Groundwater for Sustainable Development 2-3 (2016) 7–15

Contents lists available at ScienceDirect

Groundwater for Sustainable Development

journal homepage: www.elsevier.com/locate/gsd

Research paper

Data integration as the key to building a decision support system for groundwater management: Case of Saiss aquifers, Morocco

Siham Laraichi [a,*], Ali Hammani [a], Aziz Bouignane [b,1]

[a] Hassan II Institute of Agronomy and Veterinary Medicine, Rural Engineering Department, PO. Box 6202, Madinat Al Irfane, Rabat, Morocco
[b] River Basin Agency of Moulouya, PO. Box 703, Oujda, Morocco

ARTICLE INFO

Article history:
Received 3 December 2015
Received in revised form
21 March 2016
Accepted 13 April 2016
Available online 20 April 2016

Keywords:
Decision support system
Groundwater
Heterogeneous data
Data integration
Spatiotemporal data warehouse
Saiss aquifers

ABSTRACT

The widespread of Information Technology tools and the increasing of data production in groundwater management allowed the development of different information systems. Even if this field requires various disciplines, organizations and authorities in charge of groundwater have often worked with operational and specific information systems focusing on a single discipline. Driven by the need to integrate all available data from distributed sources to get relevant information at different levels of the decision-making process, many researches try to deal with this issue. This paper aims to show the role of data integration in building spatiotemporal data warehouse for supporting decision making in groundwater governance and management. The paper presents the approach adopted in this work which consider the geographical unit 'water point' as the model heart and the axis of groundwater data analysis. In order to illustrate this approach, Saiss aquifers have been chosen as a study area. Actually, Saiss aquifer system comprises two superimposed and connected aquifers. These aquifers play a significant role in the socio-economic development of Fez–Meknes region through the supply of drinking water and the satisfaction of irrigation water demand. From the 80s, with decreasing precipitation and increasing groundwater use, Saiss aquifers have known strong overexploitation. Spatiotemporal data warehouse of Saiss aquifers constitutes a platform for developing decision support systems, geo-analytical tools, and intelligent mobile applications, able to quick response to multidimensional queries of groundwater managers and users.

1. Introduction

Even though the debate on the importance of information systems for supporting decision making just started in recent years, the evolution and proliferation of Decision Support Systems (DSS) continue to mark different areas. Significant efforts have been made over the last two decades for the development of geospatial DSS to produce geo-decisional information (Bauzer-Medeiros et al., 2006; Jost, 2001).

The effectiveness of the decision making is based on the provision of relevant information and adapted analysis tools. In the field of groundwater resources management, the amount of available digital data sets is increasing, especially for geospatial data, produced in the framework of numerous studies and researches. Organizations and authorities in charge of groundwater, like Basin Authorities, are actually used to work with operational and specific information systems. This solution, especially if data is variable in time and space, currently turns out an obsolete and outdated solution in the decision making efficiency. This is because of their disparity and the lack of interoperability of these information systems that often become underused or containing outdated and insignificant data in the absence of other complementary data.

Driven by the need to integrate all available data from distributed sources, many researches were focused on data integration process. Thus, data integration is the combination of heterogeneous data residing at different sources, and providing the user an integrated and reconciled view of data in order to extract explicit information at different levels of the decision-making

Abbreviations: ABHS, River basin agency of Sebou; CPU, Central processing unit; DSS, Decision support system; DW, Data warehouse; ETL Process, Export, transfer and load Process; GIS, Geographic information system; ICT, Information and communication technologies; MDA, Multi-dimensional analytical; MPIWRD, Master plan of integrated water resources development; OLAP, On-line analytical processing; SOLAP, Spatial on-line analytical processing; STDW, Spatiotemporal data warehouse; WP_ID, Water point identifier.

* Corresponding author.
E-mail addresses: siham.laraichi@gmail.com (S. Laraichi), a.hammani@iav.ac.ma (A. Hammani), azingat@yahoo.fr (A. Bouignane).
[1] Present address: Taourirt province, Taourirt center, Taourirt, Morocco.

http://dx.doi.org/10.1016/j.gsd.2016.04.003
2352-801X/

process (Cali et al., 2004).

Moreover, data integration gives also users the opportunity to make an integrated analysis of prior information, a single geographical referencing of spatial data, a specialization of data through relationships with geo-referenced geometric entities, and mutual enrichment with semantic and geometric properties (Butenuth et al., 2007). Objectively, a real integration of geographic data is much more than a simple overlay data in a geographic information system (GIS). It is rather towards exchange between individual objects in various information systems (Butenuth et al., 2007).

Data warehouse is considered one of the effective solutions for data integration from different sources. Several definitions have been used to describe what a data warehouse is. Inmon (1996) defined the data warehouse, as "a subject-oriented, integrated, time-variant, non-volatile collection of data in support of management's decision making process".

The purpose of data warehouse is to provide from integrated data the basis for management reports, decision making support, data mining and sophisticated on-line analytical processing (OLAP) through quick answering to multi-dimensional analytical (MDA) queries (Nilakanta et al., 2008).

OLAP tools are not robust to analyze spatial and temporal data. GIS tools are helpful in spatial data analysis but still are not good enough to make full use of spatiotemporal datasets. Therefore, the new approach is to couple OLAP and GIS functions. The huge potential of this alliance allows the use of data warehouses techniques simultaneously with OLAP analysis and reporting tools, dashboards and data mining and maps visualization. This alliance is called Spatial OLAP systems, or SOLAP (Maceachren and Kraak, 2001). Several researches have been oriented towards the study of possible combinations of different OLAP and GIS technologies.

The field of Spatial Data Warehouses (SDW) has been emerging since the past decade due to the need to analyze large volumes of spatial data. The research in this field has been especially on conceptual models, materialization of spatial indexes, aggregation operations and SOLAP.

Spatial data warehouses aim at effective and efficient querying of spatial data. Decision-makers in water resources management and groundwater especially, often need to get the global picture, but when they see unexpected trends or variations, they need to drill down to get more details to discover the reason for these variations. For example, varying flow at a source that could be due to several reasons like drought, pumping, etc.

The most critical components in data warehouse design are the process of Extract, Transform and Load data (ETL), and data warehouse modeling (Franconi and Sattler, 1999). Data warehouse design has distinguished characteristics compared to traditional database design. This is, in part, because data warehouse design depends upon already existing database systems from which data are extracted, transformed, and aggregated (Nilakanta et al., 2008). The multidimensional conceptual data model can be physically achieved either by using a relational database approach or making use of a specialized multidimensional database.

In the literature, several researchers understood the importance of data integration into a spatiotemporal data warehouse in different fields such as forest inventory (Van Damme, 2010), agriculture (Vidal et al., 1997), and road traffic (Bauzer-Medeiros et al., 2006) and also in water management (Jost, 2001; Vidal et al., 1997). Driven by the need of groundwater data mastering and management some researchers started investing in building information systems and spatiotemporal data warehouses for groundwater (Refsgaard et al., 2010). Convinced of the importance of a unified set for storing and managing groundwater data, some organizations like the Australian Water Commission, launched a project to implement a National Groundwater Information System (NGIS) in 2009 (El Sawah et al., 2011). Despite these efforts, the approach remains unknown and at least not well mastered by groundwater managers and needs to be further strengthened.

The contribution of this paper consists to demonstrate that data integration is an important and an efficient solution to face data distribution and heterogeneity and to build DSS related to groundwater management. To deal with these issues, this paper presents the complexity of data integration process, and the approach adopted to construct DW model. In order to illustrate this approach, Saiss aquifers have been chosen as a study area. The paper will also explain the operating possibilities of this DW as a platform for a DSS development.

2. Materials and methods

2.1. Data integration in a spatiotemporal data warehouse

Data integration in a spatiotemporal data warehouse is recognized in the scientific community specialized in information and communication technologies (ICT), by its complexity in its different stages. This work would mainly focus on the process of DW design and development adapted to groundwater.

Few studies have been carried out on this topic in the field of groundwater resources management. Thus, the approach taken in this work is the result of a reflection on the most appropriate way to store, organize and explore variable data in time and space in order to give managers an interactive platform of data. This platform should initially allow viewing and understanding the phenomena occurring in the study area, and then to help make more suitable decisions regarding integrated water resources management. This approach is adapted to the type of collected data and deals with different constraints faced during data treatment.

The approach proposed in this work consists of 9 steps as explained in Fig. 1:

1. Identification and knowledge of what information is needed for decision making purposes of groundwater managers: This information might be qualitative and quantitative on groundwater supply and current and future water demand and eventually their evolution in time and space. *Example of query*: *What could be the evolution of the groundwater level in the south of the aquifer since the 80s?*
2. Choosing the appropriate solution: Thinking on possible methods to meet the diagnosed needs and expectations, and the choice of the most suitable solution for the DSS building;
3. Data collection: Collection of data relating to groundwater resources from entities responsible for data and information production in various forms and from various sources. Any data type is welcome initially;
4. Data analysis: The analysis of collected data is a very important step in deciding the adequacy with the need for information and the consistency of the DW, and the degree of its success;
5. Metadata definition: Due to the importance of data on data (metadata), this step is considered as a key step in DW building. Metadata are used to store and inquire about information on data in the DW like its origins, sources, how were calculated … etc. Metadata can not only enable the data warehouses design, but also the user interface of the DW (Mattison, 1996; Wu et al., 2001);
6. DW design: DW design and conception is an important step because of its importance in defining the consistence and the

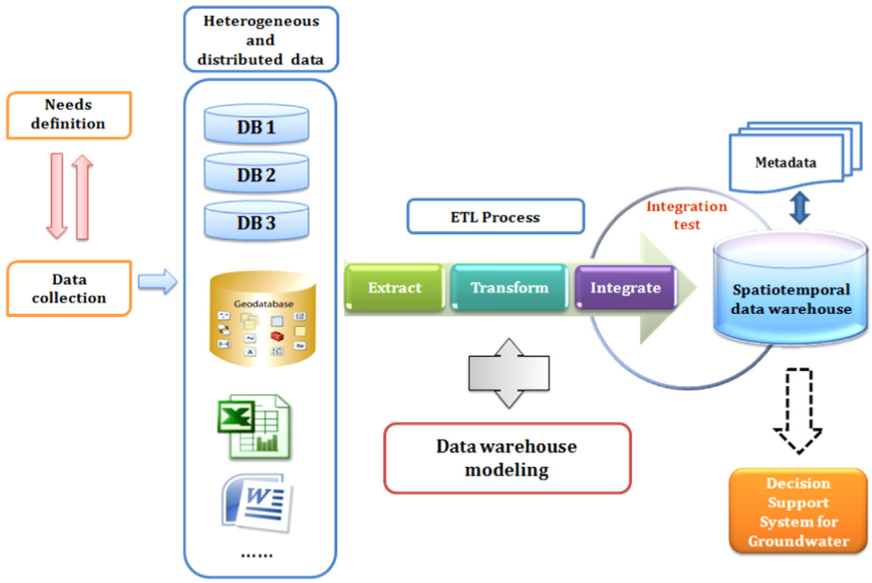

Fig. 1. Approach of a DSS building for groundwater management.

organization of data into the DW. For this work, we have chosen the multidimensional architecture. This type of modeling can offer several advantages such as the ability to build new tables from existing ones;

7. Tools definition: Without fast and reliable tools, the integration process would be tough and difficult. Therefore, tools definition whether it's for ETL process or DW design and development. Several tools allow it, either commercial or open source. Currently, it can be found more than twenty of ETL commercial tools in the market which are designed to handle high performance and scalability requirements. The best known are Informatica Power Center, SAP Data Service, Microsoft® SSIS and IBMs Information Server. The minimum prices start from $45,000 to around $150,000 per CPU for an Enterprise package (http://www.etltools.net/, 2016). Talend, Pentaho and CloverETL are examples of solutions available in open source ETL tools. These products are typically free to use, however, most users find those editions to be too limited for real-life situations. The support, training and consultancy are what the companies need to pay for. Less expensive options exist, but they often are limited in support for heterogeneous environments (Microsoft®, Oracle®) and sometimes charge extra for additional facilities and services such as Metadata management and Data Quality.

Due to their easy accessibility for free and based on the specificity of our data (spatio-temporal data), the choice fell on Open Source tools. Tools used in the present work are defined below based on the task type: (Table 1).

8. ETL process: ETL process with all stages of data extracting from their original sources, transforming it to the target formats, and loading it into the DW for the preparation of the implementation of DSS. One of the most important steps in any data processing task is to verifying data quality including converting any measure data to the same dimension using the same units so that they can later be joined. The final step is to load transformed data into the data warehouse in XML format. Even the use of ETL tools, ETL process was done in this paper mostly manually because of the huge heterogeneity and complexity of data. It was necessary to verify and validate each step of the procedure;
9. Integration test: The final step of this process is to make data integration and compliance tests through simple queries allowed by GIS. These tests provide an opportunity to make final adjustments in DW. It is important to note that although the design of the DW is an earlier step in the ETL process, we opted to perform these two operations in parallel and simultaneously, in order to adjust if there are any changes depending on the relevance of data.

2.2. *Application area*

To enhance the approach presented above, Saiss aquifers have been chosen as an application area to demonstrate how data integration is the key to provide adequate information for groundwater managers. The choice of working in this area is a multicriteria choice. Saïss aquifers are located in Saïss plain, commonly

Table 1
Tools used for data integration in DW.

Task	Tool	Description
Data integration (ETL process)	QGIS	QGIS is a user friendly Open Source Geographic Information System (GIS) licensed under the GNU General Public License. QGIS is an official project of the Open Source Geospatial Foundation (OSGeo). It runs on Linux, Unix, Mac OSX, Windows and Android and supports numerous vector, raster, and database formats and functionalities.
Spatial data warehouse modeling	PostGIS/PostgreSQL	PostGIS is a spatial database extender for PostgreSQL object-relational database. It adds support for geographic objects allowing location queries to be run in SQL. PostgreSQL is one of the best management systems databases, especially in terms of storage and massive data insertions.

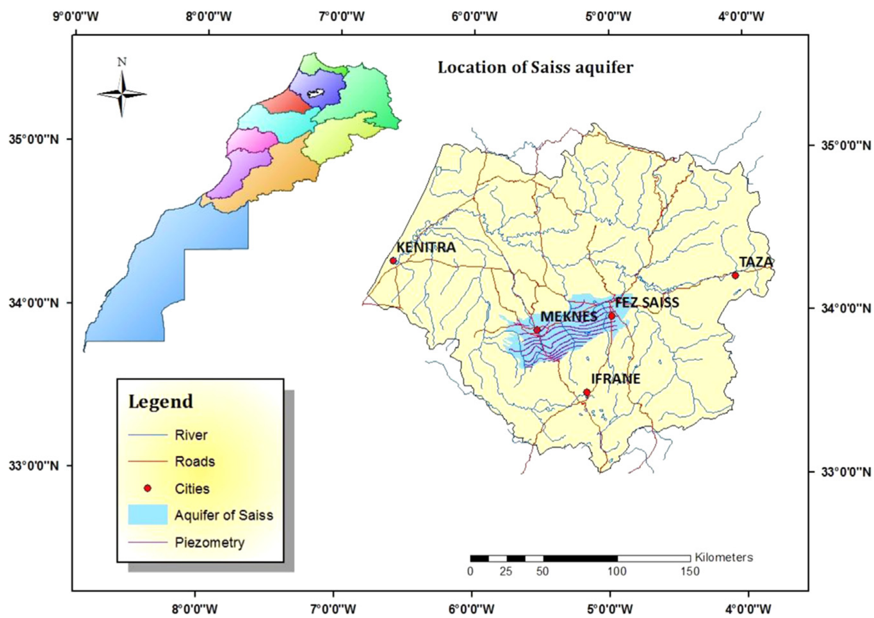

Fig. 2. Location of Saiss aquifers.

known as the plain of Fez–Meknes, in Sebou river basin. Theses aquifers play an important role in socio-economic development of its region through the supply of drinking water of Fez and Meknes cities, and satisfying the irrigation water demand of an about 49,725 ha (Regional Directorate of Agriculture of Fez-Boulemane and Meknes-Tafilalet, 2014).

The Saiss plain covers the center of Sebou basin, about 100 km from west to east and 30 km from north to south, with a total area of approximately 2100 km^2 (Fig. 2). It is bordered to the north by the prerif mountains and south by the Atlas area.

Through the Fig. 3, Belhassan et al. (2010) tried to explain and illustrate the hydrogeological context of Saiss plain. In fact, Saiss aquifer system comprises two superimposed aquifers:

- Unconfined aquifer is circulating in the sands, sandstones, conglomerates and Sahelian Pliocene lacustrine limestones and locally in the travertine. Tortonian deposits of marl are the nature impermeable bedrock of the aquifer. Their thickness can reach 900 m and put in charge the deep aquifer. The average depth of the water table from the ground surface level is 25 m. The dominant flow is generally from south to north (River Basin Agency of Sebou, 2006, 2009);
- Confined aquifer circulating in dolomitic limestones of the Lias and will start charging under the thick series of impermeable marl Miocene. The depth from the floor of the Liassic formations varies between 0 m in the south of the basin and over 1000 m at its northern limit in contact with the prerif formations. The thickness of the Liassic formations is highly variable and ranges from 0 m and 300 m (River Basin Agency of Sebou, 2006, 2009).

In addition to direct abstraction from boreholes catching the deep confined aquifer, each abstraction from the unconfined aquifer affects the confined aquifer. In fact, these two layers are connected in some places, through faults and flexures or indirectly by upward leakage. This is why groundwater table lowering (0.2–0.4 m per year) is less sensitive than the deep layer (1–3 m per year), while the abstraction is more concentrated on shallow groundwater. This issue is causing the withdrawal of artesian field of more than 70% of the extent of the plain (River Basin Agency of Sebou, 2006, 2009).

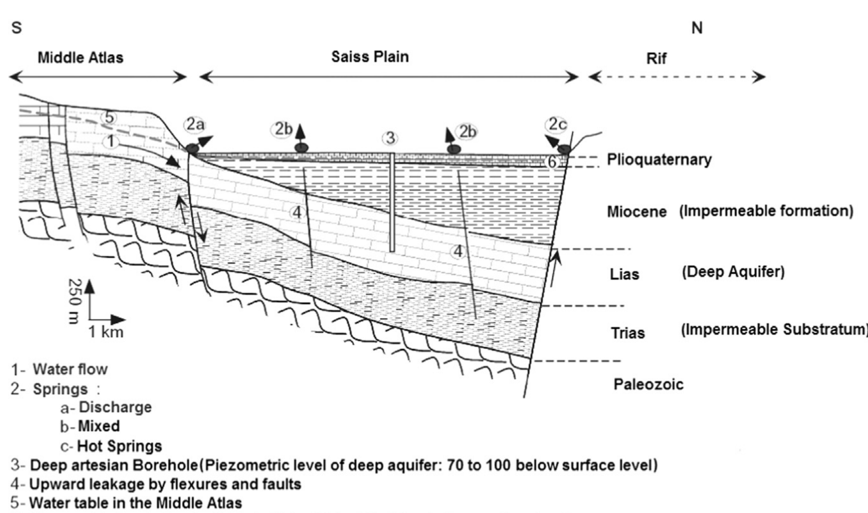

Fig. 3. Hydrogeological context of Saiss plain (extracted and revised from Belhassan et al. (2010)).

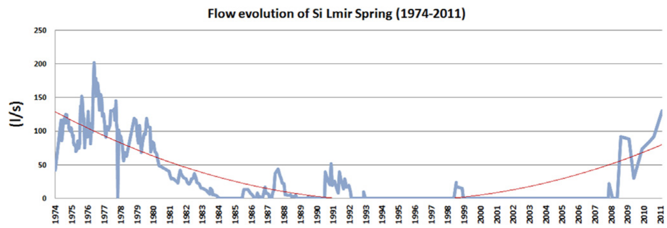

Fig. 4. Evolution of the flow of Si Lmir spring (*l/s*).

Due to the combined effects of the decrease of precipitations since the 80s, which has reduced the natural recharge of groundwater, and increased groundwater pumping for irrigation, Saiss aquifers suffer from overexploitation and the groundwater level continues to decline. The groundwater balance established by in the framework of the MPIWRD of Sebou, showed an annual deficit of 100 million cubic meters (River Basin Agency of Sebou, 2006, 2009). The drawdown of groundwater level has affected the flow of many springs such as Si Lmir which is located in the piedmont of the Middle Atlas Mountains (Figs. 3 and 4).

This issue has attracted the interest of the scientific community and was an object of many researches and studies, producing a large mass of dispersed and heterogeneous data which need to be integrated into a single data warehouse to be used appropriately and efficiently to come out with relevant information. Regarding its importance and the alarming situation of Saiss aquifers, the safeguard of these groundwater resources is one of the main priorities of the water strategy in Morocco. The Sebou River Basin Agency (ABHS) which is the authority in charge of water resources planning and management the basin scale is foreseeing to implement an aquifer contract, as a new groundwater governance and management approach. This contract is being prepared in coordination with all stakeholders concerning with Saiss aquifers. The implementation of this approach cannot succeed without the existence of relevant information on resources and uses.

2.3. Why a STDW for Saiss aquifers?

A significant number of researches and studies were focused on the issue of the Saiss's aquifer overexploitation. Several inventories of water points were made in Saiss plain, mainly by the ABHS (River Basin Agency of Sebou, 2004). As a result of all these works and inventories, significant mass of data varying in space and time. Because of their heterogeneity and their differences, these data remained underutilized and even obsolete, losing their values and the opportunity to have information inducing a better decision making in terms of groundwater resources management in the Saiss.

This application concerns the development of a STDW related to Saiss aquifers. This DW is considered as a platform to develop a DSS like a SOLAP in order to improve groundwater resources management. It focuses on the resource, rather than abstraction. Nevertheless, it can provide information on the type of use. Because of the type of modeling which is multidimensional modeling, it is possible that the model be extended further and more tables might be added to store also data relating to abstraction and use in more details.

If we imagine that we are in the position of the groundwater resource manager of Saiss, among the fundamental questions that he would like to have an answer are:

- What is the evolution trend of the piezometric level at the aquifer?
- What are the mechanisms inducing these variations?
- Is the depth of water table varies with the geographical location change?
- What is the evolution trend of the number of water points in time?
- In what year has recorded the maximum deficit in the aquifer?
- What is the most vulnerable area in terms of water scarcity?
- In which areas are the boreholes concentrated?

There are many questions that a DW can answer through advanced and multi-criteria queries.

3. Results

3.1. Data collection and analysis

A number of studies and surveys were investigated on the issue of the groundwater overexploitation in the plain of Saiss. Several inventories of water points at Saiss plain have been carried out by ABHS, and organized in multiple databases and digital spatial data. Each inventory was focused on a specific part of the Saiss plain. Thus, a considerable mass of spatiotemporal data were generated as the result of all these investigations and stored separately in computers without being able to be used together to get out the relevant information helping in decision making.

Data are stored in different sources (alphanumeric databases, geodatabases and independent files), and in different formats (sheets, shapefiles, rasters…). Data used in this work come mainly from (River Basin Agency of Sebou, 2009; Belhassan et al., 2010; River Basin Agency of Sebou, 2004). Some of these data get a metadata, but the majority does not have any information about data. We cannot talk about data heterogeneity and disparity without thinking to data quality.

Data analysis concerned more than 10,000 water points. Due to data quality and in order to get the same similarities in terms of data, only 9230 water points have been taken into account in the present model.

There are five types of water points encountered in Saiss Plain:

- Well: A hole in the ground made to gain access to an aquifer to obtain water with an average of 40 m depth and 1.5–2 m of diameter. Wells represents like 85% of the number of water points considered in this work;
- Borehole: A deep hole (average 100 m depth) with a small diameter (0.3–0.4 m). Boreholes represents like 7% of the total;
- Well – Borehole: A well developed to borehole for more groundwater abstraction. It represents 5% of the total;
- Piezometer: It's a non-pumping well of small diameter. It is used to measure the elevation of a water table or water quality. It represents 1% of the total;
- Spring: Area where there is a concentrated discharge of ground water that appears as a flow of water at the surface. It represents 2% of the total.

Analysis of collected data has allowed us to distinguish fixed and variable data either in time or space. Data are classified as follows: (Table 2).

3.2. Modeling of the STDW for Saiss groundwater

Regarding data types and multidimensional queries required to response to different issues, the best architecture to design a STDW for Saiss groundwater is a multidimensional database. This structure consists on facts and dimensions tables' linked by keys.

Data analysis showed that several data are associated to water points, such as measure data which vary in time (piezometry,

Table 2
Data classification.

Fixed data	• Data on physical environment (topography, geology, hydrogeology) • Data on socio-economic environment • Data on water point (coordinates, province, municipality, digging date, type, use) • Hydrological and climatic data (hydrographic network, hydro-climatological stations network) • Data on water infrastructures
Variable data in time or/and space	• Data on piezometry (The measure of water level or pressure below surface level) • Data on water point depth (water point depth below surface level) • Data on the permeability: The property of sediments and rocks that allows the movement of water through them • Transmissivity: A measure of the capability of the entire thickness of an aquifer to transmit water; • Data on water points facilities (type of pumps and motors)

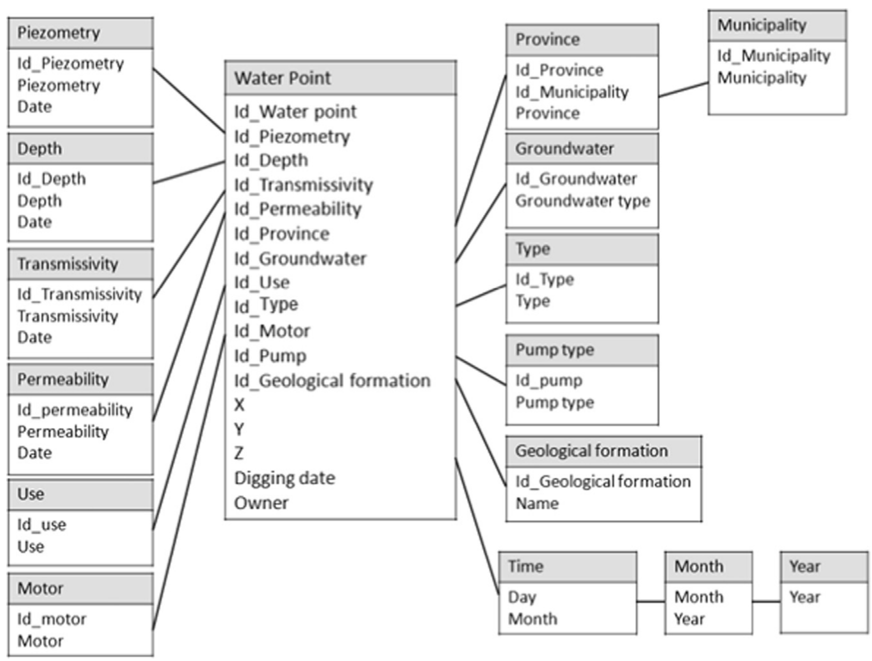

Fig. 5. Conceptual data model of the STDW of Saiss aquifers.

water depth, permeability, etc.), facilities data (type of pumps, type of motors, etc.) and geographic data as well (coordinates, province, municipality, and geological formations). Time dimension has three levels of granularity (day, month and year). In order to adopt an adequate design taking into account this observation, and to allow easy data handling and use, "water point" is taken in this paper as the geographical unit which all data float around. This small geographic unit is the central and the main table of the model (fact table). This table is linked to dimension tables through foreign keys. Conceptual data model of STDW of Saiss aquifers, developed in the the framework of this work, is presented in Fig. 5.

3.3. Exploitation and products of STDW of Saiss aquifer

The main aim of the DW is to answer to dimensional, bi-dimensional or multidimensional queries. These queries allow the join between fact table and dimensions tables, depending on the query subject which might be a measure.

The DW of Saiss aquifer was implemented in PostGIS as a system of spatial database management. DW exploitation has been done through the functionality of QGIS software which is an Open Source GIS. In addition to allowing visualization and spatial viewing, this interface would lead to do multi-criteria and extensive analysis even to make simulation by developing applications. Some of the outputs of this STDW operating are presented below:

Depending on the request and the number of dimensions, the product or the information output might be maps, graphs or tables. As illustration, we try to answer some queries and differently to show the possible ways of information presenting.

- R1: What is the distribution of the types of sampling points on the geological formations of Saiss plain?

The analysis of this request shows that it's a bi-dimensional request. The first dimension is "space" and got three levels of granularity (Saiss plain, geological formation, and water point coordinates). The second one is a thematic dimension "type of water points". Given that the spatial dimension is necessary to respond to this request, water points mapping appears the most suitable presentation. The map presented in Fig. 6, is a product of the use of the STDW developed through the overlapping of layers to take into consideration the different dimensions of the request and their granularities.

The map shows the distribution of the types of water points over different geological formations on the saiss plain. There are 4 types of water points represented in the map (wells, boreholes, well-bores and piezometers).

- R2: What is the piezometric evolution of the groundwater of Saiss at a specific piezometer from 1970?

As well as the first request, this query seems containing three dimensions. The first one is a thematic dimension "piezometery"

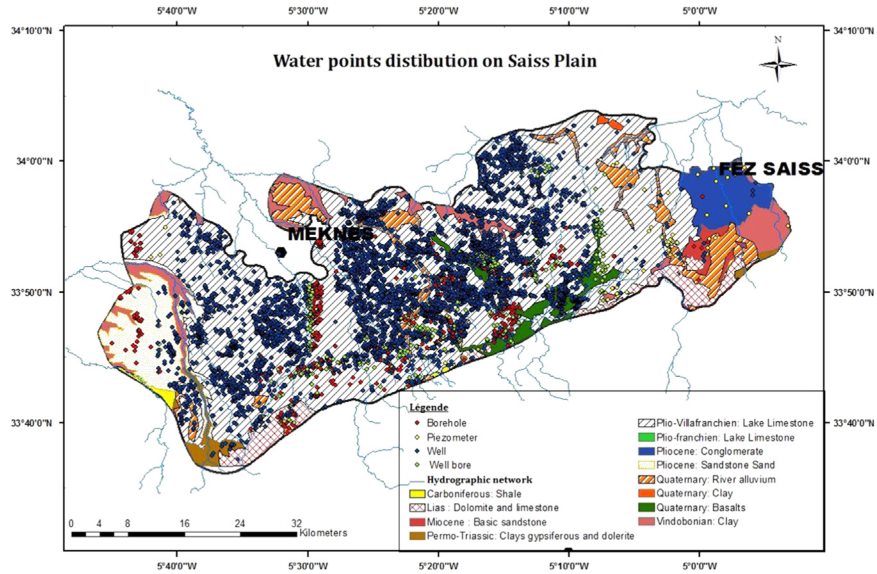

Fig. 6. Map presenting the distribution of water points types on Saiss plain (from STDW Saiss aquifers).

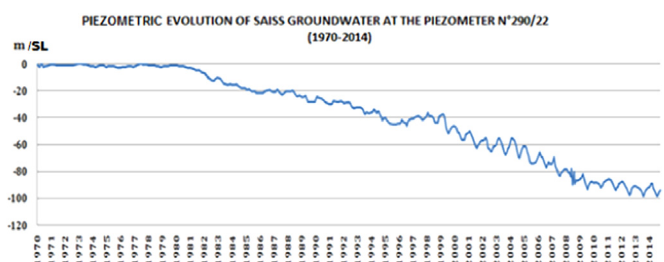

Fig. 7. Piezometric evolution of Saiss groundwater at the piezometer N°290/22 (from the STDW of Saiss aquifers).

which is a measurement, the second is “time” indicated by “from 1970”, and the third one is “space” indicated by the number of the Piezometer or the water point. Given that time dimension is the most important one in the request and because the question is to represent the water level evolution in one water point, the best way to represent the response seems to be at form of graphic.

The graphic which object of the Fig. 7 is another type of the STDW of Saiss aquifers mining product. It shows the piezometric evolution of groundwater at the piezometer N°290/22 for 44 years (from 1970 to 2014). Due to the drought period which started in 1980, this piezometer recorded a significant decrease of 100 m in 34 years with an average rhythm of almost 3 m per year.

- R3: Compare the evolution of two springs' flows on the plain of Saiss since the 80's.

For this query, the representation would be better with a graph comparing the evolution of the two springs' flows. The example of two springs has been taken (Bittit and Ain ribaa). They are

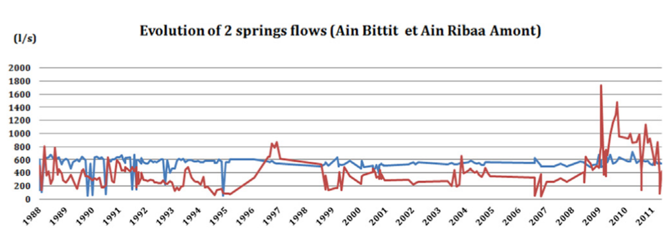

Fig. 8. Graph on the evolution of 2 springs' flows in Saiss plain: Bittit and Ribaa Amont. (from the STDW of Saiss aquifers).

overflowing springs of free lias located along the contact zone of the plain with the Middle Atlas (Fig. 3). The flows of these springs have been compared during the period since 1988–2011 (Fig. 8). It is obviously possible to have this chart next to a map showing the location of sources and even to have the time series of flow variations in tables.

Fig. 8 as the product of the STDW of Saiss aquifers shows that the flows of these important springs in the Saiss plain change in time almost in the same way from 1988 to 2008. The trend of Bittit flow seems standing regular around 600 l/s over the period since 1988–2011, with small changes, especially some decreasing during dry years such as the years of 1988, 1990, 1992 and 1995. Regarding Ribaa Amont, it recorded a significant flow decrease in 1995 (dry year) and an upturn in wet years (1996 and 2009). The flow reached till 1700 l/s in 2009. The statement is that there is a positive correlation with precipitations and springs flows.

4. Discussion

The application of data integration and building a STDW for Saiss groundwater resources shows that data integration was the largest and the most complex step in the process. This is the step requiring and consuming the most of time and effort. Its complexity consists also on requiring knowledge in various disciplines such as database design and management, water resources and data integration technologies. Given its virtues and its ability to offer the user, manager or decision maker, a unique view and synthesized selected, treated and organized data into a unique and homogeneous dataset, data integration is considered as an efficient and even the key for DSS building for groundwater management. Groundwater STDW can include other various kinds of data such as socioeconomic data and be extended to response to more specific queries. This dataset could be used as well as a data platform for constructing groundwater models and be coupled to hydrogeological models.

Nevertheless, data integration is facing a lot of constraints, limitations and challenges. During this process, many constraints were encountered, particularly related to data heterogeneity and distribution, spatiotemporal data management, data quality and uncertainty and STDW model management.

4.1. Data heterogeneity and distribution

Data heterogeneity is considered as one of most complicated issue in data integration. From data analysis, three kinds of data heterogeneity are distinguished:

- *Syntactic heterogeneity*: information is in different languages and formats. *Example*: piezometric information for one well could be found at (*.shp, spreadsheet, tables, raster…);
- *Schematic or structural heterogeneity*: a big part of data is stored and organized in different databases which have its own structures and models. The constraint in this case is managing all databases structures to provide a unique schema which could hold on the most interesting and useful data for building the data warehouse. *Example*: three databases linked to water points were made in 3 different areas at Saiss plain. Each one has its proper model and schematic.
- *Semantic heterogeneity*: Disagreement on the meaning, interpretation on intended use of data in different data sources. *Example*: a water point could be identified by different way (name, ID, WP_ID, Code…) like shown below:

WP_ID	***Name_spring***	***Code_PP***
54/15	*Aioun Ablouz*	*AEPI61_P1*
114/22	*Ain Akkouss*	*AEPI63_P1*
Spring 1	*Spring 2*	*Spring 3*

WP_ID: Water point identifier (ABHS identifier).
Name_spring: Name of the spring.
Code_PP: Water point Identifier (identifier chosen during the elaboration of the inventories).

4.2. Spatiotemporal data management

Several collected data were alphanumeric and had to be specialized, particularly the water points. As indicated above, these data exist in different sources with different formats. Some water points have coordinates (X, Y) but in different projection systems. Most of them have a “Lambert_Conformal_Conic” projection. In order to get a unified projection and to allow the superposition of the various spatial data layers, a unique projection in the DW is “Lambert_Conformal_Conic”.

Regarding the dimension “Time”, it was necessary to check all measurement data sets, to unify them to avoid any type of heterogeneity. For example, different time series of piezometric measure were found of one piezometer. Another example is it found that different measures of water depth have been taken simultaneously at the same piezometer. The updating of these time series considered also as a constraint involving the viability and relevance of STDW in supporting decision making.

4.3. Uncertainties in data integration process

Apart from the difficulties relating to data heterogeneity and distribution, a crucial question arises as to data reliability and quality, and therefore the relevance of produced information. In this work, the uncertainties that arose during data integration process lies in the source of data, in data itself as well as in each step of ETL process (Sarma et al., 2009). Each step is considered as an uncertainty intrusion occasion.

Uncertainty is not the issue of data and data integration processes only, but it is present also in the translation of multidimensional queries through the DSS to information, particularly via keywords or indexing.

4.4. STDW development and management

Given their capability to provide relevant and concise data for decision makers and groundwater managers and users, spatio-temporal data warehouses is actually gaining in importance. However, the development and the management of STDW are facing many challenges in relation with:

- The modeling process: this phase is delicate in the sense of needing to determine dimension and fact tables based on the STDW user requirements, and to define data aggregation and granularity level. The integration of new data does not cause any difficulty if granularity has been well defined. If the granularity is too large, reload operation is required. Then the model is updated;
- The development of logical schema, storing and physical implementation of the STDW;
- The choice of adequate tools in order to gain time, money and efforts;
- The improvement of multidimensional queries performance through SOLAP tools based on defining the right index structures (spatial or non spatial). Enhancing queries performance could save time to managers, especially in emergency situation such as in drought management;
- The refresh and the maintenance of the STDW to remain it updated and useful for users.

Actually, all these issues need to be studied further and be subject of future researches.

5. Conclusion

Studies and research in the field of groundwater resources continue to produce data of various types, often stored in different data management systems. These data might have several dimensions (thematic, spatial and temporal). These data remain underutilized and obsolete in the absence of data integration into a single, synthesized and easy to handle system.

Data integration consisting in the process of the extraction, transforming and integration of heterogeneous data from different sources in a data warehouse is an efficient solution for data mining and analysis to produce information helping to real-time decision. In the field of groundwater, data are varying in time and space. In this case, we talk about a spatiotemporal data warehouse. A water resources manager would be interested to follow the evolution of the resource in its spatial environment to master, manage and plan it in an integrated and sustainable manner. A groundwater user is interested as well as the manager on the evolution of this resource which is often in over exploitation, particularly in arid and semi-arid areas, in order to manage its use and to be aware of the risks linked to groundwater overexploitation and their impacts.

The schematization of the spatiotemporal data warehouse in a hierarchical relational schema allows in one hand to take account of dimensions granularity, and in the other hand to meet the different multi-criteria queries that can be asked by the manager through a friendly interface of a SOLAP constructed based on the DW. A SOLAP on groundwater would enrich and support research, planning and management of groundwater in conjunction with surface water. This SOLAP as a DSS tool, could be developed also to be a research and educational tool and intelligent mobile application, useful to students as well as farmers in order to understand easily the issue of the overexploitation through, for example, the simulation of different scenarios of the groundwater evolution.

Data integration is an effective solution in terms of storage, organization and prioritization of data into a single dataset in the field of groundwater planning and management. However, it still facing many difficulties and limitations related to data heterogeneity and disparity, spatial and temporal data management and uncertainties occurred during the DSS development, particularly during ETL process and the DW building.

References

Bauzer-Medeiros, C., Carles, O., Devuyst, F., Hugueney, B., Joliveau, M., Jomier, G., Manouvrier, M., Naija, Y., Scemama, G., Steffan, L., 2006. Towards a data warehouse for road traffic. In: EDA'06, Versailles, France. CNRS- LAMSADE. Paris-Dauphine.

Belhassan, K., Hessane, M.A., Essahlaoui, A., 2010. Groundwater-surface water interactions: catchment area of Oued Mikkes (Morocco). Hydrol. Sci. J. 55 (8), 1371–1384. http://dx.doi.org/10.1080/02626667.2010.528763.

Butenuth, M., Gösseln, G.V., Tiedge, M., Heipke, C., Lipeck, U., Sester, M., 2007. Integration of heterogeneous geospatial data in a federated database. ISPRS J. Photogramm. Remote Sens. 62, 328–346.

Cali, A., Calvanese, D., De giacomo, G., Lenzerini, M., 2004. Data integration under integrity constraints. Inf. Syst. 29, 147–163.

El Sawah S., Hicks A., Manger P., Athanasiadis I.N., Croke B.F.W., Jakeman A.J., 2011. A web-based platform for integrated groundwater data management, in: Proceedings of the 19th International Congress on Modelling and Simulation, Perth, Australia. ⟨http://mssanz.org.au/modsim2011⟩.

Franconi E., Sattler U., 1999. A data warehouse conceptual data model for multidimensional aggregation: a preliminary report, in: Proceedings of the Workshop on Design and Management of Data Warehouses. Heidelberg, Germany.

Inmon, W.H., 1996. Building the Data Warehouse, second ed. John Wiley & Sons, New York.

Jost, C., 2001. Decision support system for concerted water management. An integrated approach to social behavior and risks in Loyalty Islands. Int. J. Geomat. 11 (3–4), 443–467.

Maceachren, A.M., Kraak, M.J., 2001. Research challenges in geovisualization. Cartogr. Geogr. Inf. Sci. 28 (1), 3–12.

Mattison R., 1996. Data warehousing: strategies, technologies and techniques. New York.

Nilakanta, S., Scheibe, K., Rai, A., 2008. Dimentional issues in agricultural data warehouse designs. Comput. Electron. Agric. 60, 263–278.

Refsgaard, J.C., Hojberg, A.L., Moller, I., Hansen, M., Sondergaard, V., 2010. Groundwater modeling in integrated water resources management-visions for 2020. Groundwater 48 (5), 633–648.

Regional Directorate of Agriculture of Fez–Boulemane and Meknes–Tafilalet, 2014. Feasibility study for the saveguarde of the irrigated plain of Saiss located in Fez–Boulemane and Meknes–Tafilalet regions. Meknes–Tafilalet, Morocco.

River Basin Agency of Sebou, 2006. Study of hydrogeological synthesis of the plain of Fez–Meknes: hydrogeological synthesis of two aquifer units of Fez–Meknes Basin. Fez, Morocco.

River Basin Agency of Sebou, 2009. Master Plan for integrated water resources management of sebou basin. Evaluation of the groundwater resources of the aquifer of Fes–Meknes. Fez, Morocco.

River Basin Agency of Sebou, 2004. Inventory of beholden of irrigation, drinking and industrial water from the Groundwater of Fez–Meknes plain. Fez, Morocco.

Sarma, Anish, D., Luna, Dong, Halevy, A., 2009. Managing and Mining Uncertain Data. Springer, New York.

Van Damme, M., 2010. Data warehouses in space field for the forest inventory. Data Struct. Algorithms < dumas-00538909 >.

Vidal A., Belouze P., Laloux S., Lepoutre D., 1997. Water management and spatial data integration Ingénieries - EAT. pp. 65 –8. <hal-00461024 >.

Wu, L., et al., 2001. Design of data warehouses using metadata. Inf. Softw. Technol. ⟨http://www.etltools.net/⟩ (accessed 12.02.16).