

SOLAP for Spatial Multidimensional Analysis in Decision-Support Purpose

Case study: Natural Hazards effect on forest ecosystems managing

Abdallah Bensalloua Charef

Mathematics and Computer science dept -
University of Mostaganem
Computer science lab - University of Oran
P.O 1524 El M'Naouer, 31000 Oran, Algeria
Charef_bensalloua@yahoo.fr

Hamdadou Djamilia

Computer science lab - University of Oran
P.O 1524 El M'Naouer, 31000 Oran, Algeria
dzhamdadoud@yahoo.fr

Beldjillali bouziane

Computer science lab - University of Oran
P.O 1524 El M'Naouer, 31000 Oran, Algeria
bouzianebeldjilali@yahoo.fr

Abstract - In this paper, we discuss how Spatial On-Line Analytical Processing (SOLAP) system supports spatial and multidimensional exploration of data as a Decision System Support (DSS). Our purpose in this case study is to demonstrate how this solution can generate a full analysis to build management indicators in the domain of natural hazards and its relationship with the forest management. In fact, spatial interface can facilitate viewing and manipulating geometric aspect of data and even thematic and spatial analysis offered by Geographical Information System (GIS). In other hand, the multidimensional structure makes it easy to generate reporting as graphics, charts, maps and tables.

So multidimensional modeling can help in decision making process by exploiting data available at departmental organisms such as civil protection, environmental services, maps or other data relative to the effect of natural hazards on forest ecosystems managing.

For our proposal we use data relative to the region of *Relizane "Dahra Mountains"* in the western of Algeria.

Key-words –*Decision support system; SOLAP; spatial analysis; decision making; natural hazards; forest ecosystems managing.*

I. INTRODUCTION

Recently, the popularity of spatial information created from satellite images or other spatial applications, has generated a huge volume of data which need to be efficiently analyzed for decision making goal.

Data warehouses which are defined as “collections of historical, summarized, non-volatile data”, are optimized for on-line analytical processing (OLAP). In the same context, spatial data warehouse can be considered for Spatial OLAP operators on both spatial and non-spatial data, in a system which provides capabilities to interactively explore the spatial data warehouses.

SOLAP can be considered as “a system which can fully exploit the powerful concepts brought by the multidimensional database structure and to add spatial extensions that provide highly interactive map visualization and data exploration”.

We say that the SOLAP approach facilitates the interactive nature of the analytical process. In fact, it allows to explore and to navigate the multidimensional structure at different themes (dimensions) at different levels of granularity. This can give access to all possible combination of dimensions.

This technology can be used in the domain of forest ecosystems managing. In fact, these ecosystems are affected by natural hazard. So, managers must find mechanisms for data analysis in different themes, as well as geographical distribution of hazardous phenomena.

The present paper presents the formalization and conception of a prototype of SOLAP application in the context of natural hazards effect for forest ecosystems managing. Next section makes a brief overview on the related works in the context of natural hazards management

in one part and the spatial multidimensional modeling of data cubes. After that we describe the concept of OLAP tools and then introduce and discuss the SOLAP tools and their general characteristics. Next, we will describe the modeling and implementation of our SOLAP prototype for natural hazards data exploration. Finally the conclusion will expose the limits and the potentials of this technological approach.

II. RELATED WORKS

A natural hazard can be defined as “an event resulting from the abnormal intensity of a natural agent ...” [5]. The principle of reduction of natural hazards in Algeria is to be more and more interesting since the number of these events which had been these last years especially in 2003 at “Boumerdes” department (eastern of Algiers). Since that date, new tools have been created to control land use. These tools have the objective to specify where it is forbidden or to specify certain conditions to build. We must notify here, that the damages was not only human or industrial but also ecological especially for forest ecosystems. See “*Fig. 1*”.



Fig.1 Forest area damaged by the Marchkar avalanche. [2]

In this context, Decision Support System (DSS) can be used in the aim to help decision makers to analyze indicators produced by a prototype of Spatial OLAP. These indicators must take into account the spatial distribution of the population and of dwelling places in the vicinity of forests and especially where it had already taken a natural hazard or places which presents a real risk because of a number of elements whatever historical, geological or environmental considerations. [3]

the research shows that forest ecosystems play an important role in reducing the risk of natural disasters. here, we notify the role of forests which play in reducing human vulnerability to natural disasters. in fact, there are important implications of forest degradation for human security. Forest resources and services play an important role in the achievement of human security. Whether it is through the provision of livelihood resources (e.g. food, medicine, cooking fuel, construction materials, etc.) or ecosystem services such as water purification, climate regulation and

erosion control, forests are undeniably linked to the well-being and prosperity of people. [12]

When the availability of these resources and services is undermined, the resulting insecurity can range from hunger and susceptibility to disease, to loss of income and livelihoods, to social tensions and open conflict to acute vulnerability to natural hazards.

The relationship between forest ecosystems and human security is reciprocal. While forest degradation contributes to human insecurity, the reverse is also true. [1]

Policies for disaster risk reduction and management have shifted from defense against hazards to a more comprehensive, integrated risk approach.

The European Commission released the Communication on 'A Community approach to the prevention of natural and technological disasters'. This communication proposes that EU-level actions should focus on three areas: (1) developing knowledge-based prevention policies; (2) linking actors and policies throughout the disaster management cycle; (3) improving the effectiveness of existing financial and legislative instruments. [2]

The European Union has already developed a set of instruments to address various aspects of disaster preparedness, response and recovery. These include the Community mechanism for civil protection or the European Union Solidarity Fund [2].

The research report many disasters due to natural hazards causing huge amount fatalities, and billions Euros in overall losses. During the period 1998–2009, more than 11 million people (out of a population of 590 million, approximately, in the EEA member countries) were somehow affected by disasters caused by natural hazards

Most of the disasters are thereby caused by either floods or storms.

It is essential to consider all phases (e.g. prevention, preparation and intervention) and to take into account and consequently maintain all measures (e.g. spatial planning; technical measures within the cycle of integrated risk management" *Fig.2*". [10]

According to NatCatSERVICE¹, which provides reliable long-term data series adjusted for inflation, the number of disasters in Europe has been showing an upward trend since 1980, largely due to the continuous increase of meteorological and hydrological events "*Fig.3*". [10]

In most counties around the world follow specific policies regarding forest fires. They are often incorporated into forest management plans prepared by the civil protection authorities and interior ministries in these countries.

The total area burnt in Greece at the end of the 2007 season was estimated by the Greek authorities as 225 734 ha.

¹ <http://www.munichre.com/en/reinsurance/business/non-life/natcatservice/index.html>

Moreover, in addition to the environmental damage, the fires in Greece caused a high number of fatalities.

In most cases, weather conditions are the main driver of forest fire occurrence and impacts. However, the climate cannot be controlled. So, management options to reduce the

impact of forest fires must deal with the causes of these fires, and the proper management of forest fuels.

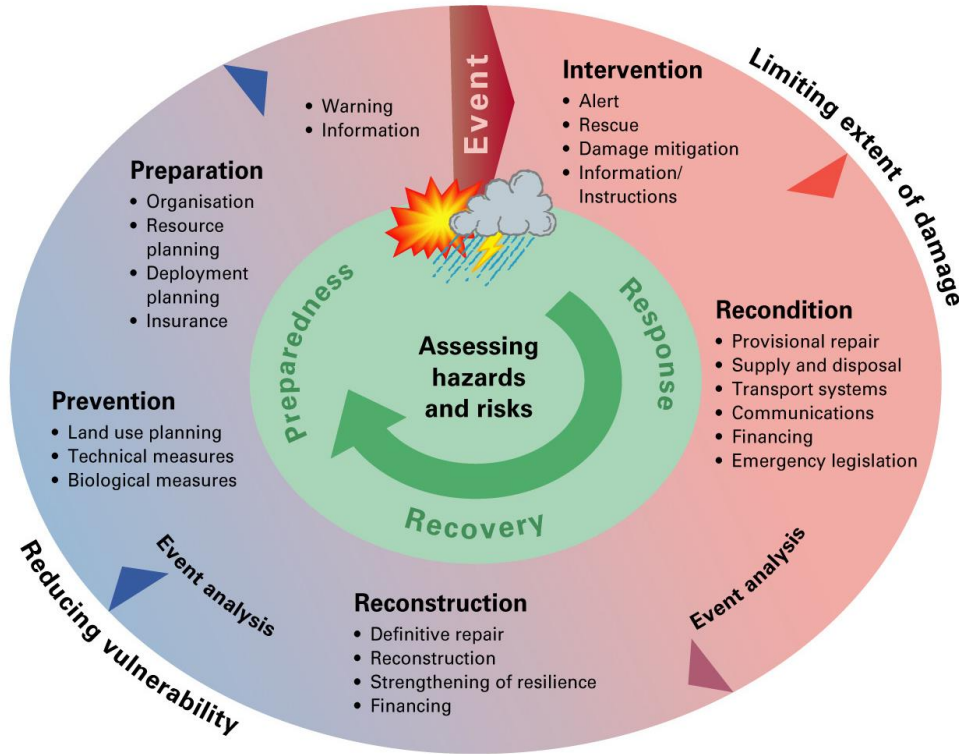


Fig.2. Cycle of integrated risk management. Source: Swiss Federal Office for Civil Protection FOCP, 2010. [10]

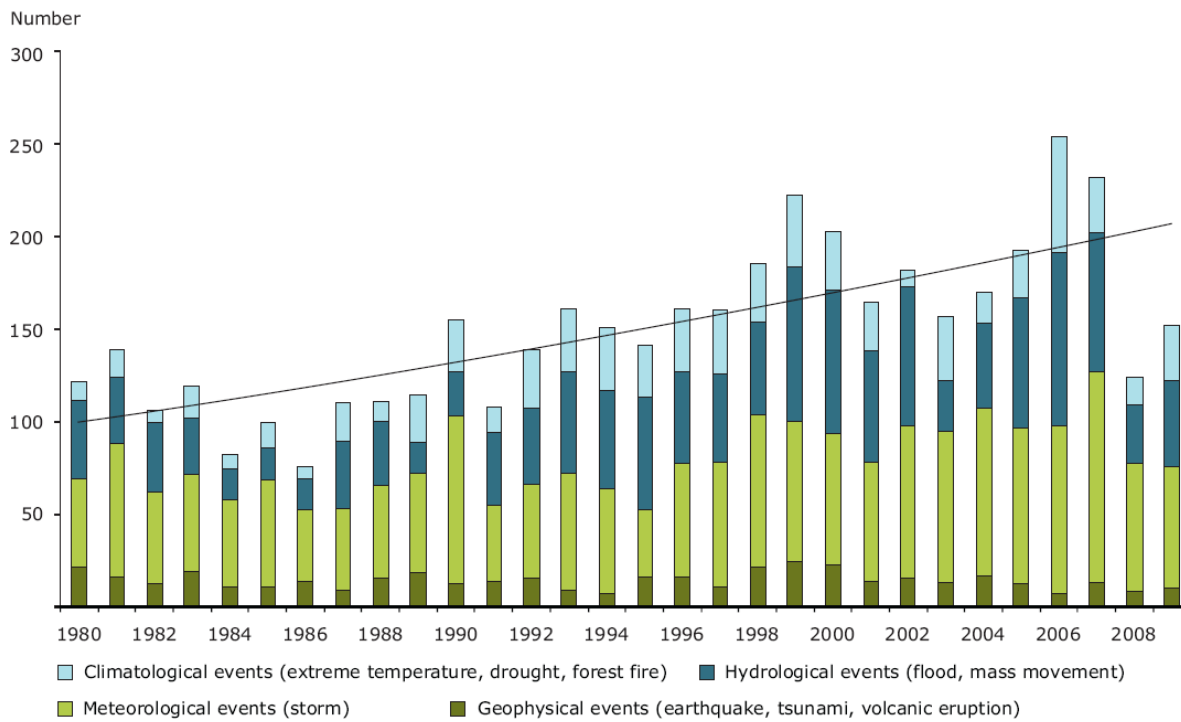


Fig.3. Disasters due to natural hazards in EEA member countries, 1980–2009. [10]

Source: NatCatSERVICE, 2010; © 2010 Münchener Rückversicherungs-Gesellschaft, Geo Risks Research, NatCatSERVICE — as at August 2010.

In this situation, the factors that influence forest fires are the human causality and weather conditions. "Fig.4"



Fig.4 Burnt scars caused by fires in Greece in 2007 mapped from MODIS satellite imagery [10]

Fire occurs in all forest ecosystems and is neither good nor bad. Neither is it neutral. The effects of fire help some members of the ecosystem and harm others. The degree to which organisms

benefit or suffer in a fire is related to the characteristics of the individual fire and the characteristics of the organism.

Important characteristics of a forest fire are the speed at which it travels, the heat at which it burns, and the height of the flames. Together with the frequency with which fire occurs, these characteristics are called the fire regime.



Fig.5 Fire in the forest ecosystem

The fire that swept through here moved through the crowns of the trees, burning off most foliage and branches. Crown fires are more common where trees are close together and winds are strong. [11]

III. OLAP AND SOLAP CONCEPT

On-line analytical processing is defined as “the name given to the dynamic enterprise analysis required to create, manipulate, animate and synthesize information from exegetical contemplative and formulaic data analysis models. This includes the ability to discern new or unanticipated relationships between variables, the ability to identify the parameters necessary to handle large amounts of data, to create an unlimited number of dimensions, and to specify cross-dimensional conditions and expressions”[1]

OLAP system is based on multidimensional approach which gives the concept of: dimension, members, measures, facts and data cube. A dimension can be considered as an axe of analyze. They can be spatial such as location names. They are also organized hierarchically into levels of granularity or details. Example: country-region-department.

A dimension is composed of members who can be aggregated to constitute the members of a higher level.

A measure is a numerical attribute analyzed by a system user by combination of a set of dimensions.

The combination of dimension members with the resulting of measures value for a particular aggregation level represents a fact [2]

A data cube is composed of a set of measures aggregated according to a set of dimensions. "Fig.6".

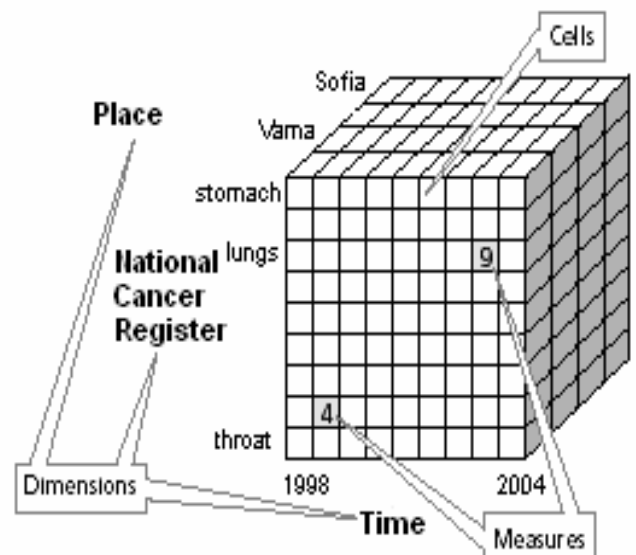


Fig.6. A multidimensional model data cube: - The cube is composed of cells that define fact attributes [13]

Every OLAP system has an architecture which can be divided into three parts: the multidimensionally structured database, the OLAP server that manages the database, and the OLAP client that gives access to the database via the OLAP server "Fig.7". This can be by exploring and analyzing the data using different visualization methods and adapted operators such as drill-down, roll-up, drill-across and swap.

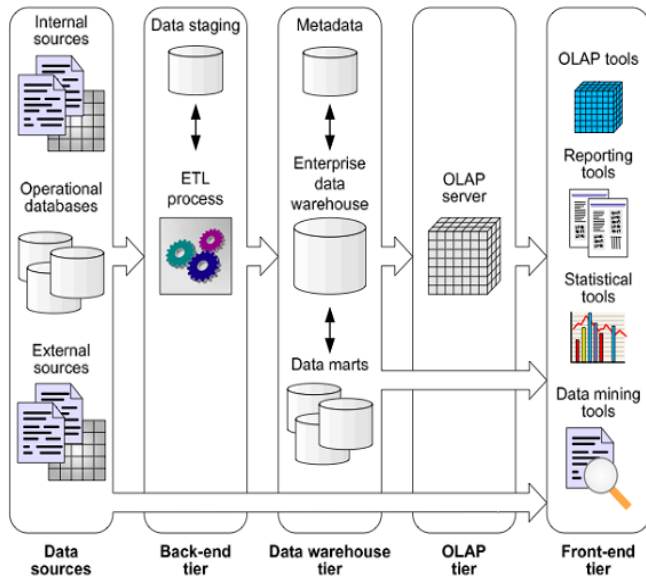


Fig.7. OLAP-based decisional system

OLAP application represents a good support for the usage of spatial dimensions in a multidimensional analysis process. However, this usage can be without consideration for the cartographic component of the data, which signifies that no spatial visualization or exploration of data will be permitted.

With the development of spatial data warehouses concepts, a new solution was then developed, which consists of combining the strengths of GIS, with the strengths of OLAP tools into a new type of application. This combination gave birth to Spatial OLAP applications which allows fully analyze of the geometric aspect of the spatial data.[4]

From these concepts a Spatial OLAP can be defined as “a visual platform built especially to support rapid and easy spatiotemporal analysis and exploration of data following a multidimensional approach comprised of aggregation levels available in cartographic displays as well as in tabular and diagram displays” [6].

SOLAP tools support the multidimensional database structure described earlier. However, with the spatial data manipulation capabilities included, three types of spatial dimensions can be used: the non-geometric spatial dimensions, the geometric spatial dimensions and the mixed spatial dimensions. "Fig. 8" presents the three types of spatial dimensions.

In the first case, which is used in conventional OLAP, no cartographic representation is associated with the members.

The second case, in order to help in graphical visualization and queries, it include geometric shapes geo referenced to the map. The third case as its name indicates, it uses nominal data and geometric representation for the dimension members.

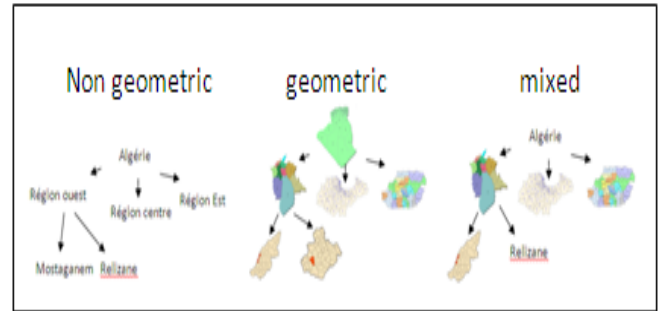


Fig. 8. The three types of spatial dimensions supported by SOLAP tools

Measures in SOLAP

In a SOLAP application, a measure can possess a geometric representation called then spatial measure, or only numerical one as in a conventional OLAP.

In a spatial measure we have two types: The first type consists of a set of coordinates, which requires a geometric operation, such as a spatial union, a spatial merge or a spatial intersection, to be computed. It is the set of all the geometries representing the spatial objects corresponding to a combination of dimension members. The second type of spatial measure results from the computation of spatial metric or topological operators. Ex: surface, distance number of neighbors.

Finally, the measure values (spatial and non-spatial) that result from the combinations of spatial and non-spatial dimension members are visualized using a SOLAP client.

A SOLAP client can be used with any type of SOLAP architecture: relational (ROLAP), multidimensional (MOLAP) or hybrid (HOLAP).

SOLAP operators

In the SOLAP application, different operators are used in order to take advantage of the multidimensional data structure. The most important operators are drill-down, roll-up, drill-across and swap.

A drill-down operation allows the end-user to navigate from a general level to a more detailed level inside a dimension

A rollup operation allows the user to navigate from a detailed level to a more general level inside a dimension

A drill-across allows to view different information but at the same level of detail.

A swap operation allows interchanging two dimensions in order to view the same information in a different way or in order to view different information.

If we manipulate the data on maps, these operators will be named spatial drilldown, spatial roll-up and spatial drill-across.

Categories of SOLAP

GIS and OLAP tool integration can be in different approaches: GIS-dominant, OLAP-dominant and total integration.

In the first approach we find full GIS capabilities, and practically no OLAP functionality except simplified access.

The second approach offers full OLAP capabilities but only cartographic GIS functionalities

The third approach necessitates fully integration of GIS and OLAP capabilities. However, Implementing a SOLAP system in this case without using SOLAP technology requires major development.[7]

IV. SOLAP APPLICATION FOR NATURAL HAZARDS

A. Modeling of the solution

We have taken the following decision model "Fig. 9"

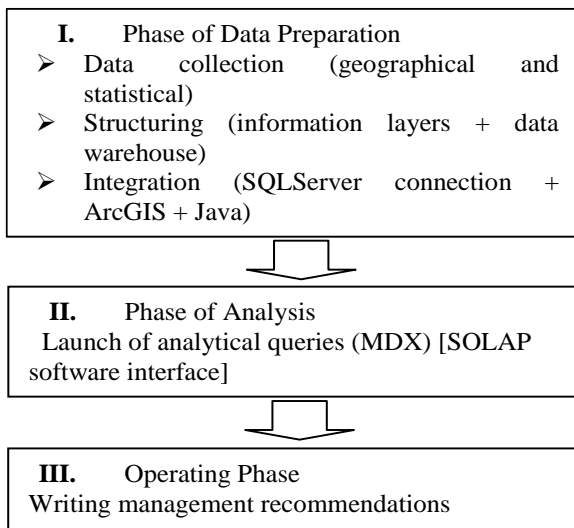


Fig.9. Decisional model for natural hazards management

To construct SOLAP architecture the first step consists in building the multidimensional database. We have done it with a relational environment. It is composed of one table called the fact table containing the measures which are values of indicators, and a set of other tables called dimension tables representing the analysis axes. [16]

The Unified Modeling Language (UML) has become the standard for modeling systems using object-oriented concepts. UML is an extensible language, in the sense it provides mechanisms such as: stereotypes, tagged values, and constraints; that allow introducing new elements for specific domains if it is needed, such database applications, business modeling, etc. [18]

A UML extension defined using a profile must be strictly additive to the standard UML semantics.

This means that such extension must not conflict with or contradict the standard semantics.

A stereotype is a model element that defines additional values (based on tagged values), additional constraints, and optionally a new graphical representation (an icon). A stereotype allows us to attach a new semantic meaning to a model element. [19]

In many works, an extension to the UML is used for MD modeling, as it easily and elegantly considers main MD properties at the conceptual level such as the many-to-many relationships between facts and dimensions, degenerate dimensions, multiple and alternative path classification hierarchies, and non-strict and complete hierarchies "Fig. 10". [14][17]

In order to facilitate MD Modeling, authors in [15] use meta-models of the spatial components "Fig. 11"

The list of indicators will serve to define the features of the MDDB. In our prototype we have used a list of management indicators such as: number of buildings, number of residents, flood zones area ... and others. [8]

Each measure value is calculated for a combination of dimensions and aggregated for each hierarchical level of the selected dimensions.

Its value constitutes one cell of the hazard's data cube: one cell is an intersection of all dimensions used in an analysis process.

To generate good indicators in this process the results will be from the combinations of geographical and non geographical information.

In our prototype, the dimensions used can be: spatial dimensions, temporal dimensions, nominative dimensions. The nominative dimensions are the dimensions without spatial representation "Fig. 12".[9]

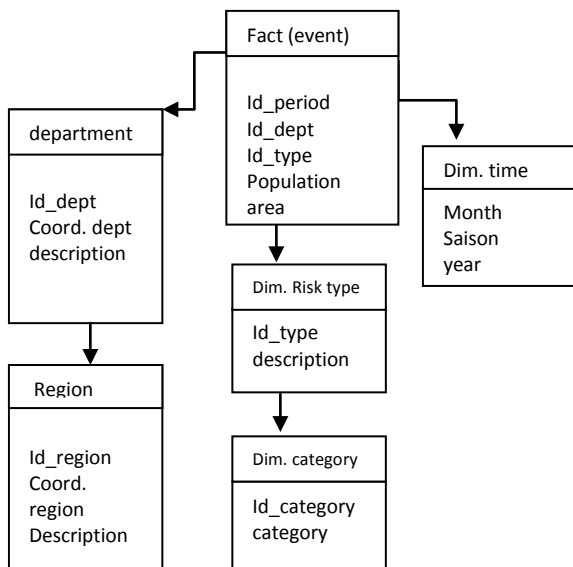


Fig. 12. Fact and dimensions tables of the Data Cube

In this model fact table contains all foreign keys of the dimension tables and all the measures. The schema is not normalized as becomes obvious, and this for facilitates analysis queries which implies some redundancies.

For small or medium-sized data volumes, such schemas have a sufficient performance because join operations are only necessary between the fact table and the related dimension tables. But for important sized data the snowflake schema is recommended: instead of modeling each dimension by one table, a table is created for each level of a hierarchical attribute.

B. Implementation of the prototype

For the implementation we used the platform SQL Server 2008 R2 which offers a set of features allowing the creation of components modules of our decision model. in parallel, to

present the geographic information, we used ArcGIS technology. In fact, ESRI demonstrated how the ArcGIS software range can work with SQL Server 2008 for the Spatial Analysis.

The "Fig. 13" demonstrates how to connect our database created on SQL Server with ArcGIS.

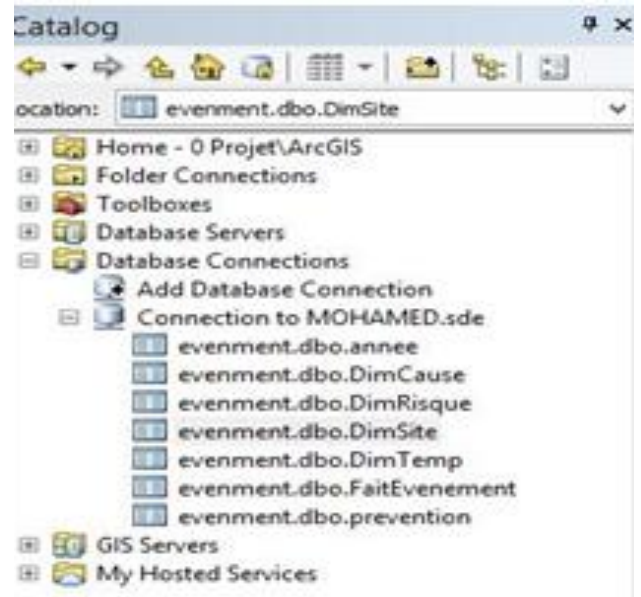


Fig. 13. Connecting the database to ArcGIS

Our prototype is developed on Eclipse software (Java platform). "Fig. 14" below shows how to connect our database to Eclipse.

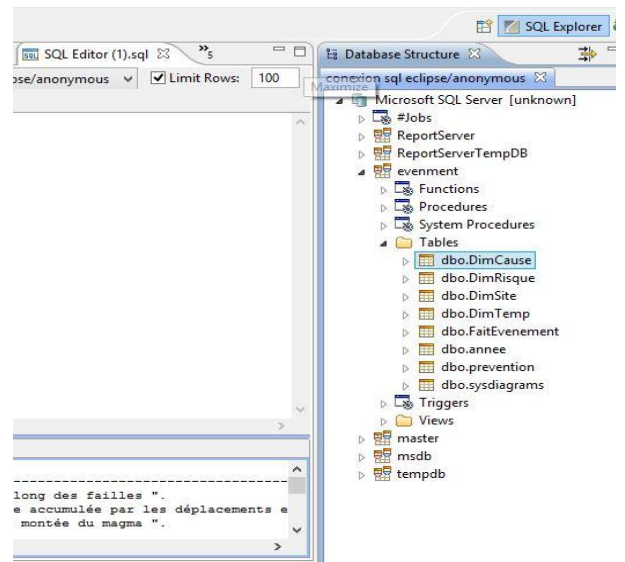


Fig. 14. Connecting the database to Eclipse

Thereafter we proceed to the creation of the multidimensional structure (data-cube) with the SQL Server

Analysis Service (SSAS). We can easily browse data on it with different axes and at different levels of detail "Fig.15".

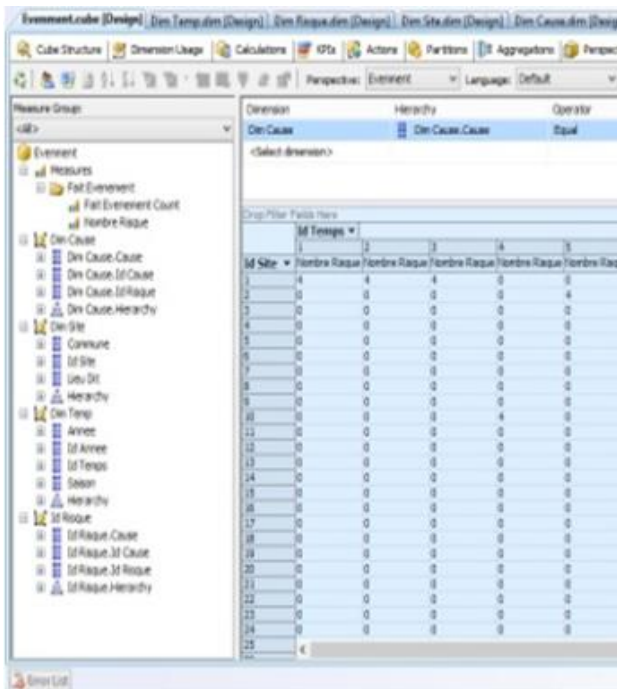


Fig.15 Browsing the Cube data

Our first purpose was to generate an integrated solution of SOLAP with spatial and multidimensional analysis. This is offered thanks to the software chosen. " Fig.16" demonstrates the final solution.

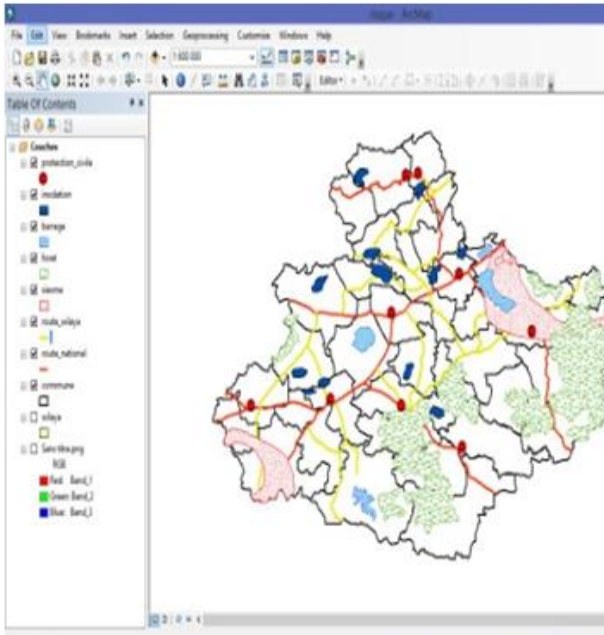


Fig.16. SOLAP prototype for natural hazards

V. DISCUSSION AND CONCLUSION

In this work, we perform a SOLAP application as a decision support. This application is intended for managers who want to make forest ecosystems managing.

Indeed, they can perform spatial and multidimensional analysis according to different themes namely the time, risk categories and location corresponding to sites and areas subject of natural hazards, by querying a multidimensional database developed in this work.

The final interface is OLAP based architecture. Indeed, we used the OLAP server (SQL Server) to benefit from all its features. Spatial analysis is performed in GIS software (ArcGIS) that supports the connection with the first in a prototype developed with Eclipse (Java platform).

The data warehouse is enriched by data for one region in western of Algeria which is "Relizane department": all combinations are calculated from the granular level to the aggregated levels of each dimension and stored in a data cube.

The data cube is modeled using extension mechanisms offered by UML. Then, it is implemented in a relational database environment (MSSQL Server 2008R2). All geographical computations have been realized with Arcview (Esri). Once the data loaded into the cube, SOLAP interface gives a lot of flexibility to generate charts, maps and cross tables. The operators drill-up, drilldown, slide and roll-up gives the possibility to set the dimensions with one or more of its members, to consult the aggregations for all types of representations.

The use of a clean SOLAP server allows benefiting all spatial and multidimensional analysis features. However the integration of multidimensional queries (MDX) poses some optimization problems.

However Spatial OLAP presents some limitative aspects. The data structure must contain all combinations for all levels of aggregations. It generates very important volumetric data problems. The number of dimensions and the number of hierarchical levels per dimension must be restricted in order to limit the amount of data.

Otherwise, algorithms for integrating multi criteria analysis can be applied into data cubes for facilitate the selection of appropriate values which can be useful for decision support aims.

REFERENCES

- [1] B éard. Y, "Merging Hypermedia GIS with Spatial On-Line Analytical Processing: Towards Hypermedia SOLAP", in Geographic Hypermedia: Concepts and Systems 2006, pp. 167-185.
- [2] B éard. Y , PROULX Marie-Jos é et RIVEST Sonia. "Enrichissement de l'OLAP pour l'analyse géographiq ue : exemples de réalisation et différentes possibilités technologiques". Revue des Nouvelles Technologies de l'Information - Entrepr is de données et l'Analyse en ligne, 2005, pp. 1-20.
- [3] El MOHAJIR et OUBENAALLA Youness. "A geo-spatial decision-support system for regional air quality management". International Journal on Computer Science and Information Systems Vol. 4, No. 1, pp. 64-74

- [4] HERNANDEZ Vera, VOSS Angi, GÖHRING Wolf et HOPMANN Cornelio. Sustainable decision support by the use of multi-level and multicriteria Spatial analysis on the Nicaragua Development Gateway. In : From pharaohs to geoinformatics Proceedings of FIG Working Week 2005 and 8th International Conference on the Global Spatial Data Infrastructure, 16-21 Avril 2005, Le Caire, Egypte.
- [5] Julien Iris, Aldo Napoli, Franck Guarnieri, "Contribution of a Spatial Multidimensional Portal for Natural Hazards Data Exploration" *EnviroInfo 2006 (Graz) Managing Environmental Knowledge*
- [6] Mohamed Bakillah. «Développement d'une approche geosemantique integree pour ajuster les resultats des requetes spatiotemporelles dans les bases de donnees geospaciales multidimensionnelles evolutives », Memoire pour l'obtention du grade de maitre es sciences. Faculte de foresterie et de Geomatique universite Laval Quebec. 2007.
- [7] M.J. Proulx, Nadeau, S. Rivest, Y. Bédard, P. Gosselin et G. Lebel, "Géomatique et santé environnementale: innovations résultant du projet ICEM/SE", Colloque AGMQ, Quebec, Canada October 2008.
- [8] S. Rivest, Y. Bédard, MJ Proulx, M. Nadeau, F. et J. Hubert Pastor, "SOLAP technology: Merging business intelligence with geospatial technology for interactive spatio-temporal exploration and analysis of data". *Journal ISPRS de photogrammétrie et de télédétection*, 2005, pp 17-33.
- [9] STOLTE Chris, TANG Diane et HANRAHAN Pat. Multiscale visualization using data cubes. *IEEE Transaction on Visualization and Computer Graphics*, 2003, Vol. 9, n°2, pp.176-187.
- [10] André Wehrli, André Jol and David Sauri. "Mapping the impacts of natural hazards and technological accidents in Europe. An overview of the last decade" Publications Office of the European Union, Luxembourg, 2010.
- [11] Catherine Raven, *the green world, forestry*, Chelsea house publishing, 2006.
- [12] Anne Hammill, *arborvitæ*, The IUCN/WWF Forest Conservation Newsletter, 2005.
- [13] Antoaneta Ivanova, Constructing Object Oriented Class for extracting and using data from data cube. *International Conference on Computer Systems and Technologies - CompSysTech'2005*
- [14] Juan Trujillo, Emilio Soler, Eduardo Fernández-Medina and Mario Piattini, A UML 2.0 profile to define security requirements for Data Warehouses. *Computer Standards & Interfaces* 31 (2009) 969-983
- [15] Mehrdad salehi, developing a model and a language to identify and specify the integrity constraints in spatial datacubes, Thèse présentée à la Faculté des études supérieures de l'Université Laval dans le cadre du programme de doctorat en sciences géomatiques pour l'obtention du grade de Philosophiae Doctor. Université Laval Québec 2009
- [16] Sandro Bimonte, Omar Boucelma, Olivier Machabert, Sana Sellami, "A new Spatial OLAP approach for the analysis of Volunteered Geographic Information", *Computers, Environment and Urban Systems* 48 (2014) 111-123.
- [17] Kamal Boulil, Florence Le Ber, Sandro Bimonte, Corinne Gra and Flavie Cernesson, "Multidimensional modeling and analysis of large and complex watercourse data: an OLAP-based solution", *Ecological Informatics* 24 (2014) 90-106.
- [18] Kamal Boulil, Sandro Bimonte and Francois Pinet, "Conceptual model for spatial data cubes: A UML profile and its automatic implementation", *Computer Standards & Interfaces* (2014), <http://dx.doi.org/10.1016/j.csi.2014.06.004>
- [19] Sergio Lujan-Mora, Juan Trujillo, Il-Yeol Song, A UML profile for multidimensional modeling in data warehouses, *Data & Knowledge Engineering* 59 (2006) 725-769