

Geohydrological GIS of geohydrological database (Wiley title)

New proposed title:

Geospatial information in hydrogeological studies

P. Wojda¹, S. Brouyère^{1,2,*}, A. Dassargues¹

¹Hydrogeology & Environmental Geology, GEO³ Department and ²Aquapôle – ULg,
University of Liège, Belgium

*Corresponding author: Serge.Brouyere@ulg.ac.be

Table of contents

ABSTRACT	2
1 INTRODUCTION	3
2 THE SPECIFICITY OF HYDROGEOLOGICAL INFORMATION	5
3 FROM PAPER TO DIGITAL STORAGE	9
3.1 GEOGRAPHIC INFORMATION SYSTEMS IN HYDROGEOLOGY	9
3.2 PROCESSING OF HYDROGEOLOGICAL INFORMATION	11
4 DIVERSITY OF HYDROGEOLOGICAL DATA MODELS	12
5 TOWARDS SEAMLESS HYDROGEOLOGICAL INFORMATION EXCHANGE	16
5.1 HYDROGEOLOGICAL DATA MODELLING.....	18
5.2 OBJECT-ORIENTED MODELLING AND UML	19
5.3 STANDARDS IN GEOSPATIAL INFORMATION	22
5.3.1 <i>Why to standardize hydrogeological data models</i>	22
5.3.2 <i>Standardization institutions</i>	23
5.3.3 <i>Geospatial metadata</i>	24
5.3.4 <i>Hydrogeological and geological domains standards</i>	26
6 CONCLUSIONS	28
7 REFERENCES	30

Key words: hydrogeological data, hydrogeological information, geomatics, UML, data modelling, digital data, geoinformation standards

ABSTRACT

The purpose of this paper is to make a review of hydrogeological information modelling objectives, techniques and problems, and finally to propose a solution for data integration and transfer. Hydrogeological data and information management is crucial for efficient integrated water resources administration, protection and exploitation. In order to assure an effective management of groundwater resources, a close cooperation and information exchange between different actors at different levels is necessary. The information flux must be optimised between local producers and potential users of hydrogeological data: regional and national institutions, as well as international and trans-boundary analysts and managers. Paper, as a carrier of hydrogeological information, has already been replaced by different digital forms. However, an enormous diversity in applied formats and representations makes hydrogeological information extremely difficult to transfer in an automatic way. If this transition is not effective, misinterpretations of hydrogeological data may appear. The proposed solution consists in creating a domain specific standard for the hydrogeological information exchange, using recent geomatics advances, taking into consideration appropriate ISO/TC211 and OGC geographic information standards. Such a new electronic carrier of hydrogeological information facilitates a dialogue between interested parties using web services and built exchange mechanisms, interdisciplinary collaboration, international harmonisation for environmental sciences.

1 INTRODUCTION

Hydrogeological data and information management is crucial for efficient integrated water resource administration, protection and exploitation. Hydrogeological data are complex because hydrogeological properties and parameters are spatially and temporally distributed. Furthermore, field work and data acquisition being very expensive and time-consuming, data retrieved from hydrogeological field tests and field campaigns should be available for any further research to reduce costs and efforts.

In order to guarantee adequate groundwater management, users and decision makers need a clear structuring of such available information. Hydrogeological data need to be accessed and transferred between different interested actors or organisations for their specific uses and applications. The transfer may be performed from a local level (on-site measurement teams), to national and international institutions dealing with water resource management issues (Figure 1). Up-scaling and downscaling of hydrogeological information from a scientific point of view being a very sensitive topic might be treated separately (see Chapter 9: Statistical Upscaling and Downscaling in Hydrology for more thorough discussion). However, technical aspects of information structuring and sharing should be covered by the GIS domain hydrogeologists.

Local specialists require and create hydrogeological data from surveys. In such context, they need hydrogeological data coming from heterogeneous sources: neighbouring sites, monitoring networks, national databases. At the same time, local specialists are also potential producers of large amounts of data obtained in the scope of their field and laboratory observations and measurements. At a higher level, to manage groundwater resources for local

communities in terms of quantity and quality, governmental administrations also need hydrogeological information. Governmental administrations must also report on European Union Water Directives implementation to national and international institutions. Finally, the management of groundwater resources is international because groundwater bodies or aquifers usually do not fit with national borders and many transboundary watersheds should be managed by water authorities from different countries. As a consequence, seamless hydrogeological information exchange in a multi-language environment is needed for management and reporting.

<Figure 1 here>

Nowadays, geological and hydrogeological information is increasingly recorded, stored and communicated in digital form. Paper, as a traditional carrier of data and information, in the form of classical hard copy media, reports, geological maps, cross sections or sketches, is less and less used for storage and transfer purposes. It is progressively becoming auxiliary to electronic carriers, the latter being coupled with computers and computer networks (Michalak, 2003a). This trend is accompanied by the growing use of computational methods to carry out geoscientific tasks (Brodaric et al., 2004). Using well-designed geospatial databases coupled with GIS allows for improving the speed and accuracy in data processing. New methods of data processing produce new forms of information, display them differently, changing our perception of the reality. Such a technological breakthrough in other related environmental domains significantly influences the hydrogeological domain.

To meet the requirements of data storage, sharing, and transfer, hydrogeological data modelling has to be performed, mainly at two levels, using specialized methodologies. First, clearly designed hydrogeological data models are needed at local and regional levels. Secondly, commonly accepted international data transfer standards should be developed and

used. Data models should also respect wider data modelling principles and standards, from information, through geoinformation and geological information, to, finally, hydrogeological information, which specialises them. Such standards increase hydrogeological information availability and transparency, by ensuring an adequate data organisation and an accurate hydrogeological content documentation by metadata.

This article presents the recent progresses in the domain of hydrogeological geoinformation storage and transfer. In the first section, the specificities of hydrogeological information are presented together with several definitions from geomatics for the sake of clarity of further explanations. Then, traditional and digital techniques of hydrogeological geospatial information storage and visualization are presented and their limitations are identified and described. Examples of several existing hydrogeological projects are summarized. In the third section, new solutions for seamless geospatial hydrogeological data management and transfer are proposed. In this regard, the most recent and ongoing projects dealing with geological and hydrogeological information modelling are presented. In the conclusions, new directions for hydrogeological system integration within wider environmental systems are outlined.

2 THE SPECIFICITY OF HYDROGEOLOGICAL INFORMATION

Hydrogeology, as one of the environmental sciences, is strongly related to other domains such as geography, geology and hydrology (Figure 2). These branches may be considered as mutually dependent, each of them having however their own particularities. The interdependence between these domains has to be taken into consideration and clear relations between information from the different fields should be identified and established. In relation

with the specific hydrogeological domain, the geographic domain should organise and deliver all the information and concepts on the localisation, spatial extent and topology (describing the relationships amongst related or neighbouring elements such as points, lines or polygons, ISO19104 DIS:2002) of any information. The hydrologic domain should organise and deliver all information on the components of the water cycle interacting with groundwater. Finally, the geological domain should organise and deliver information on geological units and structures which contain groundwater. The information content and organisation of the specific hydrogeological domain should not overlap with any of the aforementioned domains. Furthermore, hydrogeological data and processes have their own particularities and they may require a detailed categorization and specialization, with a domain data model to be developed for hydrogeological data storage and transfer. As an example, drilling, wells and piezometers engineering aspects should be treated by one group of domain specialists, groundwater chemistry by another group, and groundwater flow modelling aspects separately by other experts.

<Figure 2 here>

From a technical point of view, geospatial information can be subdivided into two main categories: generic and specific (Michalak, 2003a). Generic type of information is defined as common to many domains. For instance, geological and hydrogeological maps are elaborated on the same basis as other maps. Most of generic geospatial information coordinates are based on two-dimensional Cartesian (x,y) or any other reference systems. Rules concerning geometry and topology are the same, but the difference is in the semantic aspect and content of information. For generic information, the re-use of already established standards, communication protocols, and processing rules is very convenient and ensures a rapid development of new concepts and software applications. Hydrogeological information often requires a three-dimensional reference system (x,y,z) or a specific combination of dimensions

in 2D, for instance to describe boreholes or well depths series or hydrogeological cross sections.

In addition, since many hydrogeological properties are time-dependent, time reference and time topology issues have to be considered explicitly (x,y,z,t) . Classical time-dependent hydrogeological variables and properties are, for instance, piezometric levels, solute concentrations or pumping rates. The issue of time scale and time topology is standardized in ISO 19109 – Geographic information – Temporal Schema (ISO, 2002) and discussed in details in Michalak (2005) for geological purposes.

As compared to other related domains, one of the main characteristics of hydrogeological data is their potential variability in time and space. There are also very specific types of hydrogeological information, such as (adapted and extended from Michalak, 2003b):

- the definition and the spatial extent of specific hydrogeological features such as aquifer/aquitard/aquiclude formations and groundwater bodies, specializing geological features;
- specific physico-chemical parameters describing the underground, such as hydraulic conductivity, effective porosity, geochemical composition of groundwater, all these parameters being spatially distributed and variable in time;
- hydrogeological cross-sections where geological units are described in terms of hydrogeological properties, together with interpreted groundwater levels;
- data describing the equipment of wells and piezometers, such as casings, screens, pumps, gravel packs, sealings;
- specific hydrogeological observations and measurements, namely piezometric and groundwater chemistry measurements, recharge/discharge rates, base flow;

- descriptions and interpretations of hydrogeological field experiments such as pumping tests and tracer tests.

These specific types require specific solution in data modelling. Spatial extent of hydrogeological features should be described and encoded using a common convention of geographical features. However, hydrogeological phenomena can be 3-dimensional which combined with a spatial variability in z-direction can introduce an additional complexity: geological layers forming aquifer/aquitard/aquiclude bodies are characterized by different properties. This is being introduced in GML 3.x version (ISO 19136:2007).

Furthermore, hydrogeological data are primarily collected using sampling points, and then they are interpreted. Therefore, particular hydrogeological data structures should allow for a distinction between sampling point data, interpolated data and interpreted information. In the geological domain this solution is being partially implemented in the GeoSciML project (Sen and Duffy, 2005) through structuring of special features for handling of sampling points (or any other spatially extensive features), specimens and then observations stored as coverages (OGC, 2007). In the hydrogeological domain specific spatial interpretation features might be added to handle, for instance, hydraulic conductivity or tracer test interpretations on groundwater bodies (Wojda, 2009).

Difficulties may arise when creating such additional hydrogeological capabilities and features within existing data structures. However, crucial specifications exist to assist the hydrogeological community in creating its own Application Schema, and similar exercises have already been conducted by other communities (XMML, GeoSciML mentioned further in the chapter).

3 FROM PAPER TO DIGITAL STORAGE

Paper-based storage is still widely used as hydrogeological information carriers, as non-graphical forms such as texts, tables and forms and graphical forms such as images, maps, and cross-sections. However, corrections or updates are very difficult to implement, and they generally require the creation of completely new documents. The visualization of hydrogeological observations and measurements is also limited. Finally, since hydrogeological data are usually numerous, processing of their paper form is very time-consuming and effort-demanding.

3.1 Geographic Information Systems in hydrogeology

Recent requirements for real-time data delivery and analysis combined with automatic data transfer between interested parties, monitoring networks and remote sensors have entailed the evolution from paper to electronic carriers. Furthermore, due to the importance of hydrogeological information in water resource management, its considerable amount and financial values, the hydrogeological community needs a flexible and structured way of digital data and information storage. Such tasks as data structuring and management are more and more embedded in Geographic Information Systems (GIS), which deliver information for decision makers and specialists in environmental domains dealing with spatial and temporal information.

GIS include different components, such as data, hardware, software, procedures, operators and analytical problem statements (Meeks and Dasgupta 2004). At first, GIS were used rather to create paper maps to analyse and display geospatial data. The map content, once introduced into a computer system, was designed to correspond with its paper image. Rapidly, a very useful discovery was made that the content remaining in the system is sometimes much more

valuable than the paper representation itself. Digital records of information combined with GIS have offered the possibility to derive new information, more suitable and specific for a given problem, but also further data processing, updates, or data transformations into other formats. GIS offer the user the opportunity to capture and to collect geospatial and non-geospatial data, where data sources can be numerous such as scanned paper maps, aerial photographs, remote sensors, field observations and measurements. Moreover, data stored in GIS do not require to be cut into separate sheets linked to scales, map projects and graphical representations, only the reference system is mandatory. Hydrogeological information can be grouped in layers, dynamically processed at a chosen scale, and displayed using a desired format by superposition with other thematic layers.

There are two classical ways of storing and representing geoinformation in a digital form: vector and raster formats. Hydrogeological information being difficult and expensive to obtain, the hydrogeological continuous environment can only be sampled on the point-type basis using available drillings, piezometers, wells or other monitoring stations, or on a line-type basis, using geophysical tests. The location of such point- (e.g. well or piezometer), line- (e.g. water gallery or excavation) and polygon- (e.g. aquifer or groundwater body extent) type features is stored in a vector format. Complementarily, results of observations and measurements may be represented in the form of a discrete coverage, for instance a raster, or they may be spatially interpreted (interpolated or extrapolated) in order to create a continuous coverage, where the property varies continuously across the domain.

These general observations are particularly valid in the field of hydrogeology, where natural hydrogeological units do not necessarily fit with administrative borders, or where natural water resources should be managed and protected using integrated, multidisciplinary

approaches. GIS have significantly influenced hydrogeological field researches, laboratory activities, and observation methods. Using a structured geospatial database under GIS, any potential user is now able to easily access different hydrogeological data by selecting hydrogeological features by attributes or spatial queries. For instance, groundwater samples and groundwater quality measurements for selected wells can be accessed in order to establish a groundwater body quality status within the selected aquifer. Furthermore, piezometric measurements can be accessed and updated (Figure 3).

<Figure 3 here>

Hydrogeological studies require also numerous GIS tools when overlaying layers of information, from general geographical information, through hydrological and geological information to specific hydrogeological information. For instance, for groundwater vulnerability mapping or pollution risk mapping much information needs to be analysed. A geology map needs to be combined with topography information, soil-type and land-use maps. Moreover, a hydrological map and water table depths are overlaid and analyzed in the context of possible contaminant flow lines, advection and dispersion. In order to apply appropriate aquifer protection or decontamination measures, available water intakes localisation and exploitation schemes need to be known.

3.2 Processing of hydrogeological information

In order to conduct hydrogeological studies including water budgets, groundwater flow and contaminant transport modelling, or groundwater exploitation schemes, substantial amounts of data and information are required and, most often, automatically processed. Information processing is seen as an implementation of different algorithms in order to derive new information better suited for different uses (Michalak, 2003a). Processing of information stored in the paper form is never automatic, thus very time-consuming. For instance, a paper hydrogeological map, or a raster image as presented in Figure 3 is neither changeable nor

interactive. Any further processing of presented piezometric values or hydrogechemistry measurements would be manual, which may add errors when copied elsewhere.

4 DIVERSITY OF HYDROGEOLOGICAL DATA MODELS

Almost every administration, municipality, water and environmental authority, and research organization have developed their own hydrogeological data models, implemented in different database management systems (Rodríguez *et al.*, 2007). The collection of hydrogeological data, their verification, data validating, and the construction of the databases and data services are regulated in almost every European country (Szalkai *et al.*, 2007), as well as other non-European countries such the USA, Australia or Canada. Hydrogeological data are managed using both non-spatial and spatial systems, which are due to the wide thematic range and diverse types of requirements.

A detailed European survey on groundwater databases and related information has recently been performed by the FP6 EC eWater project (www.ewater.eu). This survey proposes a classification of the following hydrogeological data types: well, water exploitation, monitoring - time series, maps, and related metadatabases. Hydrogeological data are collected at local, regional and national levels. Local municipalities, as well as local offices, are the most active data collectors in the following countries: Italy, France, Denmark, Holland and Hungary. Some of these data are directly loaded to the national databases in Denmark and France. Otherwise, local institutions have to supply these data to the regional/national data collection authorities. At the regional level, data collection is performed by provincial and regional authorities, regional water management organizations, water management agencies of river basin authorities and regional offices of research institutes. At the national level, each country performs such data collection, except from Austria. This is the responsibility of

geological surveys in Lithuania, the Slovak Republic, the Czech Republic, Denmark, Sweden, France and Holland. In Slovenia and Hungary the situation is slightly different – two institutions manage the databases. At the national level, most of the organizations are responsible for making hydrogeological data available to the public. In several countries data are completely free, or the users can be charged for the service itself. Furthermore, there exists a big diversity of informatics systems, for GIS: from ArcGIS, through AutoDeskMap and MapInfo Professional, to GeoMedia Professional, and for DBMS: from MS Access, through SQL Server to Oracle. More detailed information and hydrogeological databases review for each concerned European country can be found in the on-line deliverables of the FP6 EC eWater project.

The diversity of standards and data schemas leads to difficulties in communication and data exchanges, which is particularly critical in the context of transboundary groundwater body management and information sharing. Different water agencies use different data formats. This situation makes it more complex to automatically exchange data coming from multiple sources or to communicate results of any hydrogeological study. Furthermore, hydrogeological data being widely used in other environmental domains and multidisciplinary studies, they are not easily available when they are dispersed in many formats and in many places. First, potential users are not even conscious about their existence due to the lack of any centralized hydrogeological data or metadata catalogue. Secondly, it is often difficult to access hydrogeological data because owners of hydrogeological information are hardly known. Data access privileges are not clearly exposed to users, or finally the proprietary data format is simply unknown or requires additional software licences. Last but not least, the lack of internationally accepted hydrogeological data storage and transfer standard makes it difficult to use open web standardized services, such as WMS, WFS, and WCS. As a

consequence, many existing powerful tools for data management, visualisation or analysis, based on these services capabilities, can not be applied.

Several interesting European and non-European projects, taken as typical examples of the existing hydrogeological data modelling are compiled in Table 1. Based on the analysis of the existing models, the following conclusions can be drawn. Despite the most common elements such as the technical description of wells and associated observation and measurements on piezometry and groundwater quantity and quality, it appears clearly that most existing data models were developed for relatively specific applications and according to different requirements. The identified models are described using different modelling designs and notations, and only a few of them use modern technologies such as Model Driven Approach and UML notation, or follow standards such as ISO 19136, described by GML for geographic data. At the ontological level, models propose different hydrogeological feature types and relationships. At the semantic level, the definitions and meanings of hydrogeological feature types are not common to all the models, leading to difficulties in further data understanding and interpretations. As a consequence, there is no existing most complete data model for the hydrogeological domain. Such a model should enable to deal with **a hydrogeological project** as a whole. First, data about the project localisation, performed hydrogeological studies, people in charge of different project aspects should be available. Second, information about groundwater natural and man-made features such as springs, sink-holes, trenches or wells, should be accessible. Furthermore, observations and measurements performed during hydrogeological field work and experiments such as pumping and tracer tests should be easily identifiable and obtainable.

<Table 1 here>

5 TOWARDS SEAMLESS HYDROGEOLOGICAL INFORMATION EXCHANGE

Due to the problems identified in the previous section, such as availability, accessibility, and exchange of hydrogeological information caused mostly by the hydrogeological data model diversity and interactions with other domains, the hydrogeological community needs to establish one public information exchange standard. Without any standardization of hydrogeological data and other types of data transfer, it is very expensive and difficult to exchange data between different producers and users in an efficient way. A first solution could be to store hydrogeological data, together with data coming from other domains in a central database (Figure 4). However, such central databases are never up-to-date, they are very expensive to maintain, and poorly enriched with additional tools, which limit their practical use by domain specialist.

<Figure 4 here>

The alternative to centralisation of information is to develop standards for data exchanges between systems. Standardisation can be considered mainly at two levels. At the local level, the same data model can be considered by different users. At the higher information exchange level, communication interfaces and exchange formats can be standardized. As it would surely be very difficult or even impossible to achieve an agreement of all users on a local data storage model, a standard for hydrogeological information exchange should be established and the use of web services should be promoted (Figure 5). A Web Service is defined as a software system designed to support automatic and interoperable machine-to-machine interaction over a network, using XML, specialized to GML-compatible application schemas (Cox *et al.*, 2002).

<Figure 5 here>

Using the option of standardization of data exchange protocols, each user or group of users can establish or keep their own hydrogeological data model at the local level. The model may be implemented in any database management system, specific to the identified needs and applications. In order to exchange hydrogeological information, particular data models should be mapped to fit standardized web services and a future hydrogeological information exchange model. The schema mapping between local and standard data models can easily be developed by geoinformation specialists. Once this work is accomplished, data exchanges may be performed by standardized OGC web services in both ways: (1) local hydrogeological data can be made available for other users and domains with GML-compatible clients, following access restrictions, (2) local users can easily access other necessary hydrogeological and non-hydrogeological data and they add them to their databases, analyses and interpretations.

As a practical example, the following “case-study” can be given. A hydrogeologist needs to perform an advanced environmental analysis in order to establish a groundwater resource exploitation scheme. In the proposed solution, specialized GML-compatible tools can be used, and needed data are delivered by different web services, using the hydrogeological data exchange standard. A Web Map Service delivers geo-referenced neighbouring maps on which the above mentioned hydrogeological features can be drawn. Additionally, a Web Coverage Service delivers a coverage with spatially distributed transmissivity and hydraulic conductivity values for the studied aquifer. A Web Feature Service delivers a detailed description of monitoring wells and piezometers, together with associated observations and measurements. To adjust the analysis, a Sensor Observation Service might deliver the latest data on piezometric head levels retrieved by automatic sensors located in the field. In order to

finalize the investigations, climatic, land-use and topography data can be provided by other services coupled with thematic distributed databases.

5.1 Hydrogeological data modelling

Management, handling, and access to hydrogeological information depend mainly on four main categories of models, namely: mental, conceptual and more formalized logical models, leading finally to physical models of hydrogeological data (Figure 6). In the following section this formalism, used as a traditional and rigorous way of developing a model in geomatics, is described for the specific case of hydrogeological data modelling.

<Figure 6 here>

The mental model contains definitions, descriptions, and understanding of concepts and physical laws governing groundwater, flow and transport processes.

The conceptual model contains identified and defined existing hydrogeological entities and objects as well as relationships between them, for instance: a well occurrence is used to sample a groundwater body, which is hosted by an aquifer. The same well is also used to exploit the aquifer, which requires additional quantity and quality observations and measurements. As a consequence, in the conceptual model, several entities have to be defined, for instance: Well, Groundwater Body, Aquifer, Geochemistry Measurement, Piezometric Head Level Measurements, Groundwater Extraction Volume, together with different relationships amongst these entities. As data are specific to the hydrogeological domain, this issue must be addressed directly by domain specialists, familiar with the geoinformation context. The conceptual model can be described using a semi-formal (free charts) or formalized (with defined semantics) notations, such as Entity-Relationship (ER), Unified Modeling Language (UML), or EXPRESS, but it does not depend on the technology nor change with different logical and physical implementations.

Based on the conceptual model, the logical model describes the structure of hydrogeological data, using object-oriented modelling, in the form of definitions of each hydrogeological entity with all its attributes, operations, methods and behaviours. All the identified relationships have to be established. These are, for instance, the associations between aquifer sampling features with samples and observations made on these samples, or relationships between different steps of hydrogeological specialized field tests, such as pumping or tracer tests. Contrarily to the conceptual model, the logical model is technology dependent (e.g. ER or UML), but it does not depend on the implementation platform.

Finally, the physical model is no longer specific to hydrogeology, it is only dependent on the implementation platform.

5.2 Object-oriented modelling and UML

In the hydro-informatics domain, object-oriented modelling may be seen as a new solution for reducing the complexity of data models and software development (Subieta, 1998). To decrease complexity and convolution in any development process, one has to use decomposition and abstraction methods. Decomposition is used to divide any problem into more elementary sub-problems that can be treated individually. Abstraction enables to eliminate or to hide less important parts of the problem within a given context and at a given level of analysis. Furthermore, object-oriented models serve to match conceptual models and physical models with natural behaviour of human beings and their perception of reality.

Several fundamental object-oriented concepts are objects, classes, inheritance, abstraction and polymorphism. The geospatial object represents an instance of a class, which is based on the object-oriented paradigm, coming from UML (OMG, 2001, 01-09-67). The class is a

descriptor of a set of objects that share the same attributes, operations, methods, relationships, and behaviour (ISO 19107). For instance, the “Well” class contains different attributes common to all the wells, such as their name, geometry, location, owner and responsible party for exploitation. The geospatial object is, for example one particular, identifiable well, located in the field, and called “well_1” as an instance of the “Well” class. Inheritance (generalization-specialization relationship), one of the most important object-oriented paradigms specifies that each super-class (more general class) in the inheritance relationship delegates all its attributes, methods, and constraints to a child-class (more specific class). Abstraction allows for simplifications by modelling and showing classes appropriate to the considered problem. The analyst can work at the most appropriate level of inheritance for a given aspect. Finally, polymorphism is a characteristic of being able to assign different meaning or usage to an object in different contexts, to have more than one form (ISO/TDS 19139). More exhaustive theoretical considerations, together with technical definitions of object-oriented concepts, with applications in the fields of geology and hydrogeology, can be found in Michalak and Leśniak (2003), Michalak (2003a and 2003b), Booch *et al.* (2002), Larman (2001), Carlson (2001), Graham (2001), Page-Jones (1999), Subieta (1999 and 1998).

A formalised language or notation must be used in order to develop object-oriented conceptual models of hydrogeological information and then to describe their structure from different points of view and at different stages of development, from requirements to implementation. Currently, the UML (Unified Modeling Language) notation has been used in many different fields from the description of business processes to environmental issues such as hydrology or hydrogeology (Muller, 2000; Quatrani, 2002). As conceptual modelling in geomatics does not require all methodologies and possibilities of the UML notation, a narrower geomatics profile has been established, consisting in technical specifications

accepted by ISO/TC211, and described in ISO 19103 (2001), with some additional information in ISO 19109, 19118, 19136 (XMML, 2006). Provided that these norms are followed, existing search, analysis and visualisation tools can be reused. Geographic objects encoded following ISO/TC211 and OGC are easily exchangeable by different users, no matter which proprietary or open source software is used.

UML developers want to address different scales of architectural complexity and different possible domains of application. Some of the fundamental advantages of using UML, as a standard conceptual schema language for hydrogeological data modelling, are that:

- both informaticians and hydrogeologists can understand the essence of the data model and its implementation (Vogt, 2002);
- it is possible to follow normative documents of the ISO 19100 series, together with standards issued by OGC, which require the use of the UML notation and provide methodologies for application schema development;
- the standards developed for other domains such as geography, hydrology, or geology can be extended or specialized to meet the needs of the hydrogeological domain, under the conditions that standards overlapping is avoided (Figure 2);
- previously developed and standardized tools for spatial data queries, data analysis, or data transfer can be reused, with no additional documentation;
- interoperable hydrogeological data exchanges between project actors will be possible using different web services for data search and delivery.

5.3 Standards in geospatial information

5.3.1 Why to standardize hydrogeological data models

In order to achieve such an easy hydrogeological information exchange, several measures have to be taken:

- a data exchange standard covering the whole hydrogeological domain has to be developed by the hydrogeological community,
- existing geoinformation and technical standards should be used,
- concepts' overlapping with other related domains have to be avoided,
- the most recent techniques, methodologies and solutions from informatics should be applied at the development, implementation and maintenance levels.

A standard is a normative document, a technical or programmatic solution developed according to consensus procedures, which has been approved by normalization institutions or accepted informally due to a very wide use (Płoski, 1999; ISO: http://www.iso.org/iso/standards_development). As the hydrogeologic domain presents its own specificities, the domain specialists have to participate in the development process. Furthermore, to ensure a very extensive use of such an exchange standard, the latter have to cover the widest possible range of hydrogeological concepts, definitions, uses and implementations.

As far as geoinformation standards are concerned, hydrogeological data should be stored using standard protocols, data formats, and clearly organized data models. The data organization must be explicit, described using standard notations such as Entity-Relationship

diagrams or Unified Modelling Language methodology, which allows for mapping between models and specialized web services.

As geomatics concepts and solutions evolve rapidly, only the newest methodologies should be used to create local, particular models, to establish hydrogeological data exchange standard and to map models. This implies the use of standardized notations and object-oriented principles at development, implementation and maintenance stages.

5.3.2 Standardization institutions

In 1994, two independent and international standardization organisations were established, to bridge the gap in geoinformation standards: the Open GIS Consortium in USA (renamed to Open Geospatial Consortium in 2004), and the ISO Technical Committee 211 in Norway (Ostensen, 1995). The Open Geospatial Consortium, Inc (OGC) is an international industry consortium of 350 companies, government agencies and universities participating in a consensus process to develop publicly available interface specifications (OGC: <http://www.opengeospatial.org/ogc>). OpenGIS Specifications support interoperable solutions that "geo-enable" the Web, wireless and location-based services. The specifications empower technology developers to make spatial information and services accessible and useful to all kinds of applications. The ISO/TC211 Geographic information/Geomatics scope is focused on standardization in the field of digital geographic information (ISO/TC211: <http://www.isotc211.org>). It aims at establishing a structured set of standards for information on objects or phenomena directly or indirectly associated with a location relative to the Earth. According to the ISO/TC211 statement, geographic information standards may specify methods, tools, and services for data definition, description and management, data acquisition, processing, analysis, accessing, and visualisation. Furthermore, ISO/TC211 standards concern data and information transfer protocols in digital/electronic form between different users,

systems and locations. They provide also a general framework for the development of domain- and sector-specific applications that use geospatial data. ISO/TC211 and OGC work very closely together with other actively engaged international professional bodies (FIG: International Federation of Surveyors, or ICA: International cartographic Association), UN agencies, and specific domain bodies (DGIWG for defence organization, ICAO for International Civil Aviation Organization).

For the hydrogeological domain the standards issued from the above-enumerated institutions should be clearly taken into account and then directly implemented to a future hydrogeological spatial data and services infrastructure.

5.3.3 Geospatial metadata

A metadata record is a file of information in different forms, usually presented as an XML document, which provides basic characteristics of a data or information resource. It provides the: “who, what, when, where, why and how of the resource”. Geospatial metadata can be used to document geoinformation resources in different formats, such as GIS files, or geospatial databases (FGDC, 2006, <http://www.fgdc.gov/metadata>). The ISO 19115 standard states that metadata give information about the identification, the extent, the quality, the spatial and temporal schema, spatial reference, and distribution of digital geographic data (ISO 19115, 2003). ISO standards apply rather to digital data, but their principles can be extended to many other forms of geographic data such as maps, charts, and textual documents, as well as non-geographic data. Metadata can also provide information about the up-to-datedness of data, the used standards, copyrights and ownership (Gaździcki, 2001). The availability of data can also be documented, together with the rules on how they can be accessed and exchanged (Batcheller, 2007, in press). Finally, metadata can describe the content of a database: its schema, users, types of data, description of semantic of data, distribution and physical

organization of data, their amount and statistics concerning data, and others (Subieta, 1999). Definition and presence of metadata, describing the existence, localisation, format, structure, and constraints of geoinformation allow for using data outside the system where the data have been created or stored. The data can be successfully shared, stored and used (Babaie and Babaei, 2005). Metadata help in the coordination of data acquisition; they inform about large datasets, reduce redundant storage, and clarify search results (Batcheller, 2007, in press).

This important metadata issue is particularly valid within an inter-country hydrogeological data sharing. First, hydrogeological concepts should be clearly described to unify their understanding while transferring. Second, hydrogeological properties should be well defined and their units of measure must be included. Finally, all hydrogeological spatial data should be accompanied by information on their coordinate system or its transformation.

Several related standards and technologies concerning metadata are continuously being developed. The ISO 19115 standard (2003) specifies metadata and ISO 19139 (2004) provides their XML implementation, producing an XML-compatible description for geographic information. Finally, the Open Geospatial Consortium Catalogue Services-Web Profile (CSW) uses Web Services technologies to manage geographic metadata (Wei et al., 2007). The emergence of XML and Web Services technologies supports the distribution and transfer of geospatial information across Internet. Nowadays, there are many free and open source tools, as well as commercial software products implementing some of these standards. They support metadata search, viewing, editing, creation, and serving (catalogue services). The examples of such tools are: tkme, MetaScribe or MERMaId for freeware/shareware tools; ArcCatalog, ArcIMS Metadata server, GeoMedia Catalog, and SMMS for commercial tools.

5.3.4 Hydrogeological and geological domains standards

To create a standard for hydrogeological information transfer compliant with ISO/TC211 and OGC principles, existing geography and geology conceptual models should first be imported. Basic hydrogeological features can be described by generic information types, describing their position, geometry and some other more specific attributes. For this purpose, the following standards and markup languages can be reused for the development of the hydrogeological data exchange standard: Geography Markup Language (GML), eXploration and Mining Markup Language (XMML) and Geoscience Markup Language (GeoSciML). GML (Cox et al., 2002), is an XML grammar written in XML Schema which provides a large variety of objects for describing features, co-ordinate reference systems, geometry, topology, time, units of measure and generalized values. The ISO 19136 standard describes GML and it is intended to be used as a basis on the top of which more specific application schemas can be constructed, such as: XMML and GeoSciML. XMML (Cox, 2004) has been developed to support online data transfer for the exploration and mining industry, by 3D Visualisation and Geological Modeling in CSIRO (Commonwealth Scientific and Industrial Research Organization) Australian organisation. GeoSciML (Sen and Duffy, 2005) has been built to exchange geoscientific information.

However, none of these standards conformant projects treats about hydrogeological information, which requires specific geoinformation types, presented in the previous section. As far as the hydrogeological domain is concerned, several identified projects focusing on hydrogeological information transfer standard are presented in Table 2.

<Table 2 here>

One of the most interesting and important for the hydrogeological domain would be GroundWater Markup Language (GWML), a GML Application Schema. It is currently the

only ongoing project concerning hydrogeological information transfer standard, completely compliant with ISO/TC211 and OGC norms. Figure 7 shows the position of GWML in the current landscape of GML and its application schemas developed specifically for different domains. GWML imports different concepts, definition and solutions from “upper” conceptual models, starting from geography, through exploration and mining industry standards to geoscientific information standards.

<Figure 7 here>

6 CONCLUSIONS

Hydrogeological information management is very important for efficient integrated water resource administration, protection and exploitation. Since hydrogeological data are expensive, complex, spatially and temporally distributed, a clear structuring and transparent storage are necessary. Decision makers and interested users should easily access groundwater information coming from multiple sources.

Currently, traditional paper storage of hydrogeological data and information is progressively replaced by electronic carriers. This trend is accompanied by the growing use of computational methods to carry out geoscientific tasks. Real-time data delivery and analysis should be combined with automatic data transfer between groundwater actors, existing databases, monitoring networks and remote sensors. This information flux is needed at diverse levels: from the local level, through regional and national levels, to the international environment.

The main identified problem for easy, time- and effort-efficient transfer of information is the current diversity in hydrogeological data storage and formats. There are many database structures, suited for particular functions, needs and priorities, which make hydrogeological resources management complicated.

In order to overcome this problem, the hydrogeological community must undertake standardization efforts for hydrogeological information storage and transfer. Such a hydrogeological standard will improve data availability and exchange, as well as it will reduce misinterpretations by users who read and display hydrogeological data. Norms and

standards coming from ISO/TC211 and OGC should be taken into account. A new hydrogeological standard should be then developed as a GML application schema, enabling data access through web-based services such as, for instance, WFS or WMS. It will be also possible to combine hydrogeological information with other related domains such as geology, geography, and hydrology. Presently, one of the most important examples being under development is GroundWater Markup Language as a specialization of GeoScientific Markup Language.

7 REFERENCES

- Babaie, H. A. and Babaei, A. (2005). Developing the earthquake markup language and database with UML and XML schema. *Computers & Geosciences* **31**, 1175-1200.
- Batcheller, J.K., (2007). Automating geospatial metadata generation – An integrated data management and documentation approach. *Computers & Geosciences in press*;
- Bernard, E.A., Steward, D.R., Le Grand, Ph. (2005). A Geodatabase for Groundwater Modeling in MLAEM and MODFLOW. ESRI International User Conference 2005, Paper 1633.
- Boisvert, E., Brodaric, B. (2007). GroundWater Markup Language (GWML): Extending GeoSciML for Groundwater. American Geophysical Union, Fall Meeting 2007, abstract #IN53C-03.
- Boisvert, E., Brodeur, J., Brodaric, B. (2005). NADM-H2O and H2O-GML: Enabling Decision Support by Extending NADM for Groundwater Information Interoperability, Digital Mapping Techniques '05 – Workshop Proceedings, <http://pubs.usgs.gov/of/2005/1428/boisvert/index.html>.
- Boisvert, E., Johnson, B.R., Cox, S., Brodaric, B. (2004). GML Encodings of NADM C1, Digital Mapping Techniques '04 – Workshop Proceedings, US Geological Survey Open-file Report 2004-1451, pp. 9.
- Booch, G., Jacobson, I., Rumbaugh, J. (2002). [OMG Unified Modeling Language Specification](#), Version 1.3 First Edition: March 2000. Retrieved on 12 August 2008.
- Brodaric, B., Gahegan, M., Harrap, R. (2004). The art and science of mapping: computing geological categories from field data. *Computers & Geosciences* **30**, 719-740.
- Cabalska, J., Felter, A., Hordejuk, M., Mikołajczyk, A. (2005). The Polish Hydrogeological Survey Database Integrator-a new GIS tool for the hydrogeological database management useful in mapping process. *Przegląd Geologiczny*, vol. **53**, nr 10/2, 917-920.
- Carleton, C., Dahlgren, R.A., Tate, K.W. (2005). A relational database for the monitoring and analysis of watershed hydrologic functions: I. Database design and pertinent queries. *Computers & Geosciences* **31**, 393-402.
- Carlson, D. (2001). Modelling XML Applications with UML: Practical e-Business Applications. Addison-Wesley, Boston.

- Carrera-Hernández, J.J., Gaskin, S.J. (2008). The Basin of Mexico Hydrogeological Database (BMHDB): Implementation, queries and interactions with open source software. *Environmental modelling & Software* **23**, 1271-1279.
- Cox, S., (2004). XMML - a standards conformant XML language for transfer of exploration data, ASEG 17th Geophysical Conference and Exhibition, Sydney.
- Cox, S., Daisey, P., Lake, R., Portele, C., Whiteside, A. (2002). Geography Markup Language Version 3.0 OGC Document Number 02-023r4, <http://www.opengeospatial.org/standards/gml>.
- Cox, S.J.D. (2001) Geologic Data Transfer Using XML, www.digitalearth.net.cn/GISConference/Geologic%20data%20transfer%20using%20xml.pdf.
- Cox, S., Daisey, P., Lake, R., Portele, C., Whiteside, A. (2002). Geography Markup Language Version 3.0 OGC Document Number 02-023r4, <http://www.opengeospatial.org/standards/gml>.
- Duckham, M., Mason, K., Stell, J., Worboys, M. (2001). A formal approach to imperfection in geographic information. *Computers, Environment and Urban Systems* **25**, 89-103.
- Fert, M., Mordzonek G., Węglarz D. (2005). The management and data distribution system of the Hydrogeological Map of Poland 1: 50.000. *Przegląd Geologiczny*, vol. **53**, nr 10/2, 940-941.
- Fowler, M., (1998). Analysis Patterns: reusable object models, Addison Wesley Longman, Menlo Park, CA.
- Gazdzicki, J. (2001). Leksykon geomatyczny – Lexicon of geomatics. Wydawnictwo Polskiego Towarzystwa Informatyki Przestrzennej, Warszawa.
- Gogu, R. C., Carabin G., Hallet V., Peters V., Dassargues A. (2001). GIS based hydrogeological databases and groundwater modelling. *Hydrogeology Journal* 9 (6): 555-569.
- Graham, I. (2001). Object-Oriented Methods – Principles and Practice. 3rd Edition. Addison-Wesley, Boston.
- Hamerlinck, J.D., Wrazien, D.R., Needham, S. (1993). Underground injection well database development for groundwater vulnerability assessment applications. Proceedings GIS/LIS 93, Volume 1, Minneapolis.
- Hughes, D.A., Forsyth, D.A. (2006). A generic database and spatial interface for the application of hydrological and water resource models. *Computers & Geosciences* **32**, 1389-1402.

- IBM, (2005). <http://publib.boulder.ibm.com/infocenter/idshelp/v10/index.jsp?topic=/com.ibm.ddi.doc/ddi.htm>
- IBM, (2003). <http://www-306.ibm.com/software/data/informix/pubs/library/dbdk/sqlt/014.html>.
- ISO, (2007). ISO 19139 (CD) – Geographic information – Metadata – XML schema implementation. ISO/TC211 Document. NSF, Oslo.
- ISO, (2003). ISO 19136 (CD) – Geographic information – Geography Markup Language. ISO/TC211 Document. NSF, Oslo.
- ISO, (2003). ISO 19115 (DIS) – Geographic information – Metadata. ISO/TC211 Document. NSF, Oslo.
- ISO, (2003). ISO 19107 (DIS) – Geographic information – Spatial Schema. ISO/TC211 Document. NTS, Oslo.
- ISO, (2002). ISO 19113 (DIS) – Geographic information – Quality Principles. ISO/TC211 Document. NTS, Oslo.
- ISO, (2002). ISO 19104 (DIS) – Geographical information – Terminology. ISO/TC211 Document. NTS, Oslo.
- ISO, (2002). ISO 19109 (DIS) - Geographic information – Rules for application schema, ISO/TC211 Document, NTS, Oslo.
- ISO, (2001). ISO 19110 (DIS), Geographic information – Methodology for feature cataloguing. NTS, Oslo.
- ISO, (2001), ISO 19103, (DTS) – Geographic information – Conceptual schema language. NTS, Oslo.
- Kristensen, P. (2004). The DPSIR Framework. Workshop on a comprehensive / detailed assessment of the vulnerability of water resources to environmental changes in Africa using river basin approach, UNEP Headquarters, Nairobi, Kenya.
- Lake, R. (2005). The application of geography markup language (GML) to the geological sciences. *Computers & Geosciences* **31**, 1081-1094.
- Larman, C. (2001). Applying UML and Patterns: An Introduction to Object-Oriented Analysis and Design and the Unified Process (2nd Edition). Prentice Hall, New Jersey.
- Maidment, D. R. (2002). Arc Hydro: GIS for Water Resources. ESRI Press, Redlands, California.
- Meeks, W. L., Dasgupta, S. (2004). Geospatial information utility: an estimation of the relevance of geospatial information users. *Decision Support Systems* **38**, 47-63.
- Michalak, J. (2005). Topological conceptual model of geological relative time scale for geoinformation systems. *Computers & Geosciences* **31**, 865-876.

- Michalak, J. (2003a). *Conceptual models of hydrogeological geospatial data – methodological foundation*. Warszawa: Państwowy Instytut Geologiczny, (in Polish).
- Michalak, J. (2003b). Geomatics in hydrogeology. *Geological Quarterly*, **47**(1), 69-76, Warszawa: Państwowy Instytut Geologiczny.
- Michalak, J., Leśniak, P. (2003). Features and coverages in hydrogeological information, *Acta Geologica Polonica*. Vol. **53**, No.3, 247-255.
- Muller, R.J. (2000). Bazy danych – język UML w modelowaniu danych. Wydawnictwo MIKOM, Warszawa, (in Polish).
- Mysiak, J., Giupponi, C., Rosato, P. (2005). Towards the development of a decision support system for water resource management. *Environmental Modelling & Software* **20**, 203-214.
- National Groundwater Committee Working Group on National Groundwater Data Standards, (1999). The Australian National Groundwater Data Transfer Standard.
- Nowicki, B., Staniszkis, W. (2002). Inteligentne systemy zarządzania wiedzą – prezentacja projektu. In: Mat. Konferencja eDemocracy, VI Konf. Miasta w Internecie, Zakopane, Poland, (in Polish).
- Open Geospatial Consortium Inc. (2007). Observations and Measurements – Part 1 – Observation schema, OGC 07-022r1, <http://www.opengeospatial.org/standards/o%2526m>
- Open Geospatial Consortium Inc. (2007). OpenGIS Sensor Model Language (SensorML), OGC 07-000, <http://www.opengeospatial.org/standards/sensorml>
- Open GIS Consortium Inc., (2003). Observations and Measurements, OGC 03-022r3, http://portal.opengeospatial.org/files/?artifact_id=1324
- Page-Jones, M. (1999). *Fundamentals of Object-Oriented Design in UML*. Addison-Wesley, Boston.
- Płoski, Z. (1999). *Słownik Encyklopedyczny – Informatyka*. Wydawnictwo Europa, Warszawa, (in Polish).
- Pokorný, J. (2006). Database architectures: Current trends and their relationships to environmental data management. *Environmental Modelling & Software* **21**, 1579-1586.
- Power, D. J. (1997). What is DSS? DSstar. *The On-Line Executive Journal of Data-Intensive Decision Support* 3.

- ProGEA, S.r.l. (2004). A WaterStrategyMan DSS, A comprehensive Decision Support System for the development of sustainable water management strategies, Bologna, Italy;
- Quatrani, T. (2002). Visual Modeling with Rational Rose 2002 and UML. Addison-Wesely, Boston.
- Rodríguez J., López, J., Gómez, M., de Mera, A., Pérez, F., Hernández, R., Iglesias, A. (2007). Inventory of hydrogeological maps and models available in partners countries. eWater Project, GAN, ECP-2005-GEO-038214.
- Sen, M., Duffy, T. (2005). GeoSciML: Development of a generic GeoScience Markup Language. *Computers & Geosciences* **31**, 1095-1103.
- Simons, B., Duffy, T., Boisvert, E., Johnson, B., Brodaric, B., Laxton, J., Cox, S., Richard, S. (2006). GeoSciML: Enabling the Exchange of Geological Map Data, *AESC*, Melbourne, Australia.
- Smith, B., Mark, D.M. (1998). Ontology and Geographic Kinds, Proceedings, International Symposium on Spatial Data Handling (SDH'98), Vancouver, Canada, 12-15 July, pp. 308-320.
- Strassberg, G. (2005). A geographic data model for groundwater systems, Ph.D. Thesis, University of Texas, Austin.
- Subieta, K. (1999). Obiektywność w projektowaniu i bazach danych. Akademicka Oficyna Wydawnicza, PJJ, Warszawa, (*in Polish*).
- Subieta, K. (1998). Słownik terminów z zakresu obiektywności. Akademicka Oficyna Wydawnicza, PJJ, Warszawa, (*in Polish*).
- Szalkai, A.R., Nagy, P., Tullner, T., et al. (2007). Organisation of ground water data management in the participating countries. EC FP6 eWater project Deliverable 3.1.
- Tindall, C.I., Moore, R.V., Bosley, J.D., Swetnam, R.D., Bowie, R., De Rudder, A. (2006). Creating and using the urgent metadata catalogue and thesaurus. *Science of the Total Environment* **360**, 223-232.
- Vogt, J. (2002). Guidance Document on Implementing the GIS Elements of the Water Framework Directive.
- XMML (2006). XMML web site: Solid Earth and Environment GRID, CSIRO, July 2006, <https://www.seegrid.csiro.au/twiki/bin/view/Xmml/WebHome>.
- Wei Y., Di, L., Zhao, B., Liao G., Chen, A. (2007). Transformation of HDF-EOS metadata from the ECS model to ISO 19115-based XML. *Computers & Geosciences* **33**, 238-247.

- Whiteside, A. (1999). UML Profile and Guidelines for OGC Abstract Models, OpenGIS Project Document 99-031, Wayland.
- Worboys, M.F. (1998). Imprecision in finite resolution spatial data. *GeoInformatica* **2**, 257-279.
- Wojda, P., (2009). Hydrogeological data modelling in groundwater studies. Ph.D. thesis. University of Liège, Faculty of Applied Sciences, Belgium.
- Wojda, P., Gardin, N., Gogu, R., Brouyère, S., (2006). Report on computerized database system. Deliverable D63 of the EU FP6 (GOCE) Project Nb. 518118-1 GABARDINE.

TABLES AND CAPTIONS

Table 1. Examples of the data models that differ in their design and notation. They were developed to respond to particular needs, specific application and functionality.

Projet/Model Name	References	Description	Original Input
<i>Underground injection well database</i>	Hamerlinck, Wrazien, Needham, 1993	<p>A GIS-based underground injection well database has been developed for the State of Wyoming. The main objective of the project was to determine geographic locations for 6700 injection wells to help in assessing their potential as point sources of groundwater contamination.</p> <p>The structure of this GIS database is based on Arc/INFO georelational vector data structure. For spatial data, represented by point, line and polygon geometries, an arc-node structure is prepared, while attribute data, describing spatial features are stored in a relational database.</p> <p>Project link: http://library.wrds.uwyo.edu/wrp/93-08/93-08.html</p>	GIS-based database structure, suiting for vulnerability assessment
<i>The Australian National Groundwater Data Transfer Standard</i>	NGC Groundwater Data Standards Working Group in the National Groundwater Committee, 1999	<p>This standardized hydrogeological data model was developed in Australia in order to unify different existing data models. The diversity in which groundwater data were stored and transferred was unnecessarily complicating natural groundwater resource management.</p> <p>The new hydrogeological model reduces the time required to reformat data, boosting significantly their productivity. It helps to overcome trans-boundary groundwater problems. Misinterpretations by users are reduced, together with confusion reading and displaying hydrogeological data.</p> <p>Project link: http://www.brs.gov.au/land&water/groundwater</p>	Standard Groundwater Data Transfer Model at the level of the whole country
<i>A geographic data model for groundwater systems</i>	Maidment, 2002; Strassberg, 2005; Bernard, et al., 2005	<p>The primary objective was to design a groundwater data model for describing, storing, visualizing, analyzing and communicating groundwater geospatial information at regional and local scales, combining surface and groundwater information. It takes advantage of the already developed ArcHydro surface water data model (Maidment, 2002).</p> <p>The model incorporates four major components, namely: “hydrogeological” features represented by points, lines polygons and multi-patches elements; “modelling” entity representing common modelling objects such as cells and elements; “surfaces” represented by rasters and TINs, used to define elevation or spatially distributed aquifer parameters; “times series” used to represent time dependent information. This data model has been extended by the Groundwater-AEM data model (Bernard et al., 2005) in order to allow using MLAEM (Multi-Layer Analytic Element Model) for groundwater flow numerical modelling.</p> <p>Project link: http://www.crrw.utexas.edu/gis/gishydro05/ArcHydroGroundwater/ArcHydroGroundwaterESRIUC2005.htm</p>	Compatibility of the model with the widely applied ArcHydro model. Possibility of groundwater flow modelling using directly the implemented data model
<i>A relational database for the monitoring and analyses of watershed hydrologic functions</i>	Carleton et al., 2005	<p>The Watershed Monitoring and Analysis Database is a relational application developed to manage hydrologic datasets. It stores and allows for manipulation of stream flow, water quality, and meteorological data. It has additional tools to assure quality of data and analyses, to correct conversion factors or finally to retrieve required data for analyses.</p> <p>The Database supports web integration and Local Area Network work, depending on the implementation platform. The on-line synchronisation can be performed.</p> <p>Project link: http://cat.inist.fr/?aModele=afficheN&cpsidt=16659730</p>	Optimized storage requirements and retrieval rates, easy web integration, data replication within LAN

Projet/Model Name	References	Description	Original Input
<i>A generic database for the application of hydrological and water resource models</i>	Hughes, Forsyth, 2006	<p>The SPATSIM (SPatial and Time Series Information Modeling) system has been developed using MapObjects. It incorporates a spatial data interface for access to the different types of information used in water resources analyses. All the information is stored within database tables with generic structures.</p> <p>The database where spatial elements (point, line and polygon types) are stored as shapefiles in the Paradox format has four dictionaries, which allows for easy data manipulation and errors limitation. It allows for storage and access of the information typically associated with water resource studies. The SPATSIM system contains several tools for the data exchange to facilitate movements of attribute data between different users.</p> <p>Project link: http://portal.acm.org/citation.cfm?id=1296607</p>	Access to data for hydrological and water resource simulations models
<i>HydroCube database</i>	Wojda et al., 2007 manuscript in preparation	<p>The HydroCube database has been developed and implemented in the Walloon region of Belgium (Wojda et al., 2007 C&G). It is partially based on HYGES database model developed previously by Gogu <i>et al.</i> (2001) in order to manage hydrogeological data, particularly in the scope of groundwater vulnerability assessment and modelling.</p> <p>The HydroCube database is based on a new formalized logical model of hydrogeological data, described by entity-relationship diagrams, and enriched with fully functional user interfaces. It enables to deal with a hydrogeological project as a whole, by managing the data about the project localisation, available hydrogeological studies, and contact people. Furthermore, necessary groundwater features, monitoring results, performed field tests descriptions and interpreted results can be stored.</p> <p>Project link: http://www.argenco.ulg.ac.be/GEO3_Hydrogeologie/banquedonnees_fr.html</p>	Completeness of hydrogeological data model, including a model for specialised hydrogeological field experiments such as pumping tests and tracer tests
<i>The basin of Mexico Hydrogeological Database</i>	Carrera-Hernández and Gaskin, 2008	<p>To manage efficiently regional water resources at the basin level, the use of both Relational Database Management System and a Geographic Information System is proposed. The Basin of Mexico Hydrogeological Database comprises data on climatological, borehole and run-off variables, providing information for the development of hydrogeological models. It allows also for geostatistical analyses using data directly from BMHDB.</p> <p>Hydrogeological data can be accessed and processed locally or remotely through open source software: postgresQL, R and GIS GRASS packages.</p> <p>Project link: http://portal.acm.org/citation.cfm?id=1379720</p>	Use of open source products, data gathering from different sources, Easiness of geostatistical analysis

Table 1. Examples of the data models that differ in their design and notation. They were developed to respond to particular needs, specific application and functionality

Table 2. Unifying hydrogeological data models and enabling information transfer.

Projct/Model Name	References	Description	Original Input
<i>The Polish Hydrogeological Survey Database Integrator</i>	Cabalska et al., 2005	<p>The PHS Database Integrator, as a practical and applied tool, has answered to the needs for a sophisticated instrument for groundwater management (Cabalska et al., 2005). This project aimed at hydrogeological data integration, and at data gathering from multiple and heterogeneous available data sources. It integrates all hydrogeological databases existing in the Polish Geological Institute, such as Groundwater Monitoring database, HYDRO Bank and Hydrogeological Map of Poland in 1:50.000 scale, the latter being developed using an Intergraph technology (Fert et al., 2005). This solution allows for an effective hydrogeological data integration, retrieval, and analysis. It reduces also the time needed for data collection and data redundancy by its transparency.</p> <p>Project link: www.pgi.gov.pl/pdf/pg_2005_10_2_10.pdf</p>	Specialized tool integrating hydrogeological information from multiple sources.
<i>GABARDINE Geospatial hydrogeological database</i>	Wojda et al., 2006	<p>In order to use the advantages of object-oriented modelling, to follow the international standards for transfer of geospatial information (ISO/TC211 and OGC), and to be compliant with the recommendations from the European Geospatial Information Working Group (Vogt, 2002), a new hydrogeological data model called GABARDINE GDB has been developed. The model has been implemented in the ArcGIS environment, as a database for a Decision Support System for the EC FP6 GABARDINE project (Groundwater Artificial recharge Based on Alternative sources of water: aDvanced Integrated technologies and management) (Wojda et al., 2006).</p> <p>The proposed holistic Project-Oriented approach enables to deal with a hydrogeological project as a whole, by managing the data about the project localisation, available hydrogeological studies in the zone of interest, contact people and contributors. Furthermore, existing hydrogeological equipment, natural and man-made groundwater access features, monitoring results, field tests performed in the zone together with their results and possible interpretations can be gathered, visualized and analyzed.</p> <p>Project link: www.gabardine-fp6.org/</p>	Hydrogeological data model described using the UML notation, following ISO/TC211 and OGC recommendations. Implemented in ArcGIS environment.
<i>eWater project hydrogeological data model</i>	Coordinated by Dr Alexei Tchistiakov	<p>The main objective of the ongoing FP6 EC eWater project coordinated by the TNO Dutch Institute is to increase the cross-border availability, accessibility and re-usability of spatial data on quality, location and use of subsurface waters. In order to achieve this objective, a multilingual WEB GIS portal is under development. The portal will be accessible for the project partners, participating countries, national river basin authorities, and water suppliers. It will give an additional value to data service providers, insurance companies, planning and controlling organizations, as well as general public, making hydrogeological data and information available. The eWater architecture complies with the INSPIRE policy (INSPIRE: http://www.ec-gis.org/inspire), and the data will be usable not only by the suppliers, but they can also be included in Water Information System for Europe (WISE: http://water.europa.eu).</p> <p>Project link: http://ewater.geolba.ac.at/</p>	First European project increasing availability and usability of hydrogeological data, using UML for data modelling and Web-based services for data exchange

Projet/Model Name	References	Description	Original Input
<i>GroundWater Markup Language</i>	Boisvert and Brodaric, 2007;	<p>The GWML project is in its very early stage of development and discussions. Many applied concepts have been inspired from NADM-C1 (Boisvert et al., 2004), and imported or derived from XMML (Cox, 2004) and GeoSciML (Sen and Duffy, 2005), following a standardised GML extension pattern for Application Schemas. The interoperability framework of GWML is based on OGC standards. It incorporates GML-based standards such as Observation & Measurements (07-022r1, 2007), SensorML (07-000, 2007) and GeoSciML.</p> <p>Due to its compliant structure, it will be possible to use it in conjunction with OGC web service standards and protocols such as Web Mapping Services; Web Feature Service; Sensor Observation Service; Web Coverage Service. A specialized collaboration is performed on-line. GWML provides a very good starting point for groundwater data interchange format. Eventually, GWML might be used as the GML Application Schema (GeoSciML derived more precisely) for groundwater information exchange.</p> <p>Project link: http://ngwd-bdnes.cits.rncan.gc.ca/service/ngwd/exploration/ngwd/gwml.html?locale=en&SESSION=PUBLIC&.</p>	GML-derived as its application schema and compliant with standards and norms issued by OGC and ISO/TC211

Table 2. Unifying hydrogeological data models and enabling information transfer.

FIGURES AND CAPTIONS

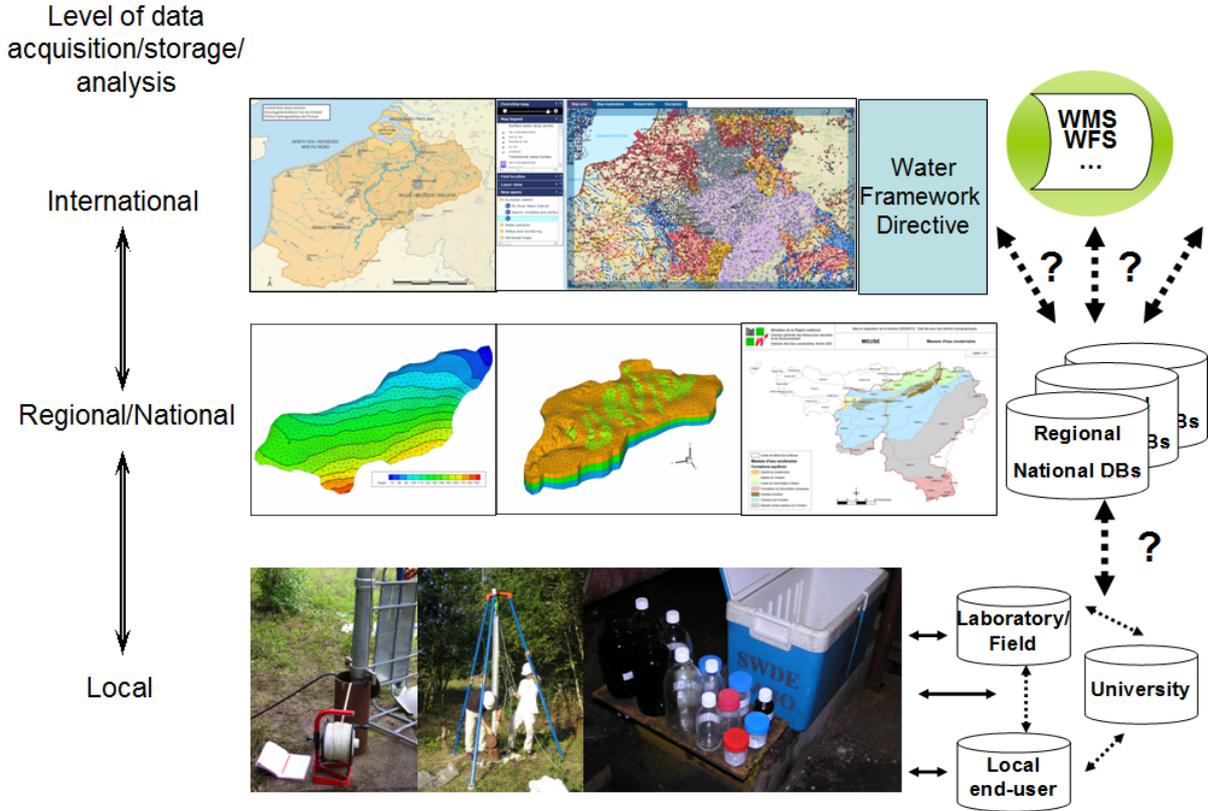


Figure 1. Hydrogeological information required and used at different levels of management, from local to international. As examples, WMS (Web Map Service) and WFS (Web Feature Service) are available services which can deliver geoinformation in a standardized format.

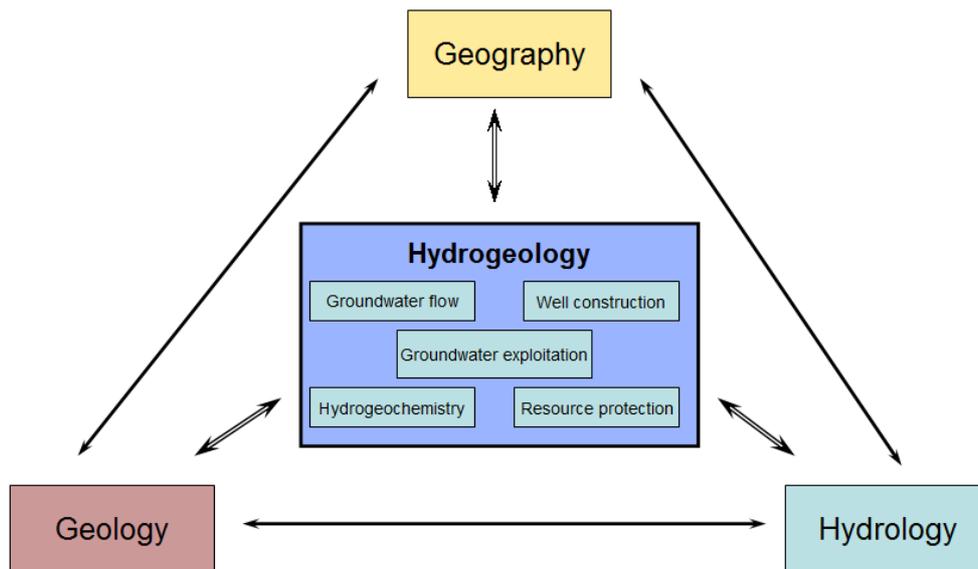


Figure 2. Interactions amongst hydrogeology-related domains: geography for space- and time-dependent problems, geology for porous/fractured media, hydrology for surface water resources. Within the single field of hydrogeology, several particular aspects should also be treated simultaneously and in conjunction with each other.

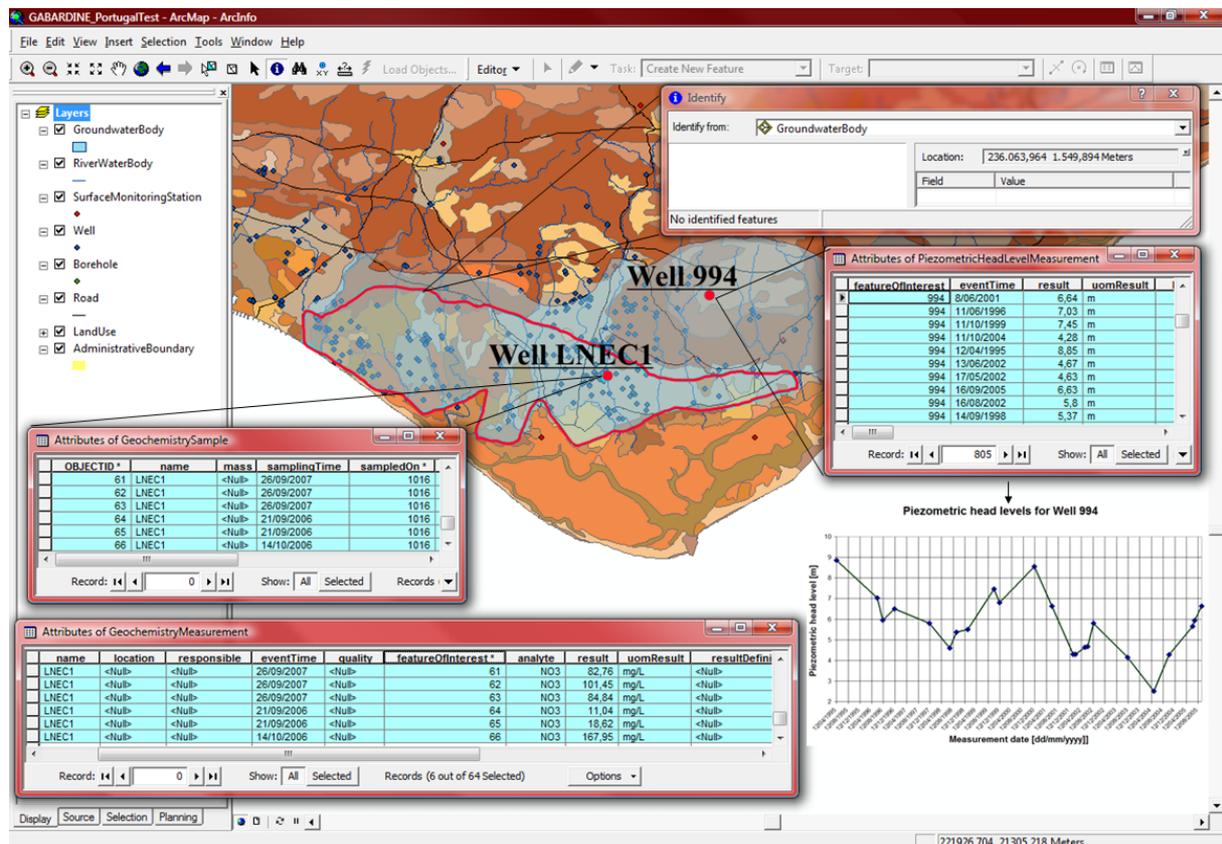


Figure 3. Example of the use of a Geospatial Database, developed for the FP6 EC GABARDINE project. On the left-hand side: for each groundwater sample (“Attributes of GeochemistrySamples” window)

taken from the “LNEC1” well, nitrate analyses are reported (“Attribute of GeochemistryMeasurements” window). On the right-hand side: piezometric head level measurements are reported for the well “994”. Observations and measurements can be exported to another software for further processing.

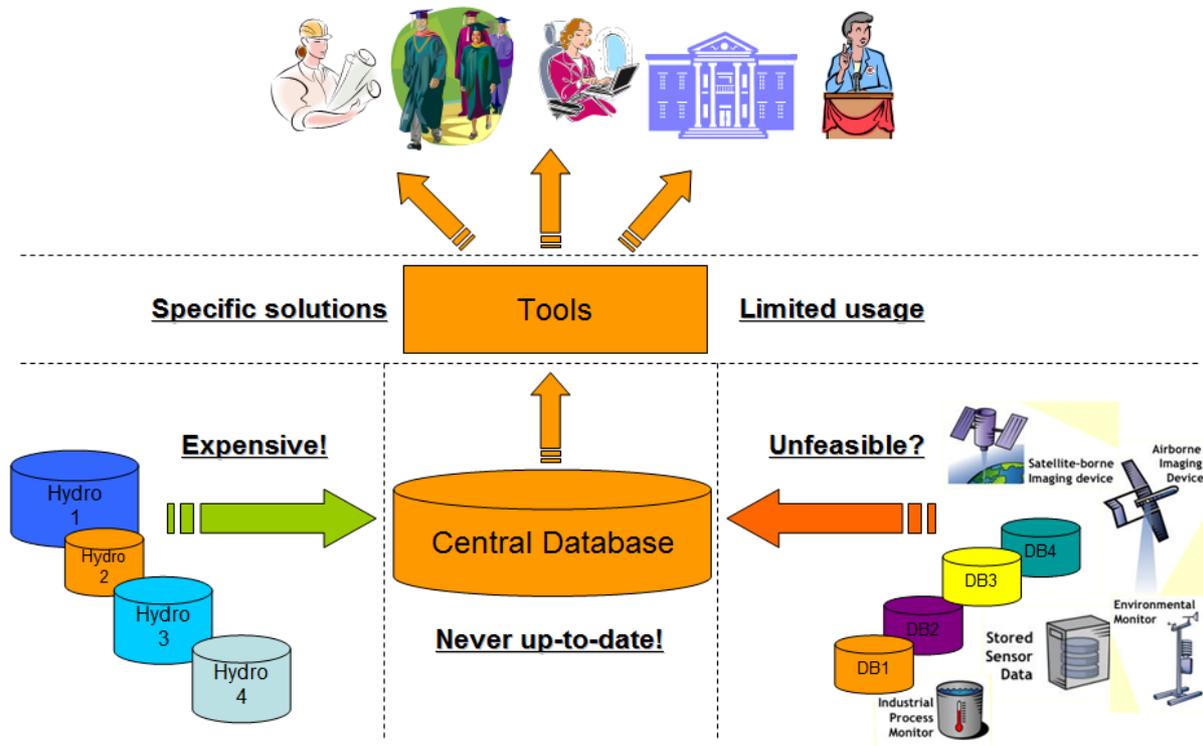


Figure 4. Without any standardization of hydrogeological data and other types of data, it is expensive and difficult to exchange data in an efficient way. A central database solution could be helpful; however it is never up-to-date, expensive to maintain, and poorly enriched with specific tools, which limit its usage for potential users (image source: WRON Australia, adapted).

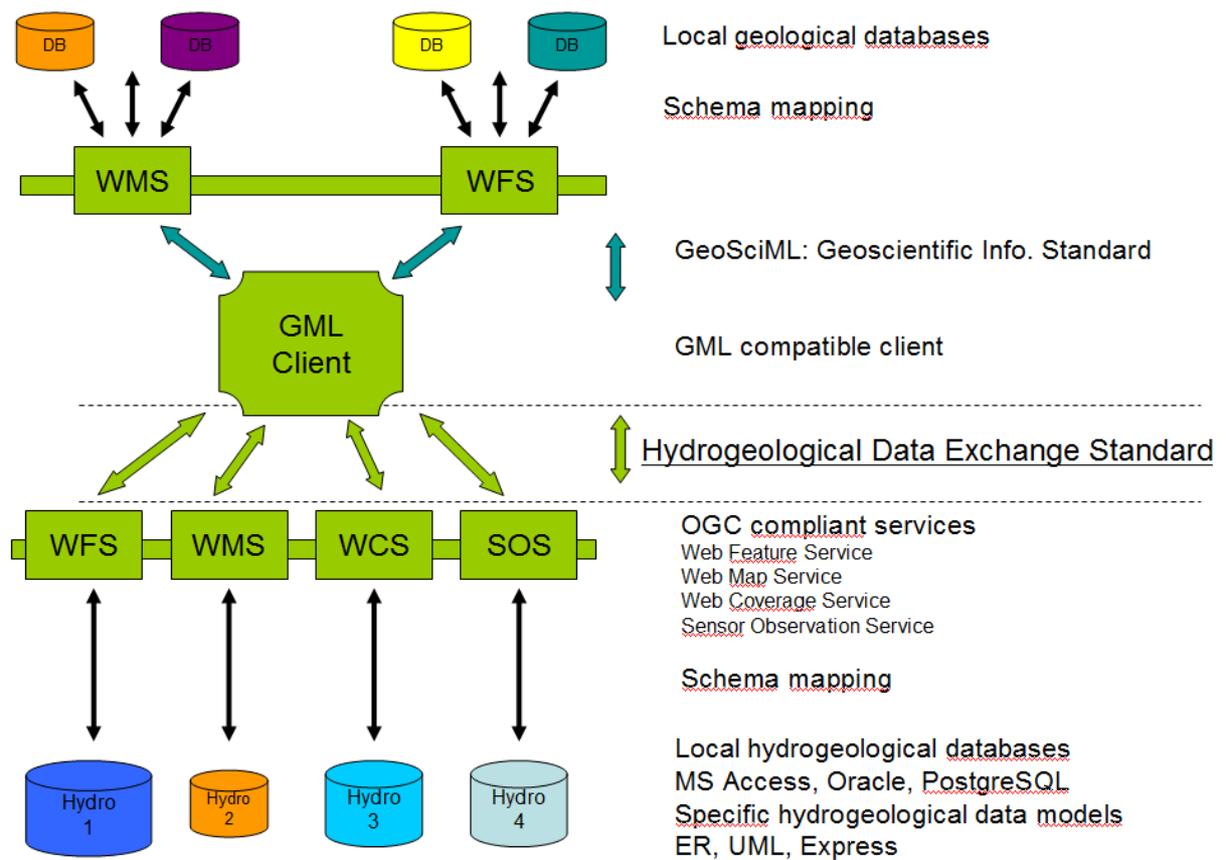


Figure 5. Hydrogeological databases with their specific data models may be mapped to standard web services, such as Web Map Service or Web Feature Service. Hydrogeological information will be then delivered to a GML-compatible client, using a standardized hydrogeological data format. The GML-compatible client is also able to read information from other sources, using other GML application schemas from various domains: geology, hydrology, geography, biology, economics...

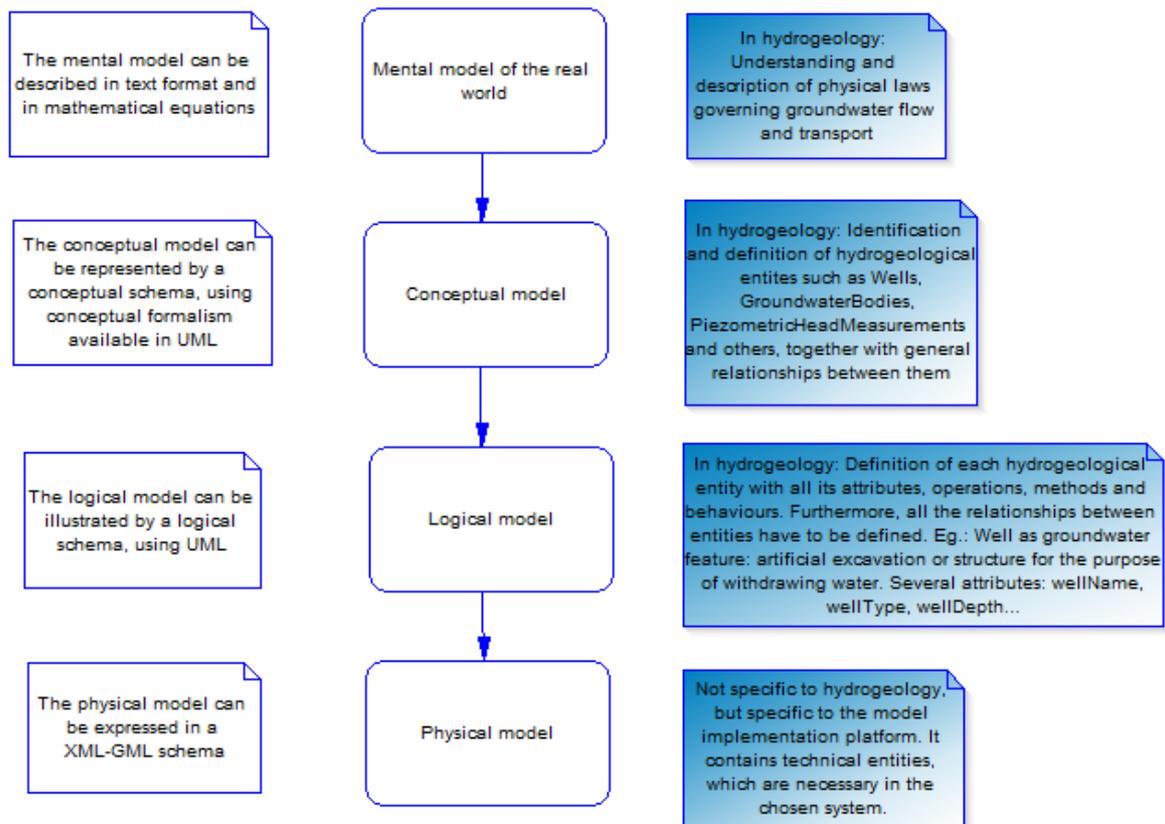


Figure 6. Different steps of data modelling with examples particular to the hydrogeological domain. A hydrogeologist has a mental model of the hydrogeological domain. In order to structure hydrogeological knowledge, a conceptual model has to be defined which consists in identifying hydrogeological entities and relationships amongst them. This can be formalized and detailed using logical models. Finally, the platform-dependent implementation process requires a physical model.

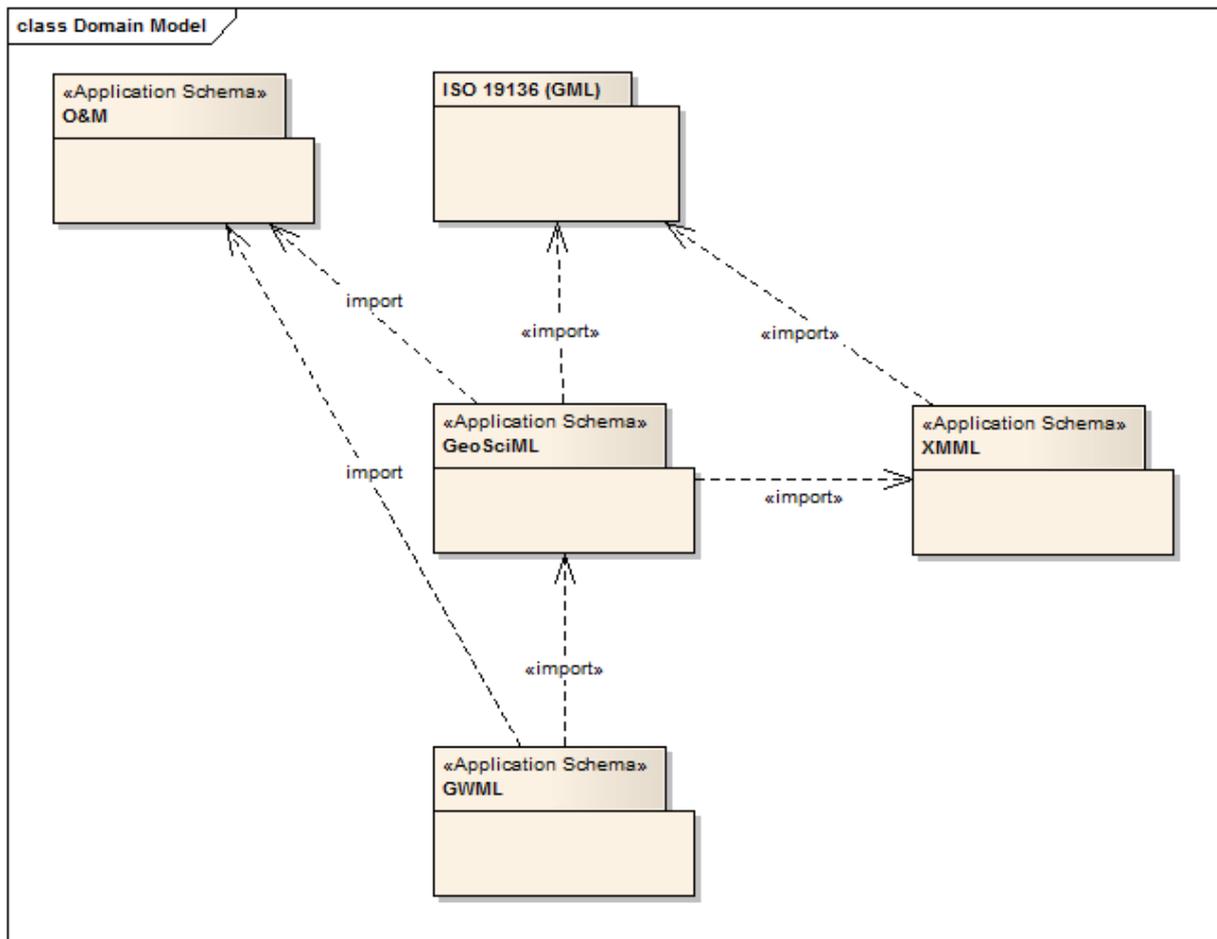


Figure 7. Simplified illustration of dependencies between the ISO 19136 standard and its derived application schemas specific to different domains.