Applying evolutionary algorithm for the study of mental disorder prevalence: the case of the Barcelona metropolitan area

Authors & affiliations:
J.A. Salinas-Pérez*, M.L. Rodero-Cosano†, C.R. García-Alonso†

†Universidad Loyola Andalucía
*jsalinas@uloyola.es

Abstract:

The spatial data analysis is a methodology which is widely used to study geographic patterns of diseases. Some of the most popular and effective tools are Local Indicators of Spatial Association. Among these, highlight Moran’s I and Getis &Ord’s G.

The last contribution in this field is a tool based on a Multi-Objective Evolutionary Algorithm, which optimizes the following spatial objectives: i) maximizing (hot spot) or minimizing (cold spot) mental disorders treated prevalence, ii) minimizing standard deviation of the prevalence, iii) minimizing the distance between spatial units.

The algorithm starts to optimize a solution that has been generated randomly by genetic operators (selection, mutation and crossover) in an iterative process by which the solutions are improved in successive generations. The results of each generation are assessed by four fitness functions which highlight the efficiency of the results to solve the spatial problem. The solutions can be different from each fitness function, so the final results will be those identified by at least two functions.

This method has been applied in regional analysis previously, but it has not yet been validated in macro urban areas, being this the first time that it is used for this. To do so, data from the

**Acknowledgements:** This paper was developed in the context of the research project PI11/02008 funded by the Institute of Health Carlos III.
Minimum Data Set for Outpatient Mental Health Centers were analyzed to identify the hot and cold spots of mental disorders prevalence in basic health areas within Barcelona metropolitan zone during 2009. The results obtained are consistent with the solutions of both classical Local Indicators of Spatial Association.

Applying evolutionary algorithm for the study of mental disorder prevalence: the case of the Barcelona metropolitan area

1. Introduction

Urbanicity bias on the morbidity of a large number of diseases due to the typical features of urban environment (Godfrey & Julien, 2005). These features are emphasized in macro-urban or metropolitan areas. The presence of spatial clusters of high or low values of incidence, prevalence or mortality in this kind of areas could indicate interesting situations of inequity for health planners and decision makers (Koschinsky, 2013). Macro-urban effects could not be considered in lower geographical scales such as municipality or regional.

The spatial analysis of diseases in macro-urban areas using well-known methods has got numerous study (Liu et al., 2012; Makita et al., 2011; Puigpinos-Riera et al., 2011), among other, with mental illnesses. The scientific literature has stated a closer relationship, positive or negative, among urbanicity and psychiatric prevalence. These associations have been mainly studied for schizophrenia and psychotic disorders (Allardyce et al., 2005; Kirkbride et al., 2007; Peräla et al., 2008; Kelly et al., 2010) or depression (Wang, 2004; Duncan et al., 2013). According to a WHO study, one third of the adult population suffers from a mental disorder each year (Kessler & Üstün, 2008). Despite mental disorders are one of the most prevalent diseases; there are not many examples of spatial analysis studies on mental health in macro urban areas (Curtis et al., 2006).

There are multiple robust spatial data analysis methods which have been widely used in epidemiological studies (Auchincloss, Gebreab, Mair, & Diez Roux, 2012). Multi-Objective Evolutionary Algorithms (MOEA) are one of the latest methods that have emerged in this field. MOEAs come from the artificial intelligence scope to find the most optimal solutions to complex multi-objective problems (Coello-Coello, Lamont, & Van Veldhuizen, 2007). Algorithms are based on the Natural Evolution and optimize the solutions iteratively throughout successive generations. An MOEA specifically designed to solve spatial problems was used to find clusters of high spatial autocorrelation values of treated prevalence of mental

Both researches had a regional geographical scale considering the municipalities as analysis spatial units, but this scale may hide relevant situations in highly populated municipalities. Thus, it would be interesting to check the accuracy of MOEA in lower scales such as macro urban or metropolitan areas with more of 1 million of inhabitants.

This paper aims the spatial analysis of the treated prevalence of mental disorders in the Barcelona metropolitan area. A MOEA is used to identify spatial clusters of high psychiatric prevalence from the results of three LISA. The study would also allow check the MOEA for the spatial analysis of macro-urban areas.

2. Methods

This research analysed the geographical distribution of mental disorders in health basic areas within Barcelona metropolitan area in 2009. The Metropolitan Area of Barcelona is a territorial entity within Catalonia Region in North Eastern which has a population of 3,239,337 (Instituto Nacional de Estadística, 2012) and it is one of the most developed Spanish Urban Areas.

Health basic areas were selected as units of analysis being a total of 207 units. These are the smallest areas where reliable statistical information can be found for precise geographical identification and location of hot and cold spots of treated prevalence of mental disorders.

We have used the 2009 database provided by the Catalanian Department of Health, Minimum Data Set for Outpatient Mental Health Centres (CMBDSMA), the variables used to calculate de treated prevalence were: sex, age, zip code and main diagnosis. The number of inhabitants (year 2009) in each health basic area was obtained from the Population of the Continuous Municipal Register by Population Unit.

Patient sex and age provided information to calculate the standard rates of treated prevalence of mental disorders (per 1,000 population) through the direct method (Rezaeian, Dunn, Leger, & Appleby, 2007) that took into consideration the population of Metropolitan Area of Barcelona.

The introduction has emphasized the importance of studying the distribution of diseases, especially to seek both high and low treated prevalence clusters. Traditionally the researchers have used Local Indicators of Spatial Association (LISA) such as Moran’s I and Getis &Ord’s G (Anselin, 1995; Ord & Getis, 1995). Currently the studies are improved the LISA
introducing other methodologies that at this point of time have not used for this analyses such as MOEA.

MOEA is a methodology used to solve complex multi-objective problems through optimization to achieve feasible and efficient solutions based on biological principles of genetic and evolution. The full technical aspects of this methodology has been widely described in previous papers (Carlos R García-Alonso, Campoy-Muñoz, & Salazar Ordoñez, 2013; Carlos R García-Alonso, Pérez-Naranjo, & Fernández-Caballero, 2011; Carlos R. García-Alonso, Salvador-Carulla, Negrín-Hernández, & Moreno-Küstner, 2010).

The algorithm search different regions of a solution space which makes possible to find a diverse set of solutions for complex problems with non-convex, discontinuous, and multimodal solutions spaces (Konak, Coit, & Smith, 2006). This methodology use three operators to generate new solutions from existing ones: selection, crossover and mutation.

This methodology is based on that of strength Pareto evolutionary algorithm 2 (SPEA2) (Zitzler, Laumanns, & Thiele, 2002). An elitism operator is applied to a predefined number of the best solutions obtained and is also “a posteriori” techniques, because the relevance of each objective is similar (Das & Dennis, 1997; Laumanns, Thiele, & Zitzler, 2006; Srigiriraju, 2000).

The initial populations selected are 100 sets of 10 health basic areas identified by their codes. The standard genetic operators are systematically used to improve the values of the fitness functions for each of the mentioned 100 sets of 10 health basic areas.

The objectives that structure the fitness functions were: maximize (for hot spots) or minimize (for cold spots) the mean of the treated prevalence of metal disorders in a set of n health basic areas (n = 10); minimize the Standard Deviation (SD) of the treated prevalence of metal disorders in the same set of n health basic areas (n = 10); minimize the minimum distance that links all the health basic areas in the solution.

Four fitness functions were designed combining the three objectives. These fitness functions were: Fine-Grained strength Pareto, weighted objectives, standard ranking selection and fuzzy evaluation of weighted objectives.

All the resulting solutions show clusters in which the treated prevalence of metal disorders is high (potential hot spots) or low (potential cold spots). Within the results we have selected the health basic areas’ clusters the most statistical frequently as the final hot and cold spot.
3. Results

The mental health prevalence in Barcelona Metropolitan Area obtained from CMBD-A has been mapped in the Figure 1. The prevalence mean is 30.24 per 1,000 inhabitants (±0.054; \( p=0.05 \)) and the standard deviation is 12.45 which indicate prevalence values widely different between areas. These differences are not also randomly distributed in the territory according to the global Moran score (Figure 2).

The MOEA has defined 29 HA as hot spot (HA in red) and 37 HA as cold spot (HA in blue) which are shown in the Figure 3 and 4. The prevalence mean of the hot spots is 46.06 and the cold spot is 14.65.

**Figure 1.** Spatial distribution of the treated prevalence of mental disorders (cases/1,000 inhabitants, 207 spatial units: Health Areas) of Barcelona Metropolitan Areas.

**Figure 2.** Basic statistics of mental disorders treated prevalence and hot/cold spots.

<table>
<thead>
<tr>
<th></th>
<th>Total Health Areas</th>
<th>Hot spots</th>
<th>Cold spots</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Mean</strong></td>
<td>30.24</td>
<td>46.06</td>
<td>14.65</td>
</tr>
<tr>
<td><strong>Standard deviation</strong></td>
<td>12.45</td>
<td>9.87</td>
<td>9.22</td>
</tr>
<tr>
<td><strong>Confidence Interval</strong>( p=0.05 )</td>
<td>±0.054</td>
<td>±0.12</td>
<td>±0.095</td>
</tr>
<tr>
<td><strong>Global spatial autocorrelation indexes</strong></td>
<td>( I=0.06 ) ( Z=8.73 \text{ s.d.} ) ( p≤0.01 )</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>
A large cluster of hot spot is located in the North of Barcelona city, into Nou Barris, Horta-Guinardó and Gràcia districts (14 HA). Other large cluster appears mainly around the Llobregat river South of Barcelona city (10 HA). The remainder of hot spots are located around the cities of Sabadell and Mollet del Vallès (North Central area), and Villafranca del Penedès (West of Metropolitan Area).

**Figure 3. Spatial distribution of hot spots of mental disorders treated prevalence in the Barcelona Metropolitan Areas.**

On the other hand, 37 HA have been classified in the analysis of cold spots. These are located more scattered than hot spots although there are two large clusters linked to Barcelona city. One of them is located in the districts of Sarrià-Sant Gervasi, Les Corts and Eixample together with the municipality of l’Hospitalet de Llobregat, and the another is in Sant Martí district and the municipality of Sant Adrià del Besòs. The last large cold spot cluster appears around Badalona and Mongat municipalities. The remainder cold spots are Vilanova i la Geltrú, Begues, Terrassa, Vilassar de Dalt and Sant Andreu de Llavaneres.
4. Discussion

The proximity of some hot and cold spots may indicate a relationship among them. It has been stated in previous studies (Sridharan, Koschinsky, & Walker, 2011; Salinas-Pérez, García-Alonso, Molina-Parrilla, Jordà-Sampietro, & Salvador-Carulla, 2012b).

The scale of the study area has allowed finding hot and cold spot of mental health cases into the highly populated municipalities such as Barcelona, l’Hospitalet or Badalona, which they had been lost in a municipality scale study. Thus, it has been possible to identify city areas with significant high or low prevalence of mental disorders and if there are also relationships between adjacent city areas of different municipalities.

The studies of spatial data analysis usually use an only analysis spatial unit such as ZIP codes, statistical units, municipalities, counties or administrative zoning. This study has used the most basic area of the health administration which it is the minimal spatial level where epidemiological data are collected. This lets us analyse infra-municipal geographical variations in large cities but, unfortunately, to lose information in the less populated towns which are joined in a same catchment area according to economy of scale criteria. In cases like this, it would be appropriate to combine different spatial units in order to obtain the maximum information.
The spatial patterns in little urban areas may help to mental health planners and decision makers to improve the health care taking into account local evidence (Lewin et al., 2009; Salvador-Carulla, Garcia-Alonso, Gibert, & Vázquez-Bourgon, 2013).

One of the classical limitations of the spatial data analysis is the called boundary problem. The values of the neighbouring areas outside of the study area are not known. This may hide possible borderline spatial clusters. However, this problem is reduced in the East of Barcelona Metropolitan Area because of the seacoast.

This study is not aimed at identifying the causes of hot/cold spots. Mental disorders prevalence has been related with various social, genetics and environment determinants in the scientific literature (Fryers, Melzer, & Jenkins, 2003). Despite there are common mental health determinants, each psychiatric diagnose has specific determinants.

5. Conclusions

The MOEA has identified and located several spatial clusters with statistically significant high or low values of treated prevalence of mental disorders in the Barcelona Metropolitan Area. This analysis has allowed finding interesting situations such as spatial clusters both in infra-municipal areas and in neighbouring areas between municipalities. Health Areas with high and low treated prevalence provide informed evidence for mental health planning and policies.

The futures lines of research are: relation the treaded prevalence of mental disorder with urban factor which allows us identifies possible risk factors associated with macro-urban areas and check if the treated prevalence of mental disorders is significantly different between rural and urban areas.

References


