Integration of Multi Criteria Analysis Methods to a Spatio Temporal Decision Support System for Epidemiological Monitoring

Farah Amina Zemri
LIO Laboratory
BP 1524 EL M' Naouer, Oran University, Algeria
zemriamina@gmail.com

Djamila Hamdadou
LIO Laboratory
BP 1524 EL M' Naouer, Oran University, Algeria
dzhamdadoud@yahoo.fr

Abstract – The present study aims to integrate Multi Criteria Analysis Methods (MCAM) to a decision support system based on SOLAP technology, modeled and implemented in other work. The current research evaluates on the one part the benefits of SOLAP in detection and location of epidemics outbreaks and discovers on another part the advantages of multi criteria analysis methods in the assessment of health risk threatening the populations in the presence of the risk (presence of infectious cases) and the vulnerability of the population (density, socio-economic level, Habitat Type, climate...) all that, in one coherent and transparent integrated decision-making platform. We seek to provide further explanation of the real factors responsible for the spread of epidemics and its emergence or reemergence. In the end, our study will lead to the automatic generation of a risk map which gives a classification of epidemics outbreaks to facilitate intervention in order of priority.

Keywords – Multi criteria Analysis Decision Support (MCAM), Spatial Data Mining (SDM), Spatial on Line Analysis Processing (SOLAP), Data warehouse (DW), Epidemiological Surveillance (SE).

1. INTRODUCTION

Epidemic prevention is a public health concern. It is a real challenge. Identification of residential areas (urban and non-urban) exposed to epidemics would help in riding of these phenomena of public health with prevention strategy and careful management. The medical management of these diseases would be more effective. The use of information technologies greatly facilitates the realization of such objective. In order to identify a good prevention strategy against epidemics and to ensure a reflect management of propagation phenomenon, a good epidemiological surveillance system must be developed for monitoring of the disease and identifying areas with epidemics outbreaks.

The article describes in section 2 our contribution. In section 3, the main works in the field of spatial decision support using MCAM are presented. Section 4 describes, EPISOLAP system and the proposed approach is illustrated in details in section 5. Multi Criteria formulation problem is given in section 6 and section 7 is devoted to the PROMETHEE method tool used for the development of multi criteria decision support system suggested. A real case study which is a first validation step of our proposed approach is detailed in section 8 and finally, we conclude our discussion in Section 9, giving some perspectives.

2. PROBLEMATIC AND CONTRIBUTION

Business intelligence provides new solutions for modeling, querying and visualization of data in an objective decision support. Multidimensional or hyper-cube models allow structuring the data for policy analysis by clarifying the notion of dimension.

The integration of spatial data into OLAP systems is a major challenge. Indeed, geographic information is frequently present implicitly or explicitly in the data, but generally under-used in the decision making process. Coupling OLAP systems and Geographical Information Systems in Spatial OLAP (SOLAP) systems is promising. We believe that the combination of SOLAP technology once designed and implemented with Multi Criteria

Analysis Methods (MCAM) is an interesting voice because it can lead to richer data analysis.

3. RELATED WORKS

In the context of single-actor decision support, several decision support systems in TP (Territory Planning) caught our attention. In [3], the system MEDUSAT is proposed for locating the site of a waste treatment plant in Tunisia. MEDUSAT combines a GIS tool allowing creation of homogenous areas determined from spatial data and common land (constituting a similarity index); these areas constitute the set of actions which are then processed by Multi Criteria Analysis Methods (MCAM).
The author in [5] proposed a decision process for water management in urban environment and in [6]; authors presented some tools for decision support in local communities in order to address water management problems. In [7], Multi Criteria analysis was used as a tool for decision making for spatial localization of areas under heavy human pressure, a case study of the department of Naama in Algeria was presented in the same work. Various decision support systems rich in spatial tools and Multi Criteria Analysis Methods were developed for management and decision making in territorial problems (water, air, natural areas, transportation, energy, waste, health planning, risk management, . . .) [13].
All these systems integrate in various levels multi criteria analysis tools coupled with GIS, but they consider criteria as independent and unable to model any interaction between them (interchangeability, correlation, preferential dependence . . .). In [14], we have already discussed significantly the inclusion of correlation criteria, in the MCDA methods, particularly"ELECTRE TRI", by introducing the Choquet integral (instead of the arithmetic sum) as an aggregation operator. In [9], the main objective is to develop a decision support system, for itinerary road modification in the case of hazardous materials transport. In [10], a multicriteria decision support system for industrial diagnosis was developed.

4. DESCRIPTION OF EPISOLAP SYSTEM

The study that was conducted by our previous works in "EPISOLAP" project [8] aimed to identify and predict the health risk according to surveillance data identified in different specialized health structures.

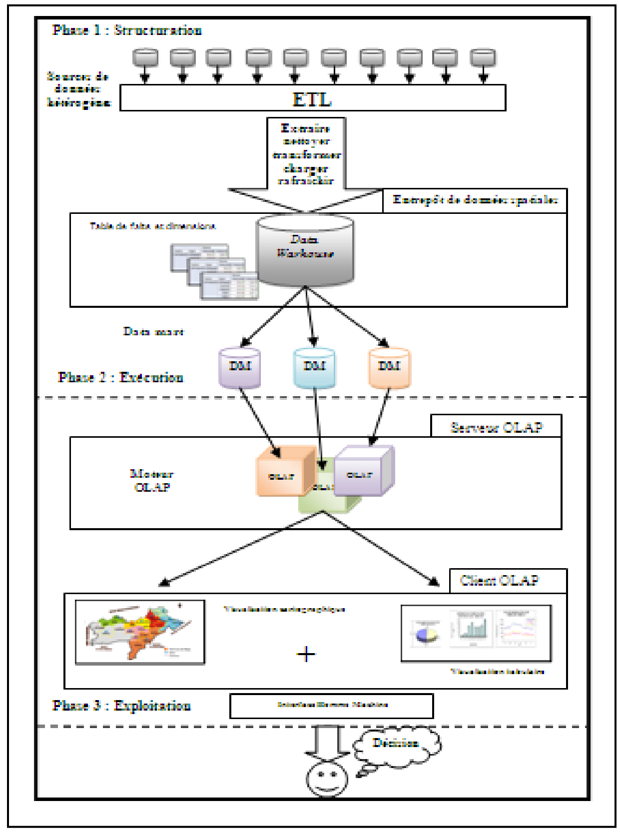

Figure 1: Decision Support System in Epidemiological Monitoring "EPISOLAP"

The analysis is done by SOLAP tool, but it is limited to alphanumeric data and does not exploit the geographical location and the link neighborhood. Our current work aims to emerge from all these data, the relevant structures of the health risk that can support the effort of surveillance, direct action eradication and strengthen the system of prevention.
More specifically, it is also spending limits of "EPISOLAP" and integrate the spatial nature of the data (here the epidemic outbreaks) and the interaction with the geographical environment especially that this is a disease (Tuberculosis) which is rapidly spreading (or the concept of neighborly relations is very important) allowing, in this application example, to explain and predict health risks threatening the population, taking into account their geographical context.
We show in Figure 2 some results obtained from the analysis of surveillance data by EPISOLAP system.

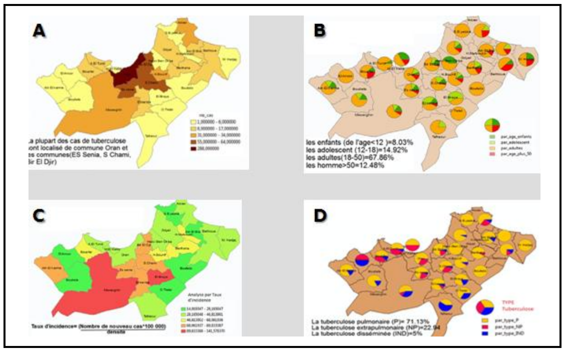

Figure 2: Geo location and Map Display on SQL server 2012 & ArcGis 10.0

A: Number of cases per municipality
B: Incidence rates by age and municipality
C: Incidence rate per municipality
D: Incidence Rates by Type of TB and municipality.

5. OUR CONTRIBUTION

In the present study, we began a new SOLAP formulation which aims to integrate Multi criteria analysis methods to answer to the requirements of epidemiological surveillance and to the most boring questions of decision makers in public health, allowing them to prevent the emergence of new epidemics outbreaks (prediction) and taking into account the socio-environmental factors favoring contamination.

The "EPISOLAP" system had as main objective the detection and localization of disease outbreaks; it remains to know which outbreaks are the most at-risk; we therefore proposed a classification of these outbreaks using different criteria which fall in the identification of health risk. We study the possibility of integration of multi-criteria analysis methods that are formal methods that have proven their efficiency in space and have demonstrated their ability to identify spatial problems. These methods have been applied in different studies conducted in our team for a decade and in different fields (Transport, Planning territory, Production Management and Industrial domain) where we come the idea of designing a spatial decision support system based on the integration of SOLAP technology and Multi criteria analysis methods whose objective is to identify the epidemiological spread phenomenon and make it controlled problem consequently increased and effective epidemiological surveillance.

- The role of **SOLAP** would be the location of outbreaks of epidemics.
- The role of **MCAM** is the classification of disease outbreaks to facilitate intervention in order of priority.

Our objective is to model the problem of epidemiological surveillance to a Multi criteria problem taking into account the various criteria affecting the spread of the disease.

6. MULTI CRITERIA FORMULATION PROBLEM

The proposed decisional model based on Multi Criteria Decision support is largely inspired from that proposed in [8]

6.1 The structuring phase

This first phase aims to identify the problem (geographical location of study area using the GIS, identification of different criteria) and the basic choices on how to approach it. This phase aims, also, to formalize two basic elements of the decisional situation:

- **Identify actions:** the identification of all the potential actions is a very significant step in any decision support approach, especially when the multi-criteria analysis method proceeds by partial aggregation. It is very important that the set of all the actions is complete because its modification during the analysis can cause a recurrence of multi-criteria analysis.

- **Identify criteria:** the list of criteria obtained by aggregating the corresponding factors (sub-criteria) should be as complete as possible. These criteria must be related to constraints and objectives used in the generation activities. The family of the most relevant criteria must verify the conditions of exhaustively, consistency and independence.

6.2 The operational phase

This second phase is the analytical process of the study. Its two main objectives are the evaluation criteria, then the aggregation of this information by a multi-criteria analysis exploiting the multi-criteria methods of classification namely (PROMETHEE family methods).

6.3 The implementation phase

This third phase is primarily the result of social acceptance. However, it also includes the implementation of the decision and the control of this implementation. The main phases and stages of the proposed model are designed in Figure 3.

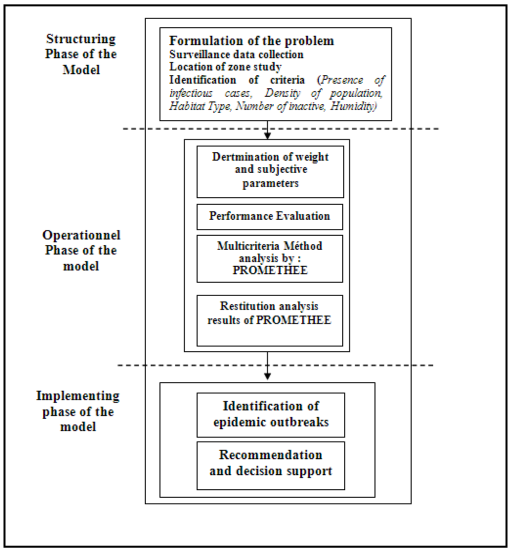

Figure 3: The proposed decisional model [4]

The functional architecture of the proposed Decision Support System based on MCAM is illustrated on the flow chart of figure 4.

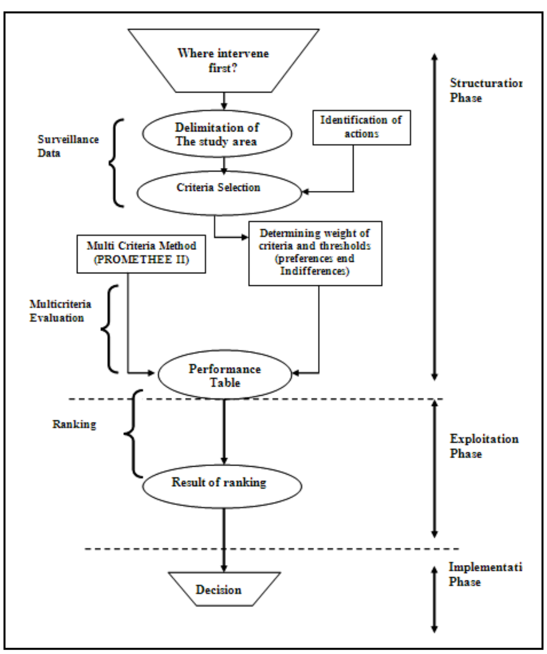

Figure 4 Diagram of Use of our proposed Decision Support System

7. PROMETHEE MULTI CRITERIA METHOD

The multi-criteria analysis method PROMETHEE (Preference Ranking Organisation Method for Enrichment Evaluations) [11] treats a γ problem. The main advantages of PROMETHEE are:

- The simplicity and intuitive aspect of the method
- The power of its preference function
- The simplicity of the operating phase of the method

The method PROMETHEE I provides the user with a ranking of different actions (outbreaks). The problem is that this method does not classify all actions. Some actions may remain unmatched. PROMETHEE II method allows removing this incomparability. The principle of this method is to establish a numerical process of comparing each action relative to all other actions. Thus, it is possible to calculate more (merit) or less (demerit) of each action compared to all others. The result of this comparison allows the orderly classification of actions. The implementation of the method can be reduced to perform the following three steps:

7.1 Choice of generalized criteria each criterion C1, C2 ... Cm be associated with a generalized criterion chosen based on a preference function and scale effects are eliminated.

7.2 Determination of an outranking relation in a second phase, it is necessary to determine a relationship outranking through a preference index that quantifies the preferences of the decision maker. The preference intensity is calculated as follows:

p (d) = 0 if d ≤qj, p(d) = (d-qj) / (pj, qj) if qj <d ≤ pj and p (d) = 1 if not "a" and "b" are two actions potential, "d" is the difference between the performance of "a" and performance of "b" (gj (a) - gj (b)). qj is the indifference threshold, and pj is the preference threshold.

7.3 Assessment of preferences the evaluation of the preference of the decision maker is ensured by the inclusion of incoming and outgoing flows.

- **Calculating the preference indicator**

$$\pi(a, b) = \Sigma W_j * P_j(a, b) / \Sigma W_j$$

Wj is the weight of criterion j

- **Calculation of incoming flows**

ϕ+ (a) =Σπ (a, x)

- **Calculation of outgoing flows**

ϕ- (a) =Σπ (x, a)

- **Calculation of global flows**

ϕ (a) =ϕ+ (a) -ϕ- (a)

The general principle of operation by PROMETHEE (I and II) is given by the flowchart illustrated by the figure 5.

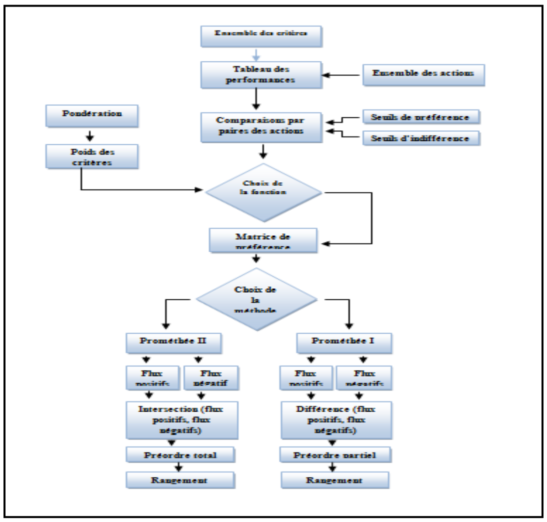

Figure 5 PROMETHEE Method Process

8. CASE STUDY

Due to demographic and socio-economic situation in Algeria, further progress is still needed to achieve the objectives of the recovery plan for the fight against tuberculosis (2006-2015) that are part of the Millennium Development Goals (MDG), and the new strategy "Stop TB" recommended by World Health Organization since 2006: *"Stop the increase in the incidence of tuberculosis and begin to reduce throughout the national territory"* [1].

Among the main measures adopted by this national program is to improve the reporting of cases of tuberculosis and their monitoring by the generalization of electronic surveillance system. Our project is integrated in this context to assist in achieving these goals by developing a spatial decision support System (SDSS) capable of carrying help in epidemiological surveillance to identify the problem of the spread of tuberculosis which is an uncontrollable kind problem.

In this regard, we note the little comparative studies that can help to determine what extent the socio-economic transformations that know the study area and environmental scenarios helped in the spread and transmission of the disease.

8.1 Identification of the study area

The field in the context of this study is the region of Oran in Algeria. The establishment of a geo-location of residential areas of population said "poor" most exposed would observe places where epidemic could spread rapidly and most widely. Professor Bouziani, epidemiologist at Bio-statistics department in the Faculty of Medicine INESSMO of Oran University, we were oriented towards the disease of tuberculosis, which still represents a continuing hazard lowering the population in the region of Oran[2] .

8.2 Considered criteria

One of the proposed solutions in this study to increase case detection was to identify populations at risk for Tuberculosis. The objective of our research is precisely to understand the dynamics of tuberculosis in Oran depending on the criteria mentioned in section 5, above to create an index of vulnerability of population living in the region, and achieve in producing a risk map of tuberculosis.

Risk is assessed according to the frequency of occurrence of one or more parameters about the case. Risk is defined by the simultaneous presence of the hazard and the vulnerability of the population

Thus, it is possible to create a risk index based on the parameters that identify the presence of the hazard (e.g. the presence of infectious cases ...) and identifiers of the population vulnerability (e.g. the density (overcrowding), location, type of housing, unemployment ...)

8.2.1 Presence of infectious cases (Medical Criteria)

The risk of Tuberculosis is defined in three classes according to the presence of infectious cases (proven cases) and the presence of cases of non-infectious (unproven) but potentially infectious tuberculosis. The hazard is so strong in the foyer in question, there is at least one infectious case; danger is low if there are no cases of tuberculosis detected as contagious but there is at least one case of tuberculosis identified in the outbreak; hazard is zero if no cases of tuberculosis have been identified throughout the year (probability of the event "with a case of tuberculosis" = 0).

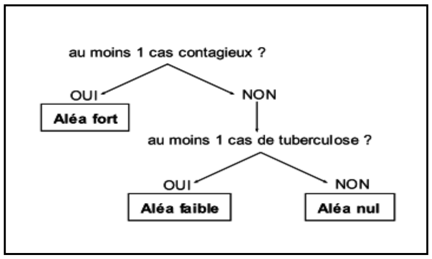

Figure 6: Different classes of risk of tuberculosis

8.2.2 Density (Demographic Criteria)

There is a little information in the literature about the "social vulnerability" and socio-economic data that may be related to the presence of tuberculosis. According to the availability of these vulnerable settings, our choice was focused on the "population density" parameter that appears the most significant. Indeed, overcrowding in a given region favors contamination and the spread of epidemics including tuberculosis enough direct contact with the patient with the disease carry the virus by air voice. A denser population is more vulnerable to be contaminated by the presence of the hazard course here is the presence of infectious cases.

8.2.3 Habitat Type (Demographic Criteria)

This criterion is calculated by the number of precarious constructions generally occupying shantytowns:

8.2.4 Number of inactive (Demographic criteria)

The number of inactive is resumed in our study on the number of unemployed people in the regions.

8.2.5 Humidity (Climatic Criteria)

Climate humidity is a very important parameter which promotes the spread of bacteria (eg the cock bacillus responsible for the TB disease). being unable to have the annual average humidity of all the regions of Oran (presence of two measure stations only: sénia and arzew) we have identified, through a weather specialist in the National Office Of Weather in Oran, a scale of 4 points (1-2-3-4) that classes the regions from the wettest (measure = 1) to the driest (measure = 4)

8.3 Performances Table

The information layers involved in our Tuberculosis risk model: the hazard, vulnerability: Presence of infectious cases, Population, Density, poverty level and the rate of humidity, will be crossed in EPISOLAP-MINING, showing areas at which interactions between a vulnerable population against Tuberculosis and patients who are likely to transmit the disease (Proven / Unproven) are most intense.

The objective of this multi-criteria modeling of the spread of Tuberculosis disease problem will result in a classification of outbreaks of epidemics from most favorable to least favorable. These areas are supposed to be most at risk, or the most "hazardous epidemiologically" that will lead to the development of a risk map. Actions in our Multi Criteria Analysis study are the foci of epidemics detected previously by the "EPISOLAP" system (26 outbreaks considered).

Table 1: Performances Table

Outbreaks	***Hazard***	***density***	***Nb of precarious Const.***	***Nb of innactifs***	***Humidity***
ORAN	723	767558	5125	9928	1
GDYEL	53	36314	101	760	2
BIR EL DJIR	114	88398	1305	2683	2
HASSI BOUNIF	60	54046	116	1356	3
ES SENIA	132	78433	570	1345	2
ARZEW	79	80761	120	788	1
BETHIOUA	3	17840	53	342	1
MARSAT EL HADJADJ	0	12449	44	304	1
AIN TURK	51	31776	197	671	1
EL ANCAR	14	9597	82	112	1
OUED TLELAT	24	16086	78	225	4
TAFRAOUI	5	12090	5	289	4

Outbreaks	***Hazard***	***density***	***Nb of precarious Const.***	***Nb of innactifs***	***Humidity***
ORAN	723	767558	5125	9928	1
SIDI CHAMI	103	71243	1282	2324	2
BOUFATIS	7	11991	11	189	2
MERS EL KEBIR	10	17149	73	357	1
BOUSFER	15	13480	241	401	1
EL KARMA	36	16507	255	422	2
EL BRAYA	8	4696	15	206	3
HASSI BEN OKBA	13	11421	51	297	3
BEN FREHA	34	17631	75	446	3
HASSI MEFSOUKH	18	9267	113	346	3
SIDI BEN YABKA	8	7133	44	116	3
MESSERGHIN	27	21896	286	629	3
BOUTLELIS	22	21303	65	564	3
AIN KERMA	11	8449	0	122	1
AIN EL BIYA	3	31778	5	395	2

8.4 Weight criteria

In the frequent case where the analysis of the consequences of potential actions led to build several criteria is multi criteria analysis to provide answers to the problem.

For any given action, and for each criterion a preference threshold p, indifference q and veto threshold v are estimated, knowing that PROMETHEE methods do not exploit the veto threshold who is a subjective parameter expressed by the decision maker. Each criterion is assigned a weight k reflecting its contribution to the final decision. The result of the analysis of the consequences is presented in a table of performance. For simplification reasons we have chosen that: p (PREF) =q (INDI) =1

Table 2: Table of subjective parameters

	Cr1	**Cr2**	**Cr3**	**Cr4**	**Cr5**
weight	3	2	1,5	2	1,5
PREF	1	1	1	1	1
INDI	1	1	1	1	1

8.5 Results of ranking

The multi criteria method that was implemented for ranking outbreaks of epidemics on the map of Oran is the method PROMETHEE II witch constructs an outranking relation value, based on the comparison of the actions in pairs; its purpose is to store the actions of the best one to the worst. Figure 7 shows a risk map which gives a classification of outbreaks of epidemics from more epidemiological risk region at the least epidemiological risk region using a legend. The result of ranking was visualized on a map using the Map Info Professional 11.0 GIS.

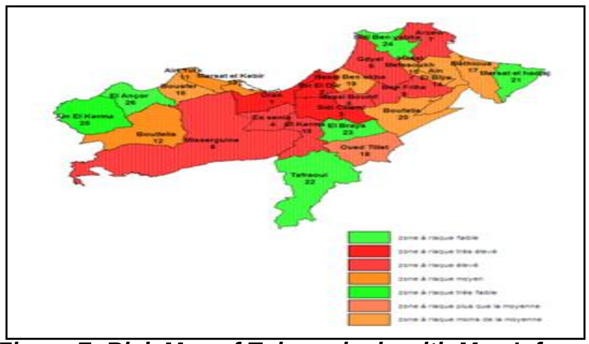

Figure 7: Risk Map of Tuberculosis with Map Info Professional 11.0

9 CONCLUSION AND PERSPECTIVES

We were able, through this case study, see how our proposed model in epidemiological surveillance "EPISOLAP-MINING and its effectiveness in terms of the provision makers of public health relevant information enabling them to see relationships between phenomena, while encouraging them to discover knowledge and produce the right decision by acting effectively in time and in space. This prediction process generated by the integration of SOLAP technology and multi criteria decision methods used in our approach makes the originality of our contribution and presents to us the purpose of our proposed model.

Even if deficiencies remain supplement to achieve our goals of departure; we have less initiated in the present paper a predictive model using multi criteria analysis to develop a risk map that helps decision makers in public health to take the necessary devices to avoid a health risk. Through this study and our future work, we try to integrate the data mining techniques to our system EPISOLAP-MINING, focusing on the creation of spatial operators that lack in our Spatial Decision Support system.

Although limitations appear from first glance, before the major achievement of such challenges, we have initiated in this paper the first nucleus of this new approach itself constitutes a first validation of our spatial decision support systems for monitoring epidemiological EPISOLAP-MINING

10 REFERENCES

[1] Agadir, F., Ali Pacha, S., Anane, T., Baough, L., Benkara, A., Boulahbal, F., Chaulet, P., Halassa, S.A., Hannoun, D., Haouichat, H., L'Hadj, M., Larbaoui, D., Nafti, S., Ouagani, D., Yala, D., Zidouni, N. (2011). Manuel de la lutte antituberculeuse à l'usage des personnels médicaux

[2] Bouziani M. Les pathologies infectieuses : Aspects épidémiologiques et prophylactique ; 2000.

[3] F.Joerin, Décider sur le territoire: proposition d'une approche par utilisation de SIG et de méthodes d'analyse multicritère, Thèse de doctorat, Université de Lausanne, 1997.

[4] Guernier Vanina, Combiner analyse spatiale et épidémiologie pour l'aide à la décision dans la lutte contre la tuberculose en Guyane française. These de doctorat de l'université d'Orléans.Disipline:Environnement et Santé soutenue publiquement le 19 janvier 2006

[5] V. Mottier, Un composant logiciel pour les systèmes informatisés de gestion des réseaux d'assainissements, Thèse de doctorat, Ecole polytechnique fédérale de Lausanne, Suisse, 1999.

[6] C.Boulemia, G.Henry et O.Pecqueur, Eléments de proposition à la mise en place d'une base de données urbaines dans une collectivité locale, Congrès de l'AUGC, Lyon,2000.

[7] A.Bensaid, M. Barki, O. Talbi, K. Benhanifia, A. Mendas, L'analyse multicritère comme outil d'aide à la décision pour la localisation spatiale des zones à forte pression anthropique: le cas du département de Naˆama en Algérie, Revue Télédétection , vol. 7, n◦ 1, pp. 359-371, 2007.

[8] Zemri, F.A., Hamdadou, D., Bouamrane, K. Towards a Spatio-temporal Interactive Decision Support System for Epidemiological Monitoring. Coupling SOLAP and Datawarehouse. Dans le procedding du Colloque sur l'Optimisation et les Systèmes d'Information, COSI'2012, Tlemcen, Algérie, 12-15 Mai 2012.

[9] Hamdadou, D. and Libourel, T. (2009) 'Couplage approche multicritère et négociation pour l'aide à la décision en aménagement du territoire', SAGEO 2009, Spatial Analysis and GEOmatics – Colloque de Géomatique et d'Information Spatiale, 25, 26 et 27 Novembre 2009, Paris.

[10] K. Bouamrane, D. Hamdadou, Kh. Yachba, Kh. Guenachi Towards a decision support system, application: itinerary road modification for road transport of hazardous materials

[11] D.Hamdadou, K.Bouamrane, A .Merouane. Un système interactif d'aide multicritères à la décision pour le diagnostic industriel International Conference on engineering of Industrial Safety and Environment "ICISE'14", January 26th and 27th, 2014. Int. J. Decision Sciences, Risk and Management, Vol. 4, Nos. 3/4, 2012

[12] Brans, J.P. and Vincke, Ph. (1985) 'A preference ranking organization method: the PROMETHEE method', Management Science, Vol. 31, No. 6, pp.647–656.

[13] M.A. Hamadouche,M.Khaladi, A.Khaldi,Y.Fekkir, D.Anteur, M.Driss, SIG, Télédétection et analyse multicritère : vers un outil de gestion et de préservation de la biodiversité du parc national de l'Ahaggar (Alg´erie), la sixième session du Congrès International.

[14] D. Hamdadou, K. Labed et B. Beldjilali, Proposal for a Decision-making process in Regional planning: GIS, Multicriterion Approach, Choquet's Integral and Genetic Algorithms, ERIMA07, European Research on Innovation Management Alliance, Biarritz, France, pp. 51-59, 2007.