A NOVEL METHODOLOGY FOR GEOVISUALIZING EPIDEMIOLOGICAL DATA

Fatiha Guerroudji Meddah

Département d'Informatique, Université des Sciences et de la Technologie d'Oran Mohamed Boudiaf, USTO-MB
BP 1505 El M'naouer 31000, Oran Algérie
fguerroudji@Yahoo.fr

Abstract: Rapid advances in geographic information systems (GIS) and related technologies have created a potential for dynamic geovisualization methods to be integrated with GIS in support of a range of decision-making tasks. Cartographic visualization is then considered as an extension of spatial analysis, and GIS is configured as a spatial decision support system. This integration can add a great deal to epidemiologic research and is essential for health policy planning, decision making, and ongoing surveillance efforts. In this context, this paper presents a novel methodology to support interactive visual exploration and analysis of epidemiological data by coupling MapInfo GIS software with the Gastner area cartogram, a particular class of map type where some aspect of the geometry of the map is modified to accommodate the studied problem. The aim of the study is to help improve public health by identifying areas of exposure and risk on tuberculosis in the city of Oran in Algeria.

Keywords: *GIS, Geovisualization, Cartograms, Tuberculosis, Epidemiology*

1. Introduction

Tuberculosis (TB) remains among the 10 leading causes of death in the world and is a public health priority in Algeria (23 000 cases in 2018) (www.aps.dz 2021). Today it is indisputable that Tuberculosis is the subject of potential studies by the medical world and particularly by epidemiologists, whose primary objective is to find solutions through the analysis of statistical data. Identifying heterogeneity in the spatial distribution of TB cases and characterising its drivers can help to inform targeted public health responses, making it an attractive approach.

However, common diseases such as tuberculosis are greatly impacted by geographical and environmental factors, we can contribute to improving public health with geovisualization solutions by identifying areas of exposure and risk, by providing relevant interpretable visual informations essential for decision making.

To geovisualize (TB) epidemiological data, we propose in this paper a methodological approach integrating GIS and anamorphic maps: cartograms. The geographic information system (GIS) is an effective tool for the organization of diseases and health data. Cartograms are maps in which the real relationships of enumeration units are distorted based on a data attribute (Slocum and others 2009) (Field 2017). Cartograms are of two types: area cartogram (Dorling 2005, 2011) and linear cartogram (Thomas 2018) depending upon the geographical feature being distorted.

Cartograms were used mainly for representing population density (Doll 2017) and electoral votes (Dominique 2005). They are employed to simultaneously convey two types of information: geographical and statistical.

However, in literature (Bhatt and others 13; Derryn 2014; Nusrat 2016; Loes Soetens 2017; Tran 2019) Cartograms are also innovative mapping techniques that allow visualization of potentially complex health relationships but are underutilized in epidemiology. As shown in (Daniel Sui 2008), it is obvious that the use of cartograms in public health can affect our understanding of reality, both cognitively and analytically.

In this context, to facilitate public health intervention, to design new tuberculosis (TB) control strategies, and to identify when TB is transmitted in Oran, the main objective of this research, is to produce epidemiological cartograms in a form adapted to the perceived reality. In order to achieve this objective, the proposed approach was defined on a mathematical model based on Gastner Newman's algorithm and Bertin's graphic semiology.

The paper proceeds with four more sections. The following section, describes a set of algorithms to construct cartogram. The third section provides methodology for producing cartogram and discusses important design consideration. Concluding remarks and future directions are offered in the final section.

2. Preliminaries :

Several algorithms have been proposed in the literature to build cartograms:

Tobler's Algorithm (Tobler 73, 86, 04), which was based on placing outlines on surface of variable or attribute being mapped, and repeatedly flatten the distribution based on regular sampling of the surface. But this method is not so successful as it produced poor shape of the county boundaries of US states after 99 iterations. Dorling's Algorithm (Dorling 93), replaces geometric features with circles. The size or radius of circle is proportional to the magnitude of the attribute being represented. This is for non-contiguous creating cartogram. Dougenik's Algorithm (Dougenik 85), was based on the gravitation concept. The polygons are shrunken and expanded at centroid to create cartogram. With iterations (approximately 8), the polygons converge much more rapidly. Other algorithms based on some order in distorting the area of polygons. Small area to large area or small attribute value to large attribute value. The disadvantage of these algorithms is that none of them preserved the shape of shape of the geometric areas. Problems like overlapping regions, lack of readability, coordinate axes dependence were still persistent.

As a solution to these problems, Gastner and Newman (Gastner 2004) proposed a simple algorithm based on elementary physics. It is called diffusion algorithm and cartogram produced is called diffusion cartogram. In this algorithm, the original input map is projected onto a distorted grid, computed in such a way that the areas of the countries match the pre-defined values. Cartogram is created by allowing the flow of data from high-density area to low-density area until the density is equalized everywhere. They express the problem as an iterative diffusion process, where quantities flow from one country to another until a balanced distribution is reached, i.e., the density is the same every where. This method allows for minimal cartographic error, while also keeping regions hapes relatively recognizable. Over the last decade, this has become one of the most popular techniques to create cartograms. Its popularity is likely due to its shape recognizability, and the availability of the software to generate these cartograms (Nusrat 2016).

3. Proposed Methodology

In this work, we propose a new visualization methodology, by integrating MapInfo GIS and cartograms to assist the analysis and the exploration of tuberculosis data (Pulmonary data) set collected by local health services DSP (Direction de la Santé Publique) in Oran, the second largest city in Algeria, covering a total area of 2,121 km^2 with 26 municipality and a specific region named “Grande Sebkha”. “DSP” in Oran collates a large amount of data generated from each interaction a patient makes with the health service. This data set includes clinical information that can be used to analyse geographical trends of tuberculosis based on a history of four successive years (2014 to 2018).

The proposed method is a decision support tool to facilitate public health interventions and design new strategies to fight against tuberculosis (TB). The geographic information system (GIS) will make it possible on the one hand to organize the data of the (TB) to identify and control the evolution of this one by taking appropriate measures with the discovery of spatial accumulation of the disease. On the other hand, cartograms are used to simultaneously convey the two types of resulting information: geographic and statistical. The visualization methodology is composed of successive steps as shown in the (Figure.1).

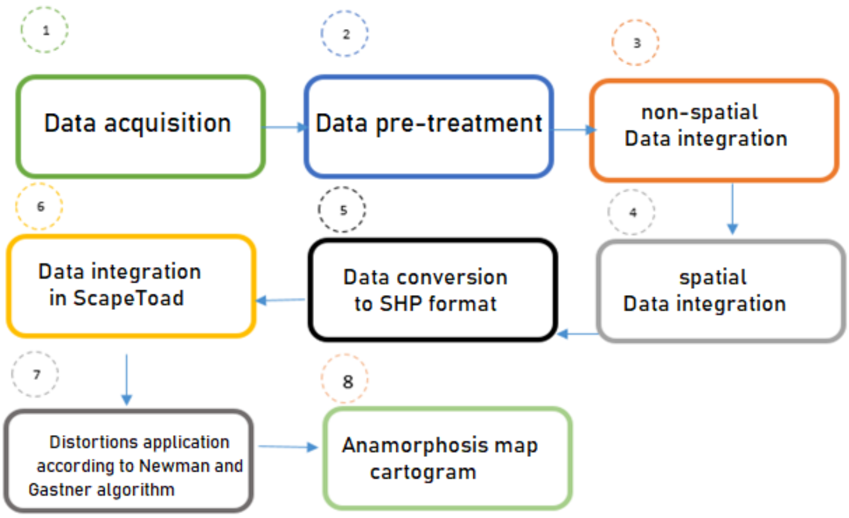

Figure 1: Proposed Methodology

To build cartograms, we have needed not only to bring together the different types of geographic data in a database, but also convert it into a form suitable for thematic interpretation. After the pre-processing phase, the recovered data were structured in the form of geographic databases (spatial and non-spatial data) that we integrated into the MapInfo GIS. The result is a map that contains a set of polygons where each polygon represents a municipality described by a set of data stored in the database. An example of this map for the cases observed in 2014 is shown in Figure 2.

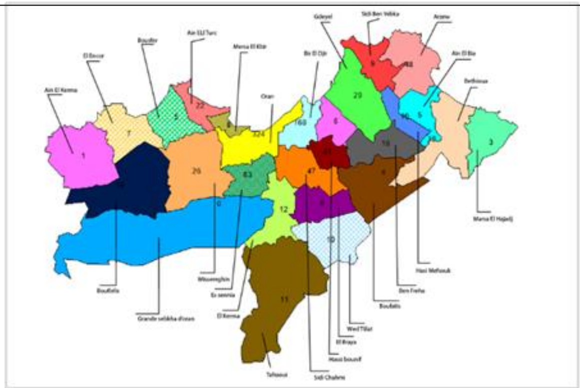

Figure 2: Example of TB cases for year 2014

The data set stored in a MapInfo database were then converted to SHP file format to support the deformations necessary for the construction of the cartogram by the ScapeToad tool. The deformation of the map is done in four steps. A computational grid is superimposed on the polygon layer. The value of the studied variable is calculated for each point of the Grid. This is a "rasterization" step. The grid is then distorted by the Gastner-Newman algorithm Through a second, finer grid (grid of diffusion), each cell is enlarged or shrunk in such a way whether the density (value of the variable / area of the cell) is the same for all cells. In our case Three (03) iterations was necessary to perform this deformation.

4. Results:

The outcomes of the procedure outlined above are given at Figure 3 which offers a series of cartograms of the number of cases of tuberculosis in the city of Oran during the period from 2014 to 2018. The size of each municipality is proportional to the number of cases reached.

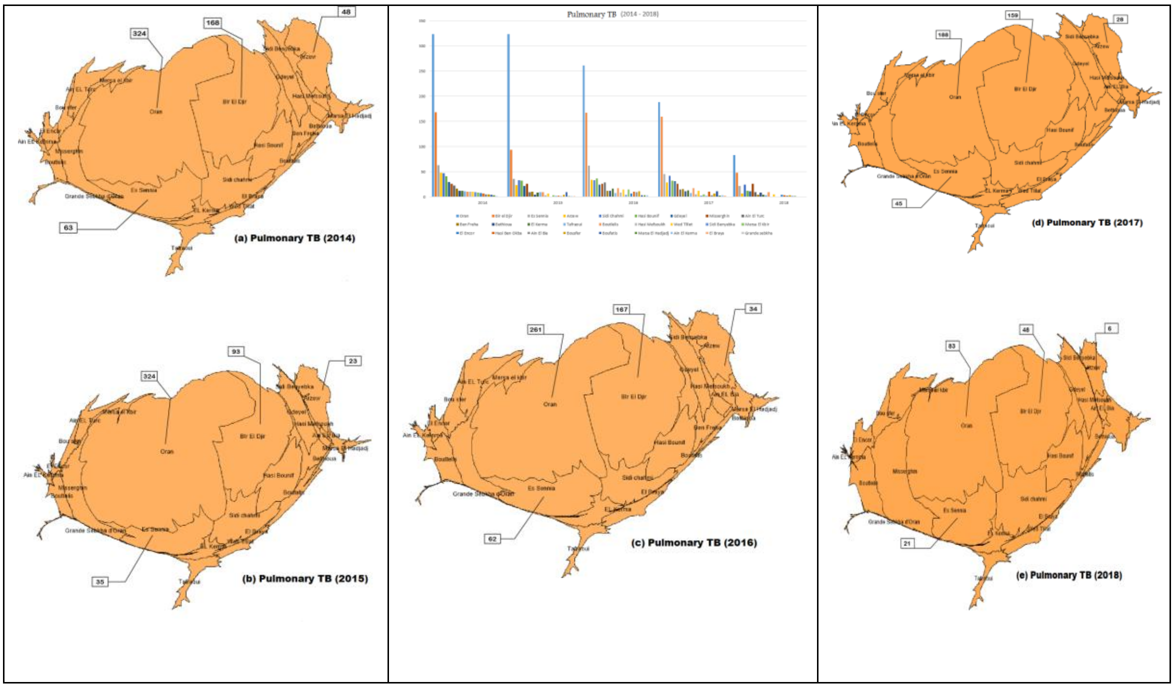

Figure.3: Generated cartograms (a(2014), b(2015), c(2016), d(2017),e(2018))

In view of the cartogram of the year 2014 we note that the number of cases of people with pulmonary tuberculosis in the municipality of Oran is clearly higher than that of other municipalities with 324 cases. The municipality of Bir el Djir being classified second with 168 cases this which justifies the deformation in the cartogram.

In same, the cartogram of 2015 shows an improvement in the number of cases reached. Even if we note that the number of cases recorded in the municipality of Oran has remained stable, the improvement is clearly noted at the level of other municipalities, particularly in the municipality of Bir el Djir which has seen a reduction of more than almost 50% of cases reached, which involved a slight deformation.

As for the year 2016, the deformation of the Bir el Djir area resurfaced with an increase of 45% of cases, despite the stability of the cases observed in the municipality of Oran, the latter still remains in the lead with the most important deformation followed by the municipality of Bir el Djir

In 2017, a slight decrease was observed in the two zones Oran and Bir el Djir, while an increase was reported in all of the other municipalities.

In 2018, the effort to fight this pathology was clearly observed in all the municipalities with significant reductions and very slight deformation.

5. Discussions:

In his seminal work on the semiotics of graphics, Jacques Bertin identified several pre-attentive visual dimensions across which sign vehicles differ, allowing for the theorization of syntactics for graphic sign systems (Bertin 1967, 1983). The original set of fundamental graphical elements, termed retinal or visual variables, included: location, size, grain, orientation, shape, colour hue, and colour value. The absolute quantitative character of a statistical information is translated by the visual variable size. Cartograms use the visual variable size to signify the equalizing variable. In our work, to size we proposed the addition of colour to reinforce the visualization of the information transmitted by the cartogram. The color as a qualitative variable will make it possible to reinforce the qualitative variable size to allow a complete visualization of the data. Results are as follows:

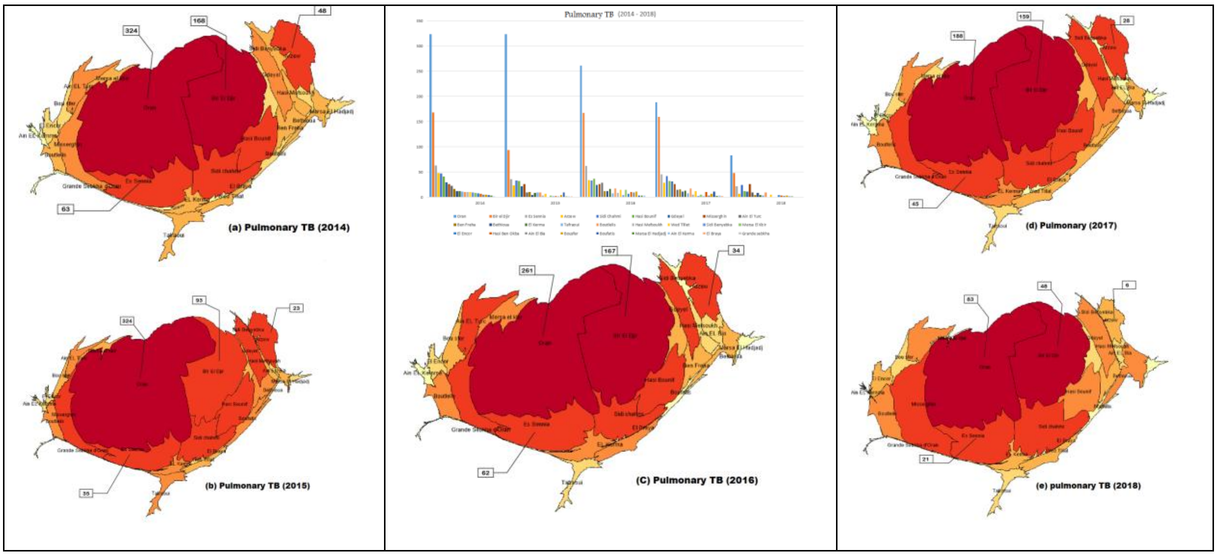

Figure.4: New visualization of (TB) Data based on graphical semiology

These anamorphoses cartograms shown in Figure 4 are more significant than the precedents in terms of reading and interpretation. So, it is very easy to simultaneously make the link between the color and the size of the municipality generated generated in order to understand the change made which corresponds to the evolution of pulmonary tuberculosis.

6. Conclusion

Maps play an important role in geographic communication. They allow large amounts of data to be displayed in parallel and in a format understandable to humans. Cartograms are a special class of map type where some aspect of the map geometry is changed to accommodate the problem. In this work, the proposed a new geovisualization methodology based on GIS and cartograms.

Cartograms are a powerful visual tool, both for communicating ideas and for facilitating data exploration. So, the proposed approach would be a great aid to epidemiologic when the cartogram construction is integrated within GIS for a geovisualization process. It will make it possible to transform the geographic space into a functional space.

7. References:

1. Andrienko, G., Andrienko, N., Dykes, J., Kraak, M.J., Robinson, A., & Schumann, H. 2016. GeoVisual analytics: Interactivity, dynamics, and scale. Cartography and Geographic Information Science, 43(1), 1–2. doi:10.1080/15230406.2016.1095006

2. Andrieu D. 2005 « L'intérêt de l'usage des cartogrammes : l'exemple de la cartographie de l'élection présidentielle française de 2002 », M@ppemonde, n°77, http://mappemonde.mgm.fr/num5/ articles/art05105.html

3. Bhatt S., Gething P. W., Brady O. J., Messina J. P., Farlow A. W., Moyes C. L., Drake J. M., Brownstein J. S., Hoen A. G., Sankoh O., And Others .2013. The Global Distribution And Burden Of Dengue. Nature 496, 7446 (2013), 504–507.

4. Clarke, K. C., McLafferty, S. L., & Tempalski, B. J. 1996. On Epidemiology and Geographic Information Systems: A Review and Discussion of Future Directions. *Emerging Infectious Diseases*, *2*(2), 85-92. https://dx.doi.org/10.3201/eid0202.960202.

5. Çöltekin, A., Bleisch, S., Andrienko, G., Dykes J. 2017. Persistent Research Challenges in Geovisualization. *International Journal of Cartography* (3) 115-139.

6. Çöltekin, A., Janetzko, H., and Fabrikant, S. I. 2018. Geovisualization. *The Geographic Information Science & Technology Body of Knowledge,* John P. Wilson (Ed). DOI:10.22224/gistbok/2018.2.6.

7. Daniel Z Sui, James B Holt. 2008. Visualizing and Analysing Public-Health Data Using Value-by-Area Cartograms: Toward a New Synthetic Framework. Article in Cartographica: The International Journal for Geographic Information and Geovisualization, DOI: 10.3138/carto.43.1.003

8. Derryn A. Lovett, Alan J. Poots, Jake T.C. Clements, Stuart A. Green, Edgar Samara sundera, Derek Bell 2014. Using geographical information systems and cartograms as a health service quality improvement tool, Spatial and Spatio-temporal Epidemiology, Volume 10, Pages 67-74, ISSN 1877-5845,https://doi.org/10.1016/j.sste.2014.05.004.

9. Döll, P. 2017 Cartograms Facilitate Communication of Climate Change Risks and Responsibilities, Earth's Future, 5, 1182–1195, https://doi.org/10.1002/2017EF000677

10.Dorling, Daniel, 2011. Area Cartograms: Their Use and Creation, VL - 59 JO - Concepts and Techniques in Modern Geography (CATMOG) - 2011/04/24, SP - 252, EP - 260, SN - 9780470979587, DO - 10.1002/9780470979587.ch33

11.Dougenik J. A., Chrisman N. R., Niemeyer D. R.: 1985. An algorithm to construct continuous area cartograms. The Professional Geographer 37, 1 (1985), 75–81.

12.Field, K. Cartograms. 2017. *The Geographic Information Science & Technology Body of Knowledge*, John P. Wilson (ed) 2017. DOI: 10.22224/gistbok/2017.3.8

13. Ian K. Duncan, Shi Tingsheng , Simon T. Perrault, and Michael T. Gastner 2021,Task-Based Effectiveness of Interactive Contiguous Area Cartograms. IEEE transactions on visualization and computer graphics, vol. 27, no. 3, march 2021

14.Kirby, Russell & Delmelle, Eric & Eberth, Jan. 2016. Advances in spatial epidemiology and geographic information systems. Annals of Epidemiology. 27. 10.1016/j.annepidem.2016.12.001.

15.M. Correll. 2018. What does" visualization literacy" mean, anyway? - Multiple Views: visualization research explained.

16.Md. Jawaherul Alam ; Stephen G. Kobourov ; Sankar Veeramoni.2015 **Quantitative Measures for Cartogram Generation Techniques**F **https://doi.org/10.1111/cgf.12647**

17..J. Smith,J.K. Hillier,J.-C. Otto,M. Geilhausen.2013. Geovisualization. Chapter in Treatise on Geomorphology, Elsevier.

18.M.-J. Kraak.2009, in International Encyclopedia of Human Geography.

19.M.T. Gartner, M.E.J. Newman.2004. Diffusion-based method for producing density-equalizing maps Previous Methods for Constructing Cartograms, PNAS, 101 (20) (2004), pp. 7499-7504

20.Nöllenburg M. 2007.Geographic Visualization. In: Kerren A., Ebert A., Meyer J. (eds) Human-Centered Visualization Environments. Lecture Notes in Computer Science, vol 4417. Springer, Berlin, Heidelberg 2007. ISBN978-3-540-71948-9 https://doi.org/10.1007/978-3-540-71949-6_6

21. Nusrat S, Alam MJ, Kobourov SG. 2016. Evaluating cartogram effectiveness. IEEE Trans Vis Comput Graph. 10.1109/TVCG.2016.2642109

22.Nusrat, S., Kobourov, S. 2016.The State of the Art in Cartograms. *Computer Graphics Forum*, *35*(3), 619-642. https://doi.org/10.1111/cgf.12932

23.Robert Laurini,2017. Geovisualization and Chorems, in Geographic Knowledge infrastructure, Elsevier,2017, Pages 223-246, ISBN 9781785482434, https://doi.org/10.1016/B978-1-78548-243-4.50011-6.

24.Röger, Christian, Krisp, Jukka. 2019. Using Cartograms for Visualizing extended Floating Car Data (xFCD), in Proceedings of the ICA, 2019/07/10, DOI - 10.5194/ica-proc-2-107-2019

25.Sandul Yasobant, Vora K, Carl Hughes, Ashish Upadhyay. 2015. Geovisualization: A Newer GIS Technology for Implementation Research in Health. February 2015, Journal of Geographic Information System 7(01):20-28,DOI: 10.4236/jgis.2015.71002

26.Thomas C. van Dijk, Dieter Lutz.2018. Realtime linear cartograms and metro maps. SIGSPATIAL '18: Proceedings of the 26th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems November 2018 Pages 488–491https://doi.org/10.1145/3274895.3274959

27.Wood, Alexander Kachkaev, and Jason Dykes. 2018. Design Exposition with Literate Visualization. IEEE transactions on visualization and computer graphics volume 25(1).

28. www.aps.dz consulted 2021