ACQUA: A Conceptual Data Model for Designing and Implementing Databases for Water Resources Management in GIS Environment

Angelo Brayner
University of Fortaleza, Department of Computer Science (MIA), Campus da Unifor, Bloco D,
Fortaleza, Ceará, Brasil, 60811-341,
brayner@unifor.br

and

Joney Rosas Cysne
Federal University of Ceará, Department of Computer Science (MCC), Campus do Pici, Bloco 910,
Fortaleza, Ceará, Brasil, 60455-760,
joney@funceme.br

Abstract

This paper proposes a conceptual data model, called ACQUA, for representing water resources. The proposed model has the capability of capturing the semantics of geographic phenomena, which are related with water resources, such as reservoirs, rivers, channels, pipelines, lakes and singularities (e.g. waterwithdrawal, waterdischarge, confluence, riverhead and monitoring-station). Using ACQUA it is possible to model the complex water resources network, its physical properties and its spatial information, within a geographic context. The model captures and represents information about an area and may be used as support for urban planning, water-related crisis management (such as, water offer and demand during a dry period) or as a basis for scenario simulation.

Keywords: Conceptual Data Models, Spatial Databases, GIS and Geographic Applications.

1. INTRODUCTION

From a database technology standpoint, a Geographic Information Systems (GIS) represents an information system with the capability of capturing the semantic of geographic phenomena, in order to storage them as geographic data in a special type of database, a spatial database. In this sense, it is possible to use database system facilities to manage geographic data.

GIS are used to collect, analyze, and present information of geographic world, describing the physical and logical properties (Shekhar et al. [1]). It may be used to support geographical data from different sources, such as cartographic data, urban and agricultural census data, satellite images, networks and surface numerical models. Viewed as a system composed by hardware, software, collection of spatial and non-spatial data and human resources (Nyerges [2]), a GIS internally presents hierarchic components distributed in layers. These layers are called *(i)* user interface, *(ii)* data input and integration, *(iii)* vector processing and raster functions, *(iv)* viewing and plotting functions, *(v)* storage and accessing data functions.

Geographical world is represented in two perspectives: field model and object model. The field model see the world like a continuous surface, and the geographical phenomena (fields) usually are represented by tessellation or matrix structures. At object model, the world is represented like a surface occupied by identified objects, which can be modeled by three basic structures: points, lines and polygons (Câmara et al. [3]). Geo-referenced data may present three distinct components: a non-spatial component which describes the phenomenon; a spatial component which describes spatial position of the phenomenon, capturing geometric and topological properties, and a time component which represents time properties of the geographic data (Silberschatz et al. [4]).

Over the last decades, water related problem has experienced meaningful changes. Nowadays, water is a scarce natural resource in worldwide scale, and endowed of economic value (Campos and Studart [5], Souza [6]). Of course, managing efficiently water resources has become a challenge. Some researchers have proposed data models for representing geographic phenomena in spatial databases to support an efficient management of environmental resources. By doing this, it is possible to use spatial database functionalities for managing water related data.

According to Borges et al [7], a geographical data model should provide *(i)* a high level of abstraction to represent geo-fields and geo-objects *(ii) a r*epresentation for all kinds of data and spatial relationships involved in geographical applications, *(iii)* definition for spatial integrity constraints, *(iv) s*upport to represent geo-classes or non geo-classes, *(v)* multiple visions for the same object, *(vi) i*ndependence between conceptual data model and physical data model.

In this sense, this paper proposes a model, called ACQUA, for representing water resources. The proposed model provides the necessary support to represent a complex water resources network, capturing hydrological and hydrographic information, and the geometry and topology of the geo-referenced objects belonging to the network. Using the ACQUA model one can represent water-resource related information about a given area allowing: identification of all sections of a given channel, pipeline or river in order to locate where hydraulic structures are installed for irrigation, for example. Furthermore, it is possible to determine the percentage of irrigated areas, to control water offer and demand during a dry period and to register data collect by monitoring stations. It is important to note that the model is able to capture information (spatial and non-spatial semantics) of any other water related geographic phenomena such as lake, well, reservoir, dam, monitoring-station, waterwithdrawal, waterdischarge, etc.

This paper is organized as follows. In the next section, we review some GIS conceptual models using different approaches. In section 3, we motivate the applicability and feasibility of our proposal by describing an example that shows how the ACQUA model organizes information about an area. In section 4, we present the core of the model to explain its coverage. In section 5, we presents a brief report about ACQUA VIEW tool, designed and implemented to manage water resources, based on concepts and principles of the ACQUA model. Section 6 discusses the applicability of our proposal. Finally, section 7 concludes the paper.

2. RELATED WORK

A conceptual spatial data model for representing geographical data poses challenges to database designers. In order to develop conceptual spatial data models, some approaches, methodologies and techniques have been used to support and to understand the complex GIS world. In general, the models are developed following the entity-relationship or object-oriented approach. Data models for representing geographic phenomena in GIS can be categorized into field-based models and object-based models or both. Each GIS data model, in particular way, can be considered an innovate model, according with its extensions, features, perspectives and applications (Erwig et al.[8], Patel et al.[9], Jones and Ware[10], Balaguer and Gordillo[11], Laender et al.[12]).

Many researches have proposed spatial data models based on the object-oriented approach. PDM (Dayal and Manola [13], Manola and Orenstein [14]), OO Spatial Model (Nabil and Subramanian [15]) and conceptual model used by SPRING GIS (Câmara et al. [3]) are some attempts to model GIS using object-oriented approach.

The PDM (Dayal and Manola [13], Manola and Orenstein [14]) - Probe Data Model - provides two basic types, entities and functions, to represent data objects. The model also provides PDM algebra to perform manipulation of PDM databases, including formation of generalization hierarchies, new entities types and multi arguments functions. The PDM allows several representations of a given spatial object, depending on the user perspective (different schemas) and takes the approach of modeling topological relationships such as “adjacent-to” explicitly.

In Nabil and Subramanian [15], the OO Spatial Model is based in a combination of data-driven approach (i.e. concentrate in the attributes of the objects) and responsibility-driven approach (i.e. concentrate in the functions or responsibilities of the objects) enhancing the ideas presented in PDM model. PDM was improved by notion of user perspective related with objects representation and implementation of a class of imprecise spatial operators such as “close-to”, “between” and “adjacent-to”. The base classes of the model are Spatial-feature (which is an abstract class that represents a template for spatial feature) and Line-segment (which is a concrete class that represents low-level spatial objects).

The model proposed for the SPRING GIS (Câmara et al. [3]) represents geo-world in three classes: Non-spatial, Geo-object and Information-plan class. The Information-plan class is a generalization of Geo-field class and Register-class. The Geo-field class and Geo-object class are the roots to Geo-classes basic hierarchy.

On the other hand, some researches have followed the entity-relationship approach where the data are described as entities, relationships and attributes. GISER (Shekhar et al. [1]) and “Roads” data model (Barnett and Carlis [16]) are examples of that approach.

GISER (Shekhar et al. [1]) - Geographic Information System Entity Relational - unifies the field-based and object-based approaches, and explicitly represents the discretization aspects. GISER attempts to support the entire GIS process, from the input of data measured and discretized to the display of this entity and all the processing that must be performed. GISER has also extended entity-relationship model to include continuous fields (EER).

It is important to emphasize that several data models for GIS were developed to represent geo-referenced phenomena by providing the support to associate “meaning” to a spatial object. In the other words, the users have to understand, to refine and to represent the semantics of the geographic phenomena and their relationships. For example, the SPRING GIS model provides the elements to model a spatial object. However, the user has to associate meanings (semantic) to this object. That is, the model provides the support to represent a sequence of lines. None of less the user has to associate a “meaning” to that sequence of lines in order to capture the information that the lines represents a given river, for example.

The “Roads” data model (Barnett and Carlis [16]) is an example of a data model driven for a specific domain. It addresses the problem of cartographic generalization or map generalization in the context of a common map features called “Roads”. Complex crossroads, highways, streets, interstate lines etc, are represented in the model. Moreover, the model includes specific types of data to represent explicitly the rules the will control the desired aesthetics of the objects at digital representation. Route, Partitioning, Section and Route-rule are the base entities of the model. Observe that in this model the spatial objects have already a meaning (e.g. road, route and section).

3. Motivating Example

We motivate the applicability and feasibility of our proposal by describing an example modeled according to the reality of a semi-arid region. The example will analyze an imaginary river, called Mandacaru River, although describing real structures. The idea is to understand a river as sequence of sections, where each section represents a stretch of the river. Each section of Mandacaru River has begin and end points. Each point is treated as a singularity point, which may have spatial and non-spatial attributes. Suppose that the Mandacaru River has the sections presented in Figure 1.

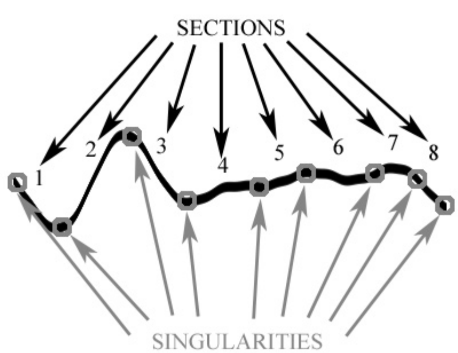

Figure 1 - Mandacaru River Topology.

The identification of each singularity helps to understand and to analyze the water related problems of each stretch of the river. It is interesting to note that a singularity which represents the end point of a section may also represents the begin point of the next section. The Mandacaru River sections (with begin and end points) are described in Table 1.

It is important to observe that a singularity may be a link-point to another section or another water network as a channel, pipeline, reservoir set, well set or another river. For example, the same singularity, called inflect point, represents the end point of the section 1 and the begin point of the section 2. The singularity inflect point may represents a sharp curve, enlargement, reduction, elevation or lowering, in a certain section in the pipe, channel or river. There is also the singularity waterwithdrawal. Waterwithdrawal are important hydraulic structures built to transfer water (e.g. using water pumps) to another hydraulic structure, such as a pipeline or channel. This channel or pipeline conducts the water to others reservoirs, wells, channels and pipelines, which can be water related objects belonging to another water resources network, that is, they do not belong to the water resources network of the Mandacaru River. In general, unfortunately dirty water, without treatment, come back to the rivers through others hydraulic structures, called waterdischarges (or sewers). For this reason, monitoring stations are installed to control water quality. Each river that reaches a principal river represents a confluence. Finally, at the river course, that begun with the singularity riverhead, there is a dam and an island before it reaches the rivermouth.

Table 1 - Mandacaru River Sections.

Begin Point (Singularity)	**Section Number**	**End Point (Singularity)**
Riverhead	1	Inflect Point
Inflect Point	2	WaterWithdrawal
WaterWithdrawal	3	Sewer
Sewer	4	Monitoring Station
Monitoring Station	5	Confluence
Confluence	6	Dam
Dam	7	Island
Island	8	Rivermouth

4. THE ACQUA MODEL

In this section, we describe the ACQUA model using the Unified Modeling Language (UML). UML is a pattern language adopted by OMG (Object Management Group) to describe a conceptual model, in a uniform way. Besides, as UML is an extensible language, it allows introducing new stereotypes for specific applications, if it is needed. Furthermore, the UML formal concepts provide appropriated support for representing the complex objects related with complex relationships that are inherent to the geographic information systems.

The ACQUA Model is focused on superficial water resources and adopts the object modeling, instead of thematic modeling, to represent water world. The proposed model has the following properties: *(i)* clear and simple representation of the objects, *(ii)* capability of capturing the semantics of any real-world object involved in a complex water resources network, *(iii)* independence between conceptual model and physical model, *(iv)* representation of geo-referenced and conventional classes, *(v)* representation of topological relationships.

Four classes compose the core of the proposed model. These classes are called *Natural_Hydrography*, *Water_Reserve*, *Water_Transfe*r and *Singularity* class. Figure 2 depicts the class diagram of the ACQUA model.

The *Natural_Hydrography* class represents water resources sculptured by the nature. This class has two subclasses: *Lake* and *River* class. The *Water_Reserve* class is used to represent hydraulic structures built to storage water for many purposes, such as irrigation, energy supply, human consumption etc. The *Water_Reserve* class has two subclasses: *Well* and *Reservoir* class. The *Water_Transfer* class models hydraulic structures used to transport water from one place to another and it also has two subclasses, called *Channel* and *Pipeline* class. Three distinct subclasses represent sections of rivers, channels and pipelines. Others water-related structures can be modeled through the *Singularity* class.

The *Singularity* class has these subclasses: *WaterDischarge*, *Riverhead*, *Rivermouth*, *WaterWithdrawal*, *Confluence*, *Inflect_Point, Island*, *Monitoring_Station*, *Dam* and *Hydroelectric_Plant* class. These subclasses that represent hydraulic structures or geographic phenomena contribute directly to water system operation, and some of these objects can be relating (linking) the others three classes. For example, the water that emerges from a riverhead (first singularity identified in a river) follows freely until may be repressed by a dam (another type of singularity). In such a case, a great reservoir is created (reservoir class). The water follows its course until it reaches the rivermouth (the last singularity of a river), but before this, others rivers may reach the principal river (confluence singularity) and many waterworks may transfer water from the river to channels or pipelines to supply others places (waterwithdrawal singularity). Note that without singularities the water network topology would be simplified, hiding important details necessary for water resources management.

The *Monitoring_Station* subclass represents control-points of the monitoring process in the water resources network. The inner sensors are used to identify stations type and the exact location where the stations have to be placed. Data input includes precipitation, evaporation, infiltration etc, and are used in conjunction with others measured data for testing alternative monitoring designs and management options. In such case, the *Monitoring_Station* subclass can be specialized in channel station, pipeline station, well station, reservoir station, climatic station or fluvial station. They provide support to certain activities such as water offers evaluation, water demand control, hydro-environmental monitoring (parameters like rivers flow, pluviometry etc), water quality analysis, operational information about hydraulic systems etc. Accordingly, the model will be used to help guide decisions on further evaluating and refining the monitoring network, and evaluating and refining action criteria.

The *Riverhead, Rivermouth, Confluence* and *Island* are subclasses to represent specific natural phenomena of rivers. *Inflect_Point* subclass represents great depressions, elevations, turnings, enlargements, reductions, waterfalls, etc, that change deeply the water way in a pipeline, channel or river. The *WaterWithdrawal* subclass can represent a hydraulic apparatus, or a system of works or fixtures, by which a supply of water is furnished for useful purposes, including dams, sluices, pumps, aqueducts, distributing pipes, etc. Another important hydraulic structure is represented by the *WaterDischarge* subclass. By means of that subclass one can represent the flow of waste and dirty water through the network. Finally, the *Dam* and *Hydroelectric_Plant* subclass that has special functions in the reservoir and river simultaneously.

Two important water-related areas are represented in the ACQUA Model and linked to *WaterWithdrawal* class and *Monitoring_Station* class. The first area associated to the *WaterWithdrawal* class represents the water offer for several purposes, for example industry, psiculture, concessionary, tourism and irrigation. The human subsistence, environmental preservation, service and industry development are influenced directly by water availability. The water offer has direct effects on water management, especially if we assume that the water demand is always higher than water offer. The situation is still more critical in semi-arid regions. By using the ACQUA model, it is possible to represent the main users of water resources in a given region. This can be done by means of the *User* class hierarchy. For example, it is possible to represent who has received grant to use water for irrigation. The second area linked to *Monitoring_Station* class deals with water quality problem. Freshwater demand is increasing rapidly, due to population growth and economic development. For that reason, the analysis of the quality is necessary to control water pureness and environment changes. The quality campaigns with its analysis and samples, measured according specific parameters, are represented by *Campaign*, *Analysis*, *Sample* and *Parameters* classes.

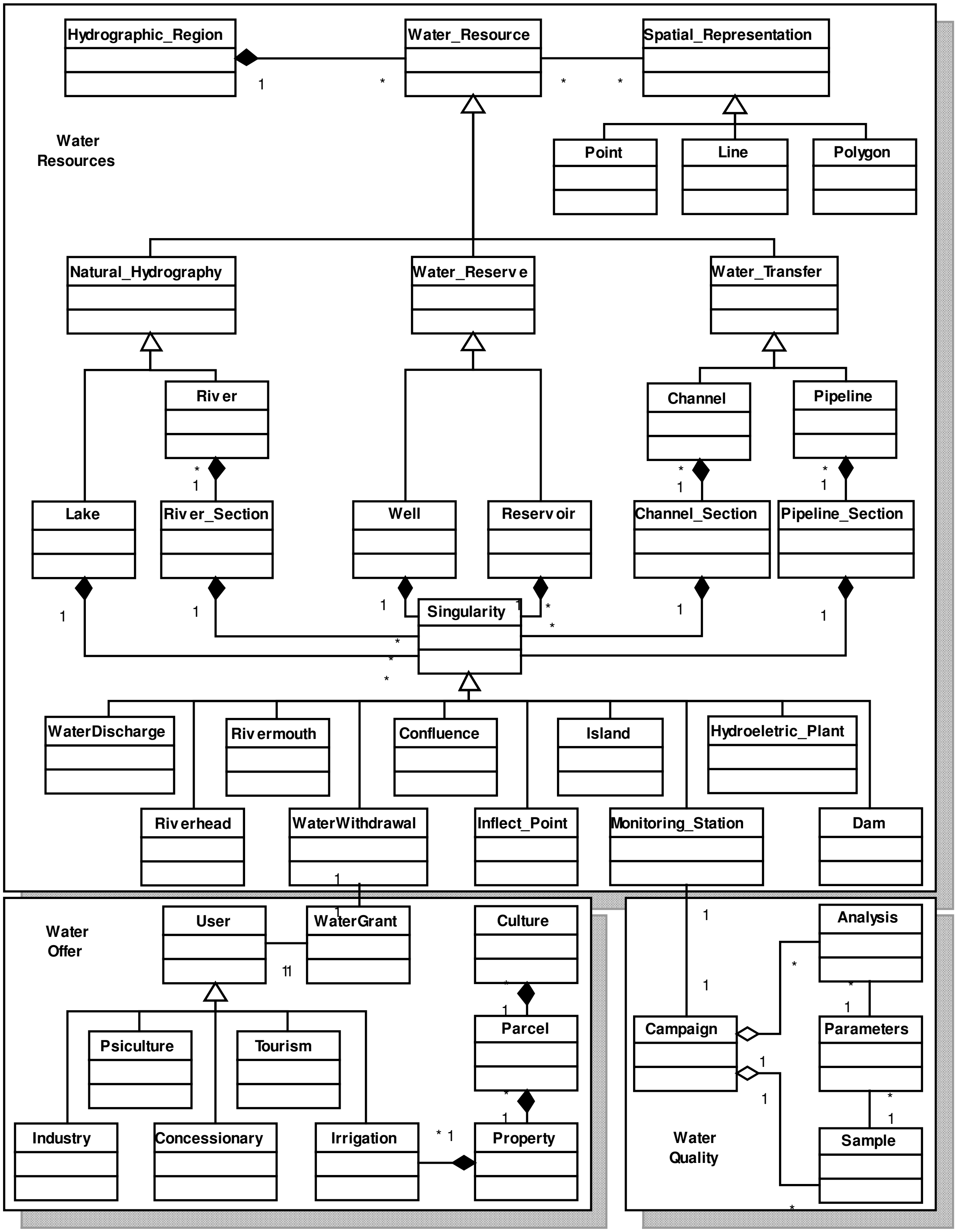

Figure 2 - The class diagram of the ACQUA model.

5. ACQUA VIEW TOOL

To demonstrate the applicability of our proposal it has been implemented a tool, called ACQUA VIEW, which gives the necessary support for managing water resources in a GIS environment. Developed on a Windows platform and using Delphi 5 components, the source code size is approximately 1.45MB. Initially, a database with water-related objects was created in ORACLE v.9i, based on the ACQUA model. Those objects represent water resources of the Brazilian State called Ceará.

The ACQUA VIEW tool has been used to explore the surface water world in an interactive form. This tool offers to final users the following functionalities: interactive analyses on georeferenced maps, “ad hoc” queries, spatial functions, graphical queries and a complete list of all attributes for each object selected on maps.

The query module offers pre-defined queries and SQL commands available through buttons, besides a complete list of tables and fields accessible by all users. The spatial functions provided by the tool answer questions like: *(i)* What are the dams or monitoring stations that we can find in a specific hydrographic region? *(ii)* What is the hydrographic region that a reservoir belongs to? *(iii)* What is the approximate distance in meters, between point A and point B, appointed on map? (See an example in Figure 3) *(iv)* What are the pluviometers found in a specific action ray? (See an example in Figure 4) *(v)* What are the singularities found in a stretch “n“ of river X?.

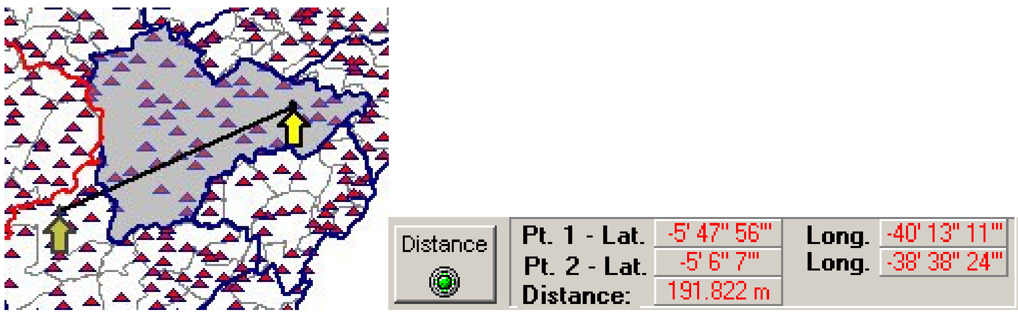

Figure 3 - Spatial Distance Function.

Figure 4 - Ray Action Spatial Function.

For a more accurate representation, evaluation and analyses, all available maps have a zoom function, up to 40x. The map is presented in Figure 5. Furthermore, when a reservoir is selected on map, for example, a short window shows the water mirror of the selected object, detailing its features (See Figure 6). Another panel shows dynamically the names of region, object and city focused, and the approximate localization (latitude and longitude) (See Figure 7). Specific colors are designated for selected regions, selectionable regions, regions in focus and cities, but can be modified by the users during running application.

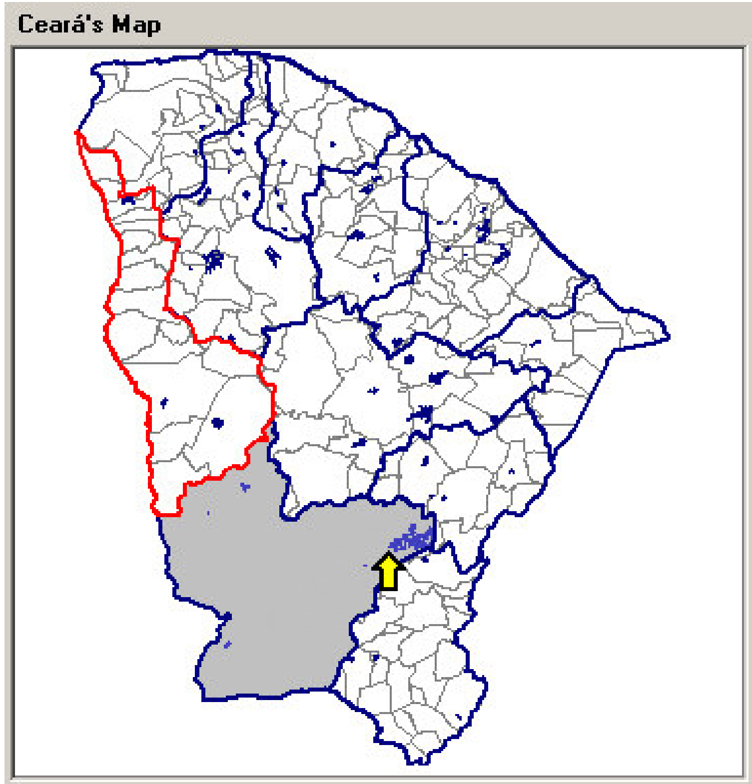

Figure 5 - Main Map.

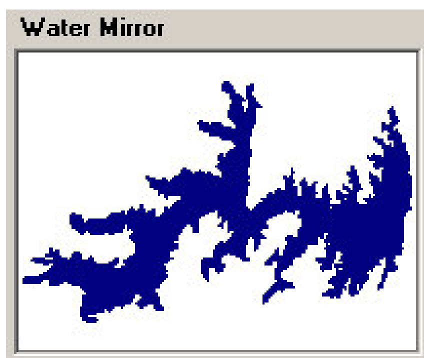

Figure 6 - Water Mirror Map.

Figure 7 - Information Panel.

6. DISCUSSION

Except "Roads" model (that presents map generalization rules and appropriate entities for representing traffic objects), the other models for modeling geo-referenced phenomena investigated are for general purposes. A data model for general purposes transfers to final users the design tasks related with identification, organization and storage of the geographic phenomena. On the other hand, a data model specific to a given area (e.g. water resources) models geographic phenomena more efficiently since it associates a meaning (semantic) to the phenomena. The gist of the ACQUA model is to capture and represent, all water resources involved in a water network (with its geometry, topology and properties).

In the models with generic representation (geo-objects), the user has to associate "meanings" (semantic) to spatial objects. In the diagram class (Figure 2), observe that the mapping class is almost intuitive when we use the ACQUA model. We will come back to Mandacaru River example (Figure 1 and Table 1) to demonstrate how the ACQUA model represents this geographic phenomenon.

In order to model the water resources network related to the Mandacaru River, an object representing the Mandacaru River is created in the River subclass, a subtype of Natural_Hidrography class and also a subtype of Water_Resource class. For each stretch of Mandacaru River is created an instance in the River_Section subclass, and for each begin and end point an instance is created in the specific subclass of the Singularity class, if it was not created yet. If you have already created an instance, only the relationships are created. Observe that the topological relationships of the water resources network of Mandacaru River are represented in the model. For example, the Dam (river sections 6 and 7) and the Island (river sections 7 and 8) are singularities identified along the course of the river.

Singularities which represent begin or end points, discovered inside or at river's margins, and its topological relationships, add details to the model and help to represent a structure nearest the real structure of the water resources network of the river. The same concept is adopted to represent a channel or pipeline network and its singularities.

As we can see the mapping class is simple and represents each object with its geometry and properties, but also considering its relationships, necessary to understand each phenomenon.

7. CONCLUSIONS

Nowadays, water is a scarce natural resource in worldwide scale. Freshwater problems are acute and worsening. Most arise from the poor management of water resources, lack of financial resources required for sustainable development and efficient utilization of resources, absence of effective regional and basin development plans and shared management, and under-estimation of the groundwater potential to supplement irrigation and drinking water supplies. A GIS should be used to provide support for an efficient water resource management. For that it is necessary providing models for capturing semantics of water resource related geographic phenomena. In this paper, we propose a data model, denoted ACQUA, for representing water-related geographic phenomena, such as lakes, wells, reservoirs, rivers, channels, pipelines and singularities.

ACQUA provides the necessary elements for modeling the complex water resources network, its physical properties and its spatial information, within a geographic context. ACQUA model also provides the information to support scenario simulation for operation of hydraulic systems and to analyze the impact of new hydraulic structures in a given hydrographic region

The proposed model has the capability to produce key information for water resource management, such as:

- The graphical representation of all the water resources of a given hydrographic region;
- How many and what irrigators have grant to use water of a certain resource;
- The water flow required for ensuring the water demand for a given region;
- The water volume that comes in a given water resource;
- The number of wells registered in a given hydrographic region;
- The consumption of water in a planted hectare for a given culture, in a given city, applying a certain irrigation method;
- The behavior between the planned and real water flow in a given year for a certain dam;
- The spatial distribution of monitored wells;
- The impact of rain incident in a hydrographic region;
- The water volume evolution of reservoirs.

Object-relational DBMSs provide the benefits of the object-oriented paradigm (for example, encapsulation, inheritance, methods and polymorphism) and the best capabilities of relational database technology into a single engine (Stonebraker and Brown [17]). For that reason, we have decided for implementing the ACQUA model using the OR database technology.

Finally, the ACQUA model has been used at FUNCEME, a governmental foundation of the Ceará State (northeast of Brazil), with recognized services given for the community in water-related problems. A database has been designed and developed at FUNCEME, using the principles of the ACQUA model to storage the information about the water resources of the Ceará State.

References

[1] Shekhar S., Coyle M, Goyal B, Liu D, Sarkar S. Data Models in Geographic Information Systems. *Communications of the ACM*, April 1997, vol 40, n.4.

[2] T. Nyerges. "Understanding the Scope of GIS: Its Relationship to Environmental Modeling", chapter 8 in Environmental Modeling with Geographic Information Systems, edited by Michael Goodchild, Louis Steyaert, and Bradford Parks. London: Oxford University Press, pp. 75-93, 1993.

[3] Câmara, G., Casanova, M., Hemerly, A. S. Geographic Information Systems Anatomy. Campinas: Computer Institute , UNICAMP, 1996. xi, 197 p. li. (In Portuguese).

[4] Silberschatz, A.; Korth H.F.; Sudarshan S. Database System Concepts. McGraw-Hill, 1996.

[5] Campos, N.; Studart, T. Water Management: Principles e Practices. Porto Alegre: ABRH, 2001 (In Portuguese).

[6] Souza Filho, F.de A.. Water Resources Management Experiences. Environment Ministry (MMA) and Water National Agency (ANA), November, 2001 (In Portuguese).

[7] Borges, K. A. V., Davis Jr., C. A., Laender, A. H. F. OMT-G: An Object-Oriented Data Model for Geographic Applications. *GeoInformatica* 5(3):221-260, 2001

[8] Erwig M., Guting R. H., Schneider M., Vazirgiannis. M. Abstract and Discrete Modeling of Spatio-Temporal Data Types. *Proc. of ACM GIS*, pages 131-136, November 1998.

[9] Patel J., Yu J., Kabra N., Tufte K., Nag B., Burger J., Hall N., Ramasamy K., Lueder R., Ellmann C., Kupsch J., Guo S., Larson J., DeWitt D., Naughton J. Building a Scalable Geo-Spatial DBMS: Technology, Implementation, and Evaluation. *Proc. of ACM-SIGMOD*, pages 337-347, 1997.

[10] Jones C.B., Ware J.M. A Data Model for Representing Geological Surfaces. *Proc. of ACM-SIGMOD*, pages 20-23, 1997.

[11] Balaguer F., Gordillo S. Refining an object-oriented GIS design model: Topologies and Field Data. Proc. of ACM-GIS, pages 56-61, November 1998.

[12] Laender A. H. F., Davis Jr. C. A., Borges K. A. V. Spatial Data Integrity Constraints in Object Oriented Geographic Data Modeling. Proc. of ACM-GIS, pages 1-6, November 1999.

[13] Dayal U., Manola F. PDM: An Object-Oriented Data Model. *Proceedings of the 1986 IEEE Data Engineering Conference*, pages 18-25, 1986.

[14] Manola F., Orenstein J. A. Toward a General Spatial Data Model for an Object-Oriented DBMS. *Proceedings of the Twelfth International Conference on Very Large Databases*, pages 328-335, August 1986.

[15] Nabil R. A., Subramanian R. The Design and Implementation of an Expert Object-Oriented Geographic Information System. *Proc. of ACM-CIKM*, pages 537-546, November 1993.

[16] Barnett L., Carlis J. V. A “Roads” Data Model: A Necessary Component for Feature-Based Map Generalization. *Proc. of ACM-GIS*, pages 58-67, November 1997.

[17] Stonebraker, M., Brown, P. Object-relational DBMSs: Tracking the next great wave. Morgan-Kaufmann. 1999.