Date of static groundwater level measurement X
UTCOffset	Double 8	UTCOffset (in hours)	no	no	none		no	
RefPt	String 30	Reference point for gw level measurement (i.e. ground, casing, other)	no	yes	none		yes	Description of datum (based on info in reports or on common knowledge)
RefPtHgt	Double 8	Height of reference point relative to ground level (+ above and - below) (in meters)	no	no	swldatum_x	Double	yes	Datum of static groundwater level measurement X
Swl	Double 8	Static groundwater level (in meters)	no	no	swl_x	Double	yes	Static groundwater level measurement X
Swe	Double 8	Static groundwater elevation (in meters above mean sea level)	no	no	none		no	
MType	Integer 4	Measurement type (i.e. estimated, pressure transducer, water level meter)	no	yes	none		yes	
Note	String 255	Note on groundwater level measurement	no	no	none		no	Notes could include corrections made to swl for pressure transducer readings

TableName **SeriesPumpingTest**
DatasetType Table
Description Table that contains time series values for pumping test
DataTheme Temporal
AliasName SeriesPumpingTest
SubtypeFieldName null
DefaultSubtype null
DSID 50

WRC Water Resources database					HAP consolidated data			
Field	**Type**	**Description**	**Unique ID**	**Domain**	**Field**	**Type**	**Basic field**	**Comments**
FeatureID	Integer 4	HydroID of related Feature	no	no	none		yes	HydroID found in Well feature class
TsTime	Date 8	Time Stamp for pumping test	no	no	date_pt	Date 8	no	Date of pumping test
UTCOffset	Double 8	UTCOffset (in hours)	no	no	none		no	
PtTime	Integer 4	Pumping test duration (in hours)	no	no	pt_dura	Double	no	Pumping test duration (in hours)
PtLoc	String 30	Location of gw level measurements (i.e. in pumping well, observation well …)	no	yes	none		no	Always in pumping well for HAP consolidated data
RefPt	String 30	Reference point for gw level measurement (i.e. ground, casing, other)	no	yes	none		no	Description of datum (based on info in reports or on common knowledge)
RefPtHgt	Double 8	Height of reference point relative to ground level (+ above and - below) (in meters)	no	no	none		no	
PtSwl	Double 8	Pumping test static groundwater level (in meters)	no	no	pt_swl	Double	no	Pumping test static groundwater level (in meters)
PtDwl	Double 8	Pumping test dynamic groundwater level (in meters)	no	no	pt_dwl	Double	no	Pumping test dynamic groundwater level (in meters)
PtYld	Double 8	Pumping test yield (in L/min)	no	no	pt_yld	Double	no	Pumping test yield (in L/min)
PtReTime	Integer 4	Pumping test recovery duration (in hours)	no	no	recov_dura	Double	no	Pumping test recovery duration (in hours)
PtReSwl	Double 8	Pumping test recovery static groundwater level (in meters)	no	no	recov_wl	Double	no	Pumping test recovery static groundwater level (in meters)
SC	Double 8	Specific capacity (in L/min m)	no	no	spec_cap	Double	no	Specific capacity (in L/min m)
T	Double 8	Transmissivity (in m2/d)	no	no	transmiss	Double	no	Transmissivity (in m2/d)
S	Double 8	Storage (no units)	no	no	storage	Double	no	Storage (no units)
Note	String 255	Note on groundwater pumping test	no	no	none		no	Notes could include info on method used to calculate hydraulic parameters

Data field correspondence between HAP consolidated data and WRC Water Resources database (cont'd)

TableName **SeriesGroundwaterQuality**
DatasetType Table
Description Table that contains time series values for groundwater quality
DataTheme Temporal
AliasName SeriesGroundwaterQuality
SubtypeFieldName null
DefaultSubtype null
DSID 49

WRC Water Resources database					HAP consolidated data			
Field	**Type**	**Description**	**Unique ID**	**Domain**	**Field**	**Type**	**Basic field**	**Comments**
FeatureID	Integer 4	HydroID of related Feature	no	no	none		yes	HydroID found in Well feature class
TsTime	Date 8	Time Stamp for groundwater sample analysis	no	no	date_analy	Date 8	yes	Date of sample analysis
UTCOffset	Double 8	UTCOffset (in hours)	no	no	none		no	
LogSampID	Integer 4	Record identifier of sample in the log table	no	no	none		no	LogSampID found in BoreholeSample table
SampleName	String 30	Sample name or identifier	no	no	sample_id	String 25	yes	Sample name or identifier
SampleDate	Date 8	Date on which the sample was collected	no	no	date_sampl	Date 8	yes	Date on which the sample was collected
Laboratory	String 30	Name of laboratory that carried out analyses	no	no	laboratory	String 20	yes	Name of laboratory that carried out analyses
Colour	Double 8	Colour (in pcu)	no	no	color	String 20	no	Colour (in pcu)
Turbidity	Double 8	Turbidity (in ntu)	no	no	turbidity	Double	no	Turbidity (in ntu)
Temperatur	Double 8	Temperature (in °C)	no	no	temperatur	Double	no	Temperature (in °C)
EC	Double 8	Electrical conductivity (in µS/cm)	no	no	conductivi	Double	yes	Electrical conductivity (in µS/cm)
TH	Double 8	Total hardness (in mg/L)	no	no	hardness	Double	no	Total hardness (in mg/L)
TA	Double 8	Total alkalinity (in mg/L)	no	no	alkalinity	Double	no	Total alkalinity (in mg/L)
TColi	Double 8	Total coliform count (no units)	no	no	tot_colif	Double	no	Total coliform count (no units)
EColi	Double 8	E-coliform count (no units)	no	no	e_colif	Double	no	E-coliform count (no units)
pH	Double 8	pH (no units)	no	no	ph	Double	yes	pH (no units)
As_	Double 8	Arsenic (in mg/L)	no	no	As_	Double	no	Arsenic (in mg/L)
Ca	Double 8	Calcium (in mg/L)	no	no	Ca	Double	no	Calcium (in mg/L)
Cl	Double 8	Chloride (in mg/L)	no	no	Cl	Double	no	Chloride (in mg/L)
CO2	Double 8	Carbon dioxyde (in mg/L)	no	no	CO2	Double	no	Carbon dioxyde (in mg/L)
CO3	Double 8	Carbonate (in mg/L)	no	no	CO3	Double	no	Carbonate (in mg/L)
Cu	Double 8	Copper (in mg/L)	no	no	Cu	Double	no	Copper (in mg/L)
F	Double 8	Fluoride (in mg/L)	no	no	F	Double	yes	Fluoride (in mg/L)
Fe	Double 8	Iron (in mg/L)	no	no	Fe	Double	yes	Iron (in mg/L)
HCO3	Double 8	Bicarbonate (in mg/L)	no	no	HCO3	Double	no	Bicarbonate (in mg/L)
K	Double 8	Potassium (in mg/L)	no	no	K	Double	no	Potassium (in mg/L)
Mg	Double 8	Magnesium (in mg/L)	no	no	Mg	Double	no	Magnesium (in mg/L)
Mn	Double 8	Manganese (in mg/L)	no	no	Mn	Double	yes	Manganese (in mg/L)
Na	Double 8	Sodium (in mg/L)	no	no	Na	Double	no	Sodium (in mg/L)
NH3_N	Double 8	Ammonia (in mg/L)	no	no	NH3_N	Double	no	Ammonia (in mg/L)
NH4_N	Double 8	Ammonium (in mg/L)	no	no	NH4_N	Double	no	Ammonium (in mg/L)
NO2_N	Double 8	Nitrite (in mg/L)	no	no	NO2_N	Double	no	Nitrite (in mg/L)
NO3_N	Double 8	Nitrate (in mg/L)	no	no	NO3_N	Double	no	Nitrate (in mg/L)
Pb	Double 8	Lead (in mg/L)	no	no	Pb	Double	no	Lead (in mg/L)
PO4	Double 8	Phosphate (in mg/L)	no	no	PO4	Double	no	Phosphate (in mg/L)
SiO2	Double 8	Silicate (in mg/L)	no	no	SiO2	Double	no	Silicate (in mg/L)
SO4	Double 8	Sulfate (in mg/L)	no	no	SO4	Double	no	Sulfate (in mg/L)
TDS	Double 8	Total dissolved solids (in mg/L)	no	no	TDS	Double	no	Total dissolved solids (in mg/L)
Zn	Double 8	Zinc (in mg/L)	no	no	Zn	Double	no	Zinc (in mg/L)
Potability	String 50	Qualitative description of groundwater potability	no	no	Potability	String 20	no	Qualitative description of groundwater potability
WaterType	String 30	Groundwater type (or hydrogeochemical facies) derived from geochemical data	no	no	none		no	Derived from analysis of reliable geochemical data only
Note	String 255	Note on groundwater sample	no	no	none		no	Note on groundwater sample

Data field correspondence between HAP consolidated data and WRC Water Resources database (cont'd)

TableName	**SeriesStepDrawdownPumpingTest**
DatasetType	Table
Description	Table that contains time series values for step drawdown pumping test
DataTheme	Temporal
AliasName	SeriesStepDrawdownPumpingTest
SubtypeFieldName	null
DefaultSubtype	null
DSID	51

WRC Water Resources database					HAP consolidated data			
Field	**Type**	**Description**	**Unique ID**	**Domain**	**Field**	**Type**	**Basic field**	**Comments**
FeatureID	Integer 4	HydroID of related Feature	no	no	none		yes	HydroID found in Well feature class
TsTime	Date 8	Time Stamp for step drawdowm pumping test	no	no	date_pt	Date 8	no	Date of step drawdowm pumping test
UTCOffset	Double 8	UTCOffset (in hours)	no	no	none		no	
SdPtTime	Integer 4	Step-drawdown pumping test duration (in hours)	no	no	sd_dura	Double	no	Step-drawdown pumping test duration (in hours)
SdPtLoc	String 30	Location of gw level measurements (i.e. in pumping well, observation well …)	no	yes	none		no	Always in pumping well for HAP consolidated data
RefPt	String 30	Reference point for gw level measurement (i.e. ground, casing, other)	no	yes	none		no	Description of datum (based on info in reports or on common knowledge)
RefPtHgt	Double 8	Height of reference point relative to ground level (+ above and - below) (in meters)	no	no	none		no	
SdPtSwl	Double 8	Step-drawdown pumping static groundwater level (in meters)	no	no	pt_swl	Double	no	Step-drawdown pumping static groundwater level (in meters)
SdPtDwl_1	Double 8	Step-drawdown pumping dynamic groundwater level (step 1) (in meters)	no	no	none		no	
SdPtDwl_2	Double 8	Step-drawdown pumping dynamic groundwater level (step 2) (in meters)	no	no	none		no	
SdPtDwl_3	Double 8	Step-drawdown pumping dynamic groundwater level (step 3) (in meters)	no	no	none		no	
SdPtYld_1	Double 8	Step-drawdown pumping test yield (step 1) (in L/min)	no	no	sd_yield_1	Double	no	Step-drawdown pumping test yield (step 1) (in L/min)
SdPtYld_2	Double 8	Step-drawdown pumping test yield (step 2) (in L/min)	no	no	sd_yield_2	Double	no	Step-drawdown pumping test yield (step 2) (in L/min)
SdPtYld_3	Double 8	Step-drawdown pumping test yield (step 3) (in L/min)	no	no	sd_yield_3	Double	no	Step-drawdown pumping test yield (step 3) (in L/min)
WE	Double 8	Well efficiency (WE = 100 * BQ/Sw) (in %)	no	no	none		no	
Note	String 255	Note on groundwater pumping test	no	no	none		no	Notes could include info on method used to calculate well efficiency

TableName	**SeriesRiverFlow**
DatasetType	Table
Description	Table that contains time series values for river flow
DataTheme	Temporal
AliasName	SeriesRiverFlow
SubtypeFieldName	null
DefaultSubtype	null
DSID	52

WRC Water Resources database					HAP consolidated data			
Field	**Type**	**Description**	**Unique ID**	**Domain**	**Field**	**Type**	**Basic field**	**Comments**
FeatureID	Integer 4	HydroID of related Feature	no	no	none		no	HydroID found in MonitoringPoint feature class
VarID	Integer 4	Variable ID from Variable Definition table	no	no	none		no	VarID of variable in TsValue
ValueID	Integer 4	Value ID related to TsValue	yes	no	none		no	Sequential number to relate riverflow values to other temporal data
TsTime	Date 8	Time Stamp for Time Series Value	no	no	none		no	Date of river flow measurement
UTCOffset	Double 8	UTCOffset (in hours)	no	no	none		no	
TsValue	Double 8	Time Series Value	no	no	none		no	River flow units and description are mentionned in Variable Definition table

TableName	**SeriesRainfall**
DatasetType	Table
Description	Table that contains time series values for rainfall
DataTheme	Temporal
AliasName	SeriesRainfall
SubtypeFieldName	null
DefaultSubtype	null
DSID	53

WRC Water Resources database					HAP consolidated data			
Field	**Type**	**Description**	**Unique ID**	**Domain**	**Field**	**Type**	**Basic field**	**Comments**
FeatureID	Integer 4	HydroID of related Feature	no	no	none		no	HydroID found in MonitoringPoint feature class
VarID	Integer 4	Variable ID from Variable Definition table	no	no	none		no	VarID of variable in TsValue
ValueID	Integer 4	Value ID related to TsValue	yes	no	none		no	Sequential number to relate rainfall event to rainwater quality
TsTime	Date 8	Time Stamp for Time Series Value	no	no	none		no	Date of rainfall measurement
UTCOffset	Double 8	UTCOffset (in hours)	no	no	none		no	
TsValue	Double 8	Time Series Value	no	no	none		no	Rainfall units and description are mentionned in Variable Definition table

Data field correspondence between HAP consolidated data and WRC Water Resources database (cont'd)

TableName **SeriesRainwaterQuality**
DatasetType Table
Description Table that contains time series values for rainfall
DataTheme Temporal
AliasName SeriesRainwaterQuality
SubtypeFieldName null
DefaultSubtype null
DSID 54

WRC Water Resources database					HAP consolidated data			
Field	**Type**	**Description**	**Unique ID**	**Domain**	**Field**	**Type**	**Basic field**	**Comments**
FeatureID	Integer 4	HydroID of related Feature	no	no	none		no	HydroID found in MonitoringPoint feature class
EventID	Integer 4	Value ID from SeriesRainfall table	no	no	none		no	ValueID found in SeriesRainfall table for event sampled
TsTime	Date 8	Time Stamp for rainwater sample analysis	no	no	none		no	Date of sample analysis
UTCOffset	Double 8	UTCOffset (in hours)	no	no	none		no	
SampleName	String 30	Sample name or identifier	no	no	none		no	Sample name or identifier
SampleDate	Date 8	Date on which the sample was collected	no	no	none		no	Date on which the sample was collected
Laboratory	String 30	Name of laboratory that carried out analyses	no	no	none		no	Name of laboratory that carried out analyses
Colour	Double 8	Colour (in pcu)	no	no	none		no	Colour (in pcu)
Turbidity	Double 8	Turbidity (in ntu)	no	no	none		no	Turbidity (in ntu)
Temperatur	Double 8	Temperature (in °C)	no	no	none		no	Temperature (in °C)
EC	Double 8	Electrical conductivity (in µS/cm)	no	no	none		no	Electrical conductivity (in µS/cm)
TH	Double 8	Total hardness (in mg/L)	no	no	none		no	Total hardness (in mg/L)
TA	Double 8	Total alkalinity (in mg/L)	no	no	none		no	Total alkalinity (in mg/L)
TColi	Double 8	Total coliform count (no units)	no	no	none		no	Total coliform count (no units)
EColi	Double 8	E-coliform count (no units)	no	no	none		no	E-coliform count (no units)
pH	Double 8	pH (no units)	no	no	none		no	pH (no units)
As_	Double 8	Arsenic (in mg/L)	no	no	none		no	Arsenic (in mg/L)
Ca	Double 8	Calcium (in mg/L)	no	no	none		no	Calcium (in mg/L)
Cl	Double 8	Chloride (in mg/L)	no	no	none		no	Chloride (in mg/L)
CO2	Double 8	Carbon dioxyde (in mg/L)	no	no	none		no	Carbon dioxyde (in mg/L)
CO3	Double 8	Carbonate (in mg/L)	no	no	none		no	Carbonate (in mg/L)
Cu	Double 8	Copper (in mg/L)	no	no	none		no	Copper (in mg/L)
F	Double 8	Fluoride (in mg/L)	no	no	none		no	Fluoride (in mg/L)
Fe	Double 8	Iron (in mg/L)	no	no	none		no	Iron (in mg/L)
HCO3	Double 8	Bicarbonate (in mg/L)	no	no	none		no	Bicarbonate (in mg/L)
K	Double 8	Potassium (in mg/L)	no	no	none		no	Potassium (in mg/L)
Mg	Double 8	Magnesium (in mg/L)	no	no	none		no	Magnesium (in mg/L)
Mn	Double 8	Manganese (in mg/L)	no	no	none		no	Manganese (in mg/L)
Na	Double 8	Sodium (in mg/L)	no	no	none		no	Sodium (in mg/L)
NH3_N	Double 8	Ammonia (in mg/L)	no	no	none		no	Ammonia (in mg/L)
NH4_N	Double 8	Ammonium (in mg/L)	no	no	none		no	Ammonium (in mg/L)
NO2_N	Double 8	Nitrite (in mg/L)	no	no	none		no	Nitrite (in mg/L)
NO3_N	Double 8	Nitrate (in mg/L)	no	no	none		no	Nitrate (in mg/L)
Pb	Double 8	Lead (in mg/L)	no	no	none		no	Lead (in mg/L)
PO4	Double 8	Phosphate (in mg/L)	no	no	none		no	Phosphate (in mg/L)
SiO2	Double 8	Silicate (in mg/L)	no	no	none		no	Silicate (in mg/L)
SO4	Double 8	Sulfate (in mg/L)	no	no	none		no	Sulfate (in mg/L)
TDS	Double 8	Total dissolved solids (in mg/L)	no	no	none		no	Total dissolved solids (in mg/L)
Zn	Double 8	Zinc (in mg/L)	no	no	none		no	Zinc (in mg/L)
Note	String 255	Note on rainwater sample	no	no	none		no	Note on rainwater sample

Data field correspondence between HAP consolidated data and WRC Water Resources database (cont'd)

TableName	**SeriesGroundwaterIsotope**
DatasetType	Table
Description	Table that contains time series values for groundwater isotopes
DataTheme	Temporal
AliasName	SeriesGroundwaterIsotope
SubtypeFieldName	null
DefaultSubtype	null
DSID	55

WRC Water Resources database					HAP consolidated data			
Field	**Type**	**Description**	**Unique ID**	**Domain**	**Field**	**Type**	**Basic field**	**Comments**
FeatureID	Integer 4	HydroID of related Feature	no	no	none		yes	HydroID found in Well feature class
TsTime	Date 8	Time Stamp for groundwater sample analysis	no	no	date_analy	Date 8	yes	Date of sample analysis
UTCOffset	Double 8	UTCOffset (in hours)	no	no	none		no	
LogSampID	Integer 4	Record identifier of sample in the log table	no	no	none		no	LogSampID found in BoreholeSample table
SampleName	String 30	Sample name or identifier	no	no	sample_id	String 25	yes	Sample name or identifier
SampleDate	Date 8	Date on which the sample was collected	no	no	date_sampl	Date 8	yes	Date on which the sample was collected
Laboratory	String 30	Name of laboratory that carried out analyses	no	no	laboratory	String 20	yes	Name of laboratory that carried out analyses
Oxygen18	Double 8	Oxygen-18 or ^{18}O (in ‰)	no	no	none		no	
Deuterium	Double 8	Deuterium or ^{2}H (in ‰)	no	no	none		no	
Carbon14	Double 8	Carbon-14 or ^{14}C (in dpm/gC)	no	no	none		no	
Carbon13	Double 8	Carbon-13 or ^{13}C (in ‰)	no	no	none		no	
PMC	Double 8	Percent of modern carbon (in %)	no	no	none		no	
Note	String 255	Note on groundwater sample	no	no	none		no	Note on groundwater sample

Appendix C

<u>Step-by-step procedure to assign HydroID</u>

Following the integration of new data in a feature class of the WRC Water Resources Database, the first step consists in assigning a new and unique identifier (ID) to each feature. This identifier, termed *HydroID*, will notably be used to define relations between features and tabular data within the database. The following steps will ensure consistent *HydroID* assignation for all records (N.B.: guidelines provided below assume the user has a certain knowledge of ArcGIS but more detailed help on general topics can be found in the ArcGIS Desktop Help menu):

1) Load relevant layer(s) (e.g. Well feature class, MonitoringPoint feature class or Aquifer feature class) containing features requiring a *HydroID*

2) Start an edit session for the relevant layer(s) (N.B.: <u>*only if using the "Assign UniqueID" tool for the first time*</u>)

3) Select "Assign UniqueID" tool in the ApUtilities menu of the ArcHydro Tools 9 toolbar

4) Select current dataframe (e.g. "Layers") as Map; select the current WRC Water Resources Database as workspace; select *HydroID* as UniqueID to assign; select the relevant layer(s) containing features requiring a HydroID; select "No" for "Overwrite existing UniqueIDs" (unless overwriting is required); select "All features" for "Apply to" (unless IDs must be assigned to selected features only); click OK (N.B.: close and save edit session after operation if the software does not prompt you to do so)

(N.B.: if the "Assign UniqueID" tool is used for the first time, two (2) new tables should be created during this process, i.e. LayerKey and ApUnique tables)

Step-by-step procedure to assign HydroCode

Following the integration of new data in the Well feature class, the first step consists in assigning a new and unique identifier (ID) to each well with reliable coordinates. The new ID, to be integrated in the *HydroCode* data field of the Well feature class of the WRC Water Resources Database, must be assigned to each new record with reliable coordinates using the vector grid based on the Ghana national borehole numbering scheme (see figure below). While different approaches can be used to derive the *HydroCode* data field (e.g. geoprocessing, object-related programming, manual assignation, …), the methodology described here should help minimize errors that can come from manual entry and it should also be be easier to implement than fully automated assignation that would require programming skills. Guidelines provided below assumes the user has a certain knowledge of ArcGIS but more detailed help on general topics can be found in the ArcGIS Desktop Help menu. Prior to ID assignation, the Ghana Topographic 50k Sheet Index (shapefile or feature class) and the Well feature class containing the existing and new records to be assigned *HydroCodes* should be loaded in ArcMap. These layers may not have the same coordinate systems, so it is recommended to convert the Well feature class to a new temporary layer (e.g. WellProjected) with the same coordinate system as the Ghana Topographic 50k Sheet Index (e.g. Geographic Coordinate System WGS1984). This can be done using the "Project" tool in the Projections and Transformations toolset of the Data Management Toolbox (N.B.: this new temporary layer cannot be stored in the database since its coordinate system is different). When the two layers are loaded in ArcMap and have the same coordinate system, the following steps will ensure consistent *HydroCode* assignation for records:

A) Topographic 50k Sheet Index ID extraction

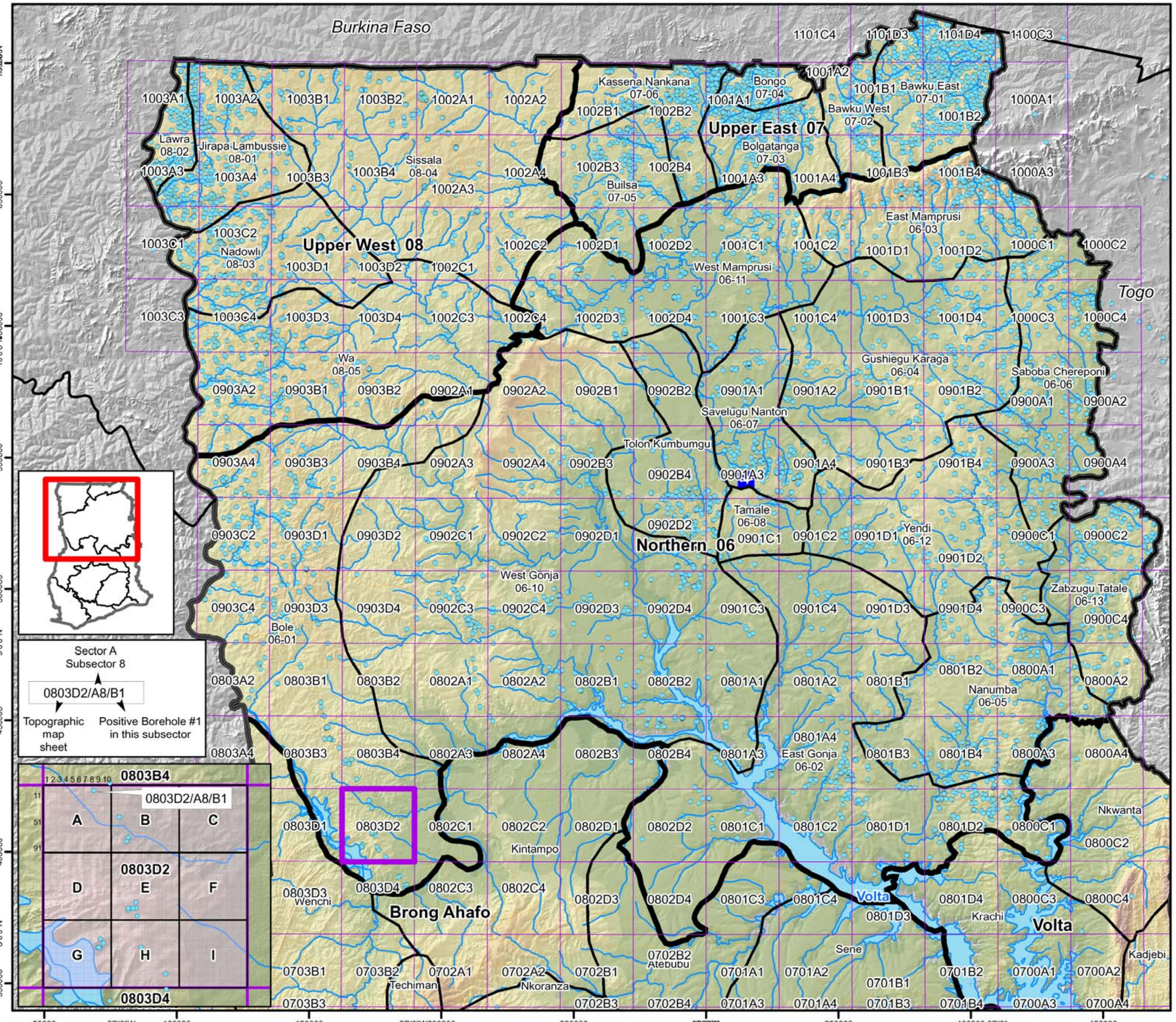
1) Extract values from the *Index50k* data field of the Ghana Topographic 50k Sheet Index for each record or feature by performing a spatial join of the Ghana Topographic 50k Sheet Index to the WellProjected temporary layer. To do so, a) right-click on WellProjected temporary layer, b) select "Joins and Relates" and then "Join …", c) select the "Join data from another layer based on spatial location" option in the pick list, d) choose the Ghana Topographic 50k Sheet Index as layer to join and e) select the "is closest to it" option for attributes assignation, g) indicate name and destination folder of the output file (e.g. WellProjectedJoin temporary layer).

2) Delete the data fields imported through the spatial join process, **except the *Index50k* data field**. To do so, a) right-click on WellProjectedJoin temporary layer, b) select "Open Attribute Table", c) scroll at the end of the attribute table to find the data fields added through the spatial join process, d) right-click on the name of the data field to be deleted and select "Delete Field" (Warning: this operation cannot be undone so be careful which field you delete).

3) Delete values in the *Index50k* data field for the records with unreliable coordinates. To do so, a) right-click on WellProjectedJoin temporary layer, b) select "Open Attribute Table", c) start an edit session (in the "Editor" toolbar), d) click on "Options" at the bottom of the attribute table frame and select "Select by Attributes…", e) build or type the following query to select records with unreliable coordinates : "HasXY" = 0, f) right-click on the name of the *Index50k* data field in the attribute table and select "Field Calculator …", g) enter the following expression to replace selected values: "", h) click OK and stop and save the edit session.

B) Addition of ID suffix for feature type

4) Add the type of feature to the *Index50k* data field by inserting a letter (i.e. B for wet borehole, H for wet hand-dug well and D for dry borehole or dry hand-dug well) at the end of the *Index50k* data field for each feature type. To do so, a) right-click on WellProjectedJoin temporary layer, b) select "Open Attribute Table",c) start an edit session (in the "Editor" toolbar), d) click on "Options" at the bottom of the attribute table frame and select "Select by Attributes…", e) build or type the following query to select records with that represent wet boreholes: "WellType" = 'Borehole' AND "WellState" LIKE 'Wet%', f) right-click on the name of the *Index50k* data field in the attribute table and select "Field Calculator …", g) enter the following expression to replace selected values: [Index50k] & "B", h) build or type the following query to select records with that represent wet hand-dug wells: "WellType" = 'HandDugWell' AND "WellState" LIKE 'Wet%' (), i) right-click on the name of the *Index50k* data field in the attribute table and select "Field Calculator …", j) enter the



Ghana Topographic Index

Scale 1:50,000 sheets

Limits

- Regions
- Districts

Infrastructures

- Airports
- Roads

Hydrography

- Rivers
- Lakes

Elevation (meters above sea level)

800 – 0m

Data Sources: Elevation form Shuttle Radar Topographic Mission (Nasa)
Administratives limits and all base map layers from SWERA
(local source to be determined)

Title

Ghana topographic sheet Index for BH ids

Project

Hydrogeological Assessment of the Northern Regions of Ghana

Project Director	Map edited by	Verified by
René Fontaine		

Client	Consultant
Water Resources Comission	SNC·LAVALIN International; Université du Québec Institut national de la recherche scientifique Eau, Terre et Environnement

Scale		File name
0 10 20 40 60 km	SLI 604138 CIDA 7038883	basemap.mxd

	Date	Description	Drawn	Reviewed
01	July 10, 2009	Preliminar	R. A.	R. F.

\PROJ\604138-Hydrogeo-Ghana\Carto\ArcGIS\1ProjetMXD\Cartes_de_travail

Ghana 1:50,000 sheets numbering is based on latitude / longitude / quadrangle - example: 08 03 D2
Each 1:50,000 map sheet is divided into 9 sectors: A to I, while each sector is divided into 100 cells of 1x1km
B stands for positive borehole, D for dry and W for hand-dug well

following expression to replace selected values: [Index50k] & "H", k) build or type the following query to select records that represent dry boreholes or hand-dug wells: "WellState" LIKE 'Dry%' OR "WellState" = 'UnknownState', l) right-click on the name of the *Index50k* data field in the attribute table and select "Field Calculator …", m) enter the following expression to replace selected values: [Index50k] & "D", n) click OK and stop and save the edit session.

C) Addition of sequential number for records that fall within the same grid cell

5) Sort all records of WellProjectedJoin temporary layer using the resulting *Index50k* data field. To do so, a) right-click on WellProjectedJoin temporary layer, b) select "Open Attribute Table", c) scroll at the end of the attribute table to find the *Index50k* data field and right-click on the name of the data field to select "Sort Ascending".

6) Add a temporary sequential ID in a new data field. To do so, a) click on the XTool Pro toolbar menu and select "Sort Features/Records" tool in Table Operations toolset, b) select the WellProjectedJoin temporary layer as input, indicate name and destination folder of the output file (e.g. WellProjectedJoinSorted temporary layer), select *Index50k* data field for sorting and click OK (N.B.: the new sequential ID data field will be called FID), c) create a new data field by clicking on "Options" at the bottom of the attribute table frame and select "Add Field …" (name: *SortedID*; type: Long Integer), d) right-click on the name of the *SortedID* data field in the attribute table and select "Field Calculator …", g) build or enter the following expression to copy sequential ID values: [FID].

7) Calculate summary statistics on the new ID field *SortedID*. To do so, a) open the "Summary Statistics" tool in the Statistics toolset of the Analysis toolbox, b) use WellProjectedJoinSorted temporary layer as input, indicate name and destination folder of the output table (e.g. SummaryStats temporary table), select *SortedID* data field as statistics field, select "First value" as statistic type, select *Index50k* data field as Case field and click OK.

8) Assign the results from the SummaryStats temporary table to each record in the WellProjectedJoinSorted temporary layer. To do so, a) create a new data field by clicking on "Options" at the bottom of the WellProjectedJoinSorted attribute table frame and select "Add Field …" (name: *Position*; type: Long Integer), b) right-click on WellProjectedJoinSorted temporary layer, c) select "Joins and Relates" and then "Join …", d) select the "Join attributes from a table " option in the pick list, select the *Index50k* data field as "field to base the join on", e) choose the SummaryStats temporary table as layer to join, select the *Index50k* data field as "field to base the join on", f) select the "Keep only matching records" join option and click OK, g) start an edit session (in the "Editor" toolbar), h) right-click on the name of the new *Position* data field in the attribute table and select "Field Calculator …", i) build or enter the following expression to copy sequential ID values: [SummaryStats.FIRST_Sort] and click OK, j) stop and save the edit session.

9) Calculate the *HydroCode* values using the following expression: [Index50k] & ([SortedID]-([Position]-1)). To do so, a) create a new data field by clicking on "Options" at the bottom of the WellProjectedJoinSorted attribute table frame and select "Add Field …" (name: *HydroCode*; type: Text; length: 30), b) right-click on the name of the *HydroCode* data field in the attribute table and select "Field Calculator …", c) build or enter the expression mentioned above to calculate values.

10) Tranfer *HydroCode* values of WellProjectedJoinSorted temporary layer to *HydroCode* data field of Well feature class. To do so, a) right-click on Well feature class, b) select "Joins and Relates" and then "Join …", c) select the "Join attributes from a table " option in the pick list, select the *HydroID* data field as "field to base the join on", d) choose the WellProjectedJoinSorted temporary layer as layer to join, select the *HydroID* data field as "field to base the join on", e) select the "Keep only matching records" join option and click OK, f) start an edit session (in the "Editor" toolbar), g) right-click on the name of the *HydroCode* data field in the attribute table and select "Field Calculator …", h) build or enter the following expression to copy *HydroCode* values: [WellProjectedJoinSorted.HydroCode] and click OK, i) stop and save the edit session.

Appendix E

Content of Appendix E (Electronic documents)

Appendix E-1: WRC Water Resources Database content report

Appendix E-2: WRC Water Resources Database schema report

Appendix E-3: WRC Water Resources Database template

Appendix E-4: Data entry form templates for WRC Water Resources Database

SNC·LAVALIN

www.snclavalin.com

SNC-Lavalin International inc.
455, René-Lévesque Blvd West
Montreal (Quebec)
H2Z 1Z3 Canada
Tel.: (514) 393-1000
Fax: (514) 861-1454