Implementation of a Geographic Information System coupled with Spatial statistical analysis For demography and epidemiology

El Hadj Hayet, Kerboua Ziari Yasmina, Mejden Kherchi Hania

Abstract— Spatial statistics and Geographic Information Systems (GIS) are widely used in various Domains (demography, social and economic data etc.) and serve as decision-making tool. For this purpose, we tried through this work to join an information system into the statistical data relative to the population in Algeria and epidemic of cutaneous leishmaniasis which takes the most important part in terms of incidences of zoonoses in Algeria, to be able to do a statistical and cartographic analysis for data.

Index Terms— Epidemiology, cutaneous leishmaniasis ,decision-making, geographic information system,spatial statistics,simulation,.

1 Introduction

An Information system (IS) is an organized set of resources (materials, software, persons, data and procedures) which allows grouping, classifying, processing and broadcasting information about a given phenomen[2]
A Geographic Information system (GIS) as defined by David COWEN, is a decision support system which places georeferenced data in a context of problems resolution [8],
Back in 1854, when John Snow [7], the father of the modern Epidemiology determined the geographical distribution of the deaths due to the Cholera by demonstrating the relation between the cases of this epidemic appeared in London in the district of Soho and the well contaminated by waste water of Broad Streett
In this way, he related forever the new science of epidemiology with the use of geographical information in order to determine the multiple relations between the environment and the epidemic.

The objective of this work is determined by:

1- Introducing the spatial approach in the field of health through technology of GIS and spatial statistics;
2- Description of the geographic distribution of the population and the notifiable diseases in Algeria through statistical and spatial data analysis;

2 Situation of the health system and the sanitary information system in Algeria2.1 Review Stage

The introduction of the free access of the care [10] in Algeria from 1974, in the whole of the structures of public health, as well as the launch of the various programs of health led to a strong improvement of accessibility to the care. The part of the spending in health in the gross domestic product is estimated at 5,79 %, although it is important, it remains low with regard to the other countries having a similar economic development and industrialized countries.

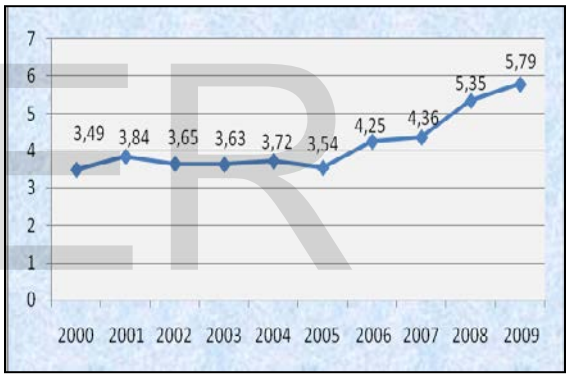

Fig1. Evolution of the part of the spending in health [9]

Despite efforts approved in terms of monitoring and information gathering by health organizations in Algeria, the health information system is failing, health information flows poorly, which will prevent a better treatment of various health information, their transmission between practitioners and the appropriate action to health problems.
Diseases for which efforts must be maintained by Algeria are represented by [13]:

- Diseases under the expanded program of immunization
- Diseases related to environmental health
- Tuberculosis
- Some zoonoses
- Sexually Transmitted Diseases
- Viral hepatitis B and C.

- *El Hadj Hayat Email:. elhadj.hayat@yahoo.fr*
- *Yasmina Kerboua Ziari*: yasminaziari@yahoo.fr*
- *Hania Mejden Kherch: kherchi_inps@yahoo.fr*

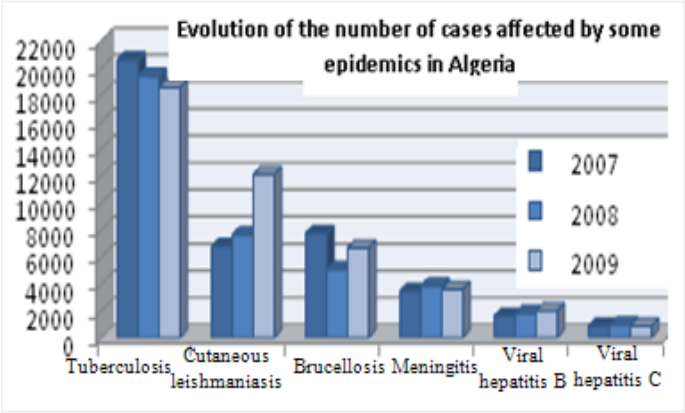

Fig2. Comparison between the incidence of Cutaneous Leishmaniasis and some other epidemics

The figure above shows the evolution of the incidence of Cutaneous leishmaniasis

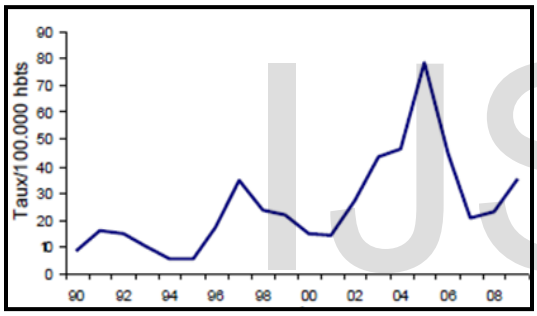

Fig3. Evolution of annuel incidence 1990-2009

3 Cartographic analysis [6] of Distribution of the population of Algeria and incidence of cutaneous leishmaniasis

The main functions of Geographic Information System are:

- The Acquisition: it is the capture of geographical data in digital form,
- The Archiving is the system of recording of the information on the Territory to transfer the data of the workspace towards the space of archiving (hard disk)
- The Analysis which is made on the geographical data includes their manipulation (modelling, statistics, etc.) to answer questions we ask ourselves.
- The Posting is the shaping by producing maps in an automatic way, to perceive the spatial relations between objects and visualize the data on the screens of computers.
- The Abstraction means conceiving a model which organizes the

Treated information concerns both statistical data (like Census of population) and field data

3.1 Geographic distribution of Algerian population [11]:

The Algerian population, estimated at 34075038 people according to the 2008 census, is distributed unequally between regions; about 63% in the north, 27% in the highlands, and only 10% in the South

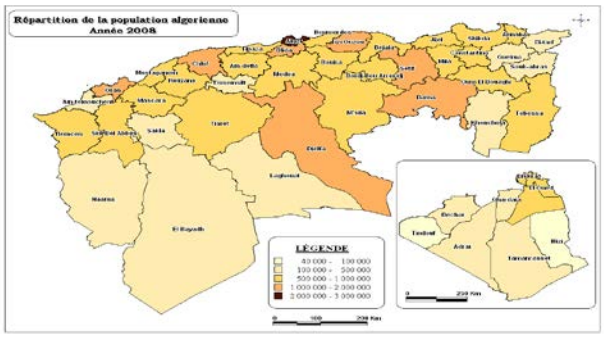

Fig4. Geographic distribution of Algerian population
Source: data of the National Statistics Office 2008

3.2 Distribution of cutaneous leishmaniasis in Algeria [12]

Regarding the notifiable diseases, Zoonosis takes the first position in the statements recorded. C**utaneous leishmaniasis** takes the most important part in terms of incidences , the number of cases was 6,755 in 2007 and 7,632 in 2008 compared to 12,097 registered in 2009, the epidemic appears to be strongly in the provinces of the interior region.

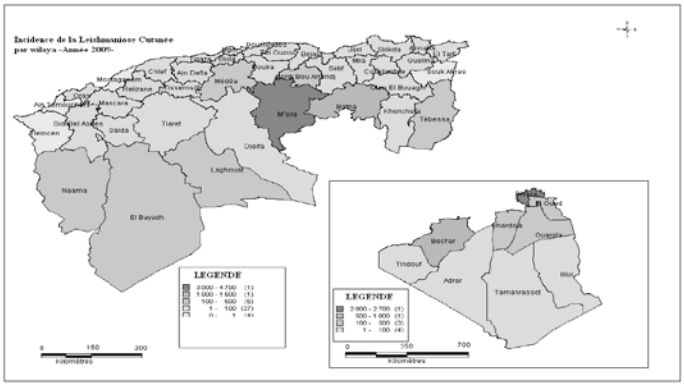

Fig5. Geographic distribution of Cutaneous Leishmaniasis incidence
Source: data of the institute National of public health 2009

This vectorial disease of epidemic nature is of a big expansion in these regions because of the insecurity and the precariousness in the agricultural programs and of management of the water which do not integrate the health dimension [5]. The lack of hygiene, the proliferation of rats, insects and the stray dogs are the main vectors of its transmission. Leishmaniose Cutanée requires a very expensive care

and causes visible aftereffects and scars even after the cure.
In spite of the efforts made by the Algerian public authorities by launching campaigns of prevention and fighting against this disease, the rate of its incidence does not stop increasing and the risk of its distribution towards others region of the North remains to be afraid. For that purpose, it is indispensable to set up more measures in order to slow down its distribution

4. Statistical analysis of spatial data

Spatial statistics[1], [3] , is a branch of statistics that focuses on the specific distribution of a phenomenon observed in a given territory (distribution of a population, epidemics, pollution... etc.). Its central concept is the spatial autocorrelation[4], which is initially the idea of Moran (1948) and measure essentially the resemblance between nearby location in space (dependence between the attributes of an individual and the attributes of the other individuals placed in the same two-dimensional space)
We can distinguish one of three following cases:

- Positive spatial autocorrelation: two neighboring places are alike more than two remote places
- Negative spatial autocorrelation: two neighboring places are alike unless two remote locations
- Absence of spatial autocorrelation: no spatial organization

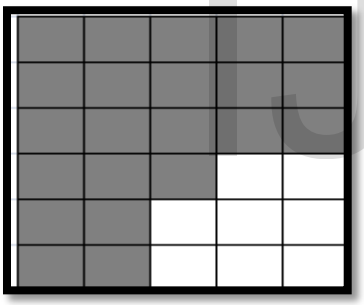

Fig 5.1 Positive spatial autocorrelation

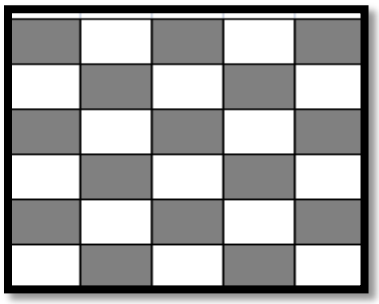

Fig 5.2 Negative spatial autocorrelation

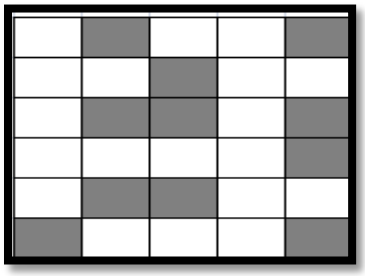

Fig 5.3 Absence of spatial autocorrelation

4.1 Adjacency matrix : it is defined as W=$w_{ij,}$ the element w_{ij} is equal to 1 if to cells are joined, and 0 otherwise. W is a symmetric and binary matrix**.**

4.2 Moran's Index: This index is the ratio of covariance of nearby locations and the global variance of the studied character

$$\boldsymbol{I} = \frac{\mathbf{1}}{\boldsymbol{p}} \frac{\sum_i \sum_j \boldsymbol{w_{ij}}(\boldsymbol{z_i} - \bar{\boldsymbol{z}})(\boldsymbol{z_j} - \bar{\boldsymbol{z}})}{\sum_i (\boldsymbol{z_i} - \bar{\boldsymbol{z}})^2} \qquad (1)$$

$$= \frac{Covarianceofthenearbylocationsinspace}{Globalvarianceofthestudiedcharacter}$$

Where : $\boldsymbol{p} = \sum_i \sum_j \frac{\boldsymbol{w_{ij}}}{\boldsymbol{n}}$ (2)

n: number of spatial units
and **z** the observed variable.

- If I=0: Absence of spatial autocorrelation, the covariance between adjoining places is hopeless (the neighborhood plays no role in the values of the character);
- if I>0: spatial Autocorrelation is positive,
- if I < 0: negative spatial Autocorrelation.

If the observations z are spatially independent random variable, the expected value of Moran's I is

$$E(I) = -\frac{1}{n-1} \qquad (3)$$

For large n the expected value is approximately 0.
The variance of Moran's I is defined as

$$Var\ (I) = \frac{n^2(n-1)S_1 - n(n-1)S_2 - 2S_0^2}{(n+1)(n-1)S_0^2} \quad (4)$$

Where:

$$S_1 = \frac{1}{2}\sum_i \sum_j \left(w_{ij} + w_{ij}\right)^2, i \neq j \qquad (5)$$

$$S_2 = \sum_i (\sum_j w_{ij} + \sum_i w_{ji})^2, \qquad (6)$$

And

$$S_0 = \sum_i \sum_j w_{ij}, i \neq j \quad (7)$$

As demonstrated in this document, the numbering for sections upper case Arabic numerals, then upper case Arabic numerals, separated by periods. Initial paragraphs after the section title are not indented. Only the initial, introductory paragraph has a drop cap.

4.3RESULTS

Moran Index	Calculated value of the test
0,21	2,41

The calculated value of moran's index is 0,21 which indicates positive autocorrelation. To confirme this result, Significance test used by comparing the Calculated values tests with the tabulate value of the normal distribution (equal to 1,96),allows us to reject the assumption according to which the studied variable is randomly distributed in the space. Consequently, we confirm the existence of a positive spatial autocorrelation, so that Algerian population is concentrated in certain regions.

6 CONCLUSION

Geographic Information System has been extended extensively to different fields, actually, it is considered as real tool of surveillance and decision support. It is an important tool in the management and monitoring of public health which requires great care by government.

The continuous increase of the incidence of the C**utaneous leishmaniasis** and the risk of its distribution towards other regions incited us to analyse its geographic distribution We noticed that its incidence is very strong in the area of the interior (highlands) which are farming areas because of insecurity in the agricultural programs there .

To look for the link between the populated regions and those affected by this epidemic we also made a cartographic analysis on the distribution of the Algerian population.

Thanks to the tools of **spatial statistics** which has a close relationship with GIS, we were able to calculate the indix of spatial autocorrelation of Algerien population: Moran's index (**0.21**),wich confirm that spatial autocorrelation is positive indicating a dependency between neighboring geographic units, it means that the Algerian population is concentrated in the Northern region by forming a group and leaving the rest of the regions of low enrollment.

ACKNOWLEDGMENT

The authors wish to thank the Algerian National Statistics Office.

REFERENCES

[1] Carlo Gaetan , Xavier Guyon: « Modélisation et statistique spatiales » Edition Springer, Paris 2008
[2] Jean-Claude Courbon : « Système d'information : structuration, modélisation et communication », Edition Intereditions, Paris1993
[3] Jean - Jacques Droesbeke , Michel Ejeune, Gilbert Saporta : « Analyse statistique des données spatiales », Edition Technip, Paris 2006,
[4] Jean-Marc-Zaninetti : « Statistique Spatiale : méthodes et applications géomatiques »,Hermès Science Publications, 2005
[5] Houria Zait , Bousaad Hamriou Leishmanioses cutanées en Algérie, 2007
[6] LARRY DANIEL : « Inside MapInfo Professional», Onword press, Canada 2002
[7] Lloyd F. Novick, Cynthia B.Morrow, Glen P.Mays (2008): Public health administration , Jones and Bartlett Publishers, United States of America.
[8] Martin Paegelow (1995): cours Initiation aux systèmes d'information géographique , University of Toulouse le MIRAL ,UFR de géographie.
[9] [Ministry of Health and Hospital Reform (2006) Master Plan for Health 2006, Algeria.
[10] Mustapha Bouziani, (2009): «Regard sur la santé publique en Algérie». Scientific report.
[11] National Office of Statistics: census of population 2008
[12] National Institute of Public Health: "Epidemiological monthly statement", 2009
[13] Rapport National sur le développement humain ». Conseil National Economique et Social, 2006