

Where is the Poverty Area? Quantifying the Neighborhood Effect in a Deprivation Index Estimation: A case Study in Quito, Ecuador

Chunzhu Wei, Pablo Cabrera Barona and Thomas Blaschke

Department of Geoinformatics – Z_GIS, University of Salzburg, Austria ·
chunzhu.wei@stud.sbg.ac.at

Full paper double blind review

Abstract

Automated zoning procedures offer efficient, systematic and objective methodologies for identifying the neighborhood effects on socio-economic statistics. However, the automatic spatial aggregation of census data over manually defined geographic units based on landscape heterogeneity characteristics are barely studied. In this study we utilize high-resolution remote sensing data and census data and apply a multi-level zoning system in order to analyze how a deprivation index differs in the corresponding urbanization environment within the Shannon's diversity Index. Our study area is the capital city of Ecuador, Quito. The results of the autocorrelation analysis show that, within the Shannon's Diversity-based multi-level zoning system, areas with a low degree of deprivation in the city center of Quito tend to be larger as the size of the neighborhood increases, and the poverty areas, which are mainly located in the north-east and south-west of Quito, differ significantly between different zoning levels. Our conclusion is that the neighborhood effect influences not only the composition of the spatial pattern and social data, but also their correlation and autocorrelation. Therefore, when analyzing the environment effect of urbanization and its influence on the society development, different levels of zoning systems should be taken into consideration.

1 Introduction

In order to deal with the Modifiable Areal Unit Problem (MAUP) (OPENSHAW 1984) associated with aggregated census data, automated zoning procedures have offered several methods for zoning boundary identification and have allowed the investigation of the neighborhood effect on health statistics, human distribution and choosing a regional-building area (FLOWERDEW 2008, MARTIN 2003, OPENSHAW 1977). Some landscape ecologists also assume that ecological processes affect the neighborhood context within a zone and the interaction between zones, which means that environmental heterogeneity may impact socio-economic variables (Wagner and Fortin 2005). Nevertheless, most design criteria that have been used to define multi-scale zoning systems for census data aggregation only take the social-economic factors into consideration, such as population size, deprivation, homogeneity of ill-health dynamics policy intervention policies (COCKINGS & MARTIN 2005)

neglecting to explicitly and systematically explore the effect of spatial landscape patterns in zonal boundary definition. It is therefore appropriate to design a strategy to analyze how spatial patterns or landscape diversity influence the socio-economic indices within multiple scale mapping units (TURNER et al. 1989, WU et al. 2000, LONGLEY AND BATTY 1996). In this study, we aim to test whether a landscape diversity-based multi-level zoning system would have neighborhood effects on the assessment of deprivation, and to quantify the neighborhood effects on the autocorrelation analysis between landscape diversity and an index of deprivation (HACKER et al. 2013).

2 Methods

2.1 Study Area and Data collection

The Ecuadorian capital city of Quito, located in the North of the country at an elevation of 2800 meters, is home to 2,671,191 inhabitants according to the last Ecuadorian Population and Housing Census of 2014. Quito is characterized by a narrow spread, which is approximately 35 kilometers in length and only about 9 kilometers in width. In recent years, along with socio-economic development, the urban sprawl of Quito city has taken place in the southern and northern locations, and some illegal constructions have emerged around the edges of the city, mainly in the hills and steep lands. This process is changing the urban growth structure. Therefore, the way in which different spatial patterns (i.e., urban form, land-use distribution, and connectivity) generate different effects on the urban structure, and how they differ in their corresponding social attachment on the environment has become a hot research topic in Quito.

For the data collection, the census data concerning the deprivation index used for the social factors analysis came from the 2010 Ecuadorian Population and Housing Census. In addition, remote sensing data (see next section) was used to extract the land use and land cover information for Quito and conduct further spatial patterns analysis.

2.2 Deprivation Index calculation

Four indicators were used for the Deprivation Index estimation in this study: the percentage of houses without access to the sewerage system, the percentage of houses without a connection to the electrical system, the percentage of houses without a connection to the drinking water supply system, and the percentage of houses without a garbage collection service. The indicators were chosen using the following criteria: i) to be related to a basic human needs approach (MIDEROS 2013, RAMÍREZ 2012), ii) to be related to socio-economic deprivation (Pampalon et al 2009), iii) to be available in a Census dataset and to be spatially represented in census blocks and iv) to represent latent deprivation problems still common in Latin American cities. These indicators were extracted from the 2010 Ecuadorian population and housing census. The indicators were geocoded and included in the shapefile of the study area census tracks. Variances Inflation Factors (VIFs) were calculated for all the indicators to identify multi-collinearities using the equation (1):

$$VIF = \frac{1}{1 - R^2} \quad (1)$$

VIF is the reciprocal of tolerance (OECD 2008) and is a measure that quantifies how much the variance is inflated. This measure is obtained from the variation of an estimated coefficient of a simple linear regression where R^2 is the coefficient of correlation representing the correlation between the indicators. If VIF is lower than 5, the indicator can be used to construct the index (OECD 2008). All the obtained VIF values were lower than 5, meaning that all the selected indicators could be included in the index.

A simple linear combination was applied to construct the index by adding the four indicators, and then a linear min max normalization was applied (MALCZEWSKI 1999) using the formula (2):

$$v_i = \frac{v_i - v_{\min}}{v_{\max} - v_{\min}} \quad (2)$$

Where v_i is the non-normalized value of the index, v_{\min} is the lowest value of the index, and v_{\max} is highest value of the index. The final result is an index with values between 0 and 1, where values closer to 1 represent a higher level of socio-economic deprivation.

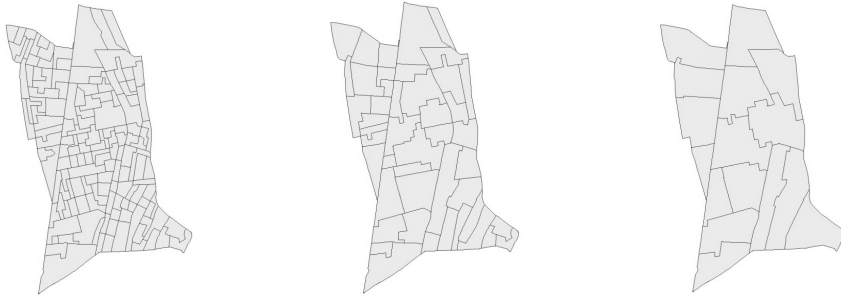
2.3 Landscape Diversity Analysis

A standard processing level 3.0 (orthorectified) was applied to the 2012 Rapid Eye imagery. After the imagery pre-process and multi-resolution segmentation, the Normalized Difference Vegetation Index (NDVI) was used to distinguish the non-vegetated areas and the vegetated areas (TAPSALL et al. 2010). The results of the landscape survey were used as a thematic input layer to further classify the non-vegetated areas into 3 classes, namely residential areas, multiple use areas and industrial areas, and to classify the vegetated areas into the 8 classes grass, shrub, mountain shrub, eucalyptus, farmland, bare soil, other green space and rivers.

For the ecological heterogeneity analysis, the Shannon's diversity index (SHDI) is usually used to quantify the landscape composition (i.e. proportions of habitat patches) (MCGARIGAL et al. 2002). Therefore, combined with the land use and land cover classification of the Rapid Eye Data, the Diversity calculation function in the software Fragstats was used to calculate the SHDI and identify the characteristics of landscape heterogeneity (MCGARIGAL 2006).

2.4 Automatic Zoning System Definition

OPENSHAW (1977) used the Automatic Zoning Procedure (AZP) to explore the scale and aggregation effects on the choice of zonal boundaries, and then applied an Automated Zone Matching (AZM) algorithm, developed from- and extending the AZP, by iteratively combining target population sizes, zonal compactness and other variables into the zone design process. The AZTool system, based on the AZM algorithm, has the advantage of minimizing the mismatch between different input zonal geographies (MARTIN 2003, FLOWERDEW et al. 2008). Taking Figure 1 as an example, the AZM can define the homogeneity based on the attributes of each boundary (Fig.1 a) and create different levels of zonal systems using homogeneity and setting the relative weights of category variables (Fig.1 b and c):



a: 146_level zoning system **b:** 46_level zoning system **c:** 16_level zoning system

Fig. 1: Representation of Automatic Zoning Algorithm

The zoning procedure in the AZTool includes: (1) Defining the threshold variable which can indicate the homogeneity of the boundary; (2) Defining the target value of the threshold variable; and (3) Defining the minimum and maximum values of the threshold variable (MARTIN 2003, COCKINGS & MARTIN 2005, FOTHERINGHAM & ROGERSON 2008). In contrast to traditional zone design strategies, which use population data as the threshold variable to define the zone, this research used SHDI as a separate threshold variables for the multi-level spatial aggregation, and the census data of demographic units containing 5000 polygons was used as the basic data. It is then required to manually set up the experimental conditions of the target value and the minimum and maximum values of the target threshold, in order to obtain a multi-level zoning system. According to many tests, we finally chose three groups of experimental conditions to create three multi-level zoning systems, namely an 800_level zoning system, a 150_level zoning system and a 50_level zoning system, with the number of levels representing the number of delineated polygons.

Table 1: The experimental conditions

Threshold	Level Number	Target	Min	Max	Zoning Level
Shannon's Diversity	(a)	100	10	600	800
	(b)	500	50	800	150
	(d)	300	1	5000	50

3 Results

The land use and land cover of Quito is shown in Figure 2. Figures 3.b to 3.d provide examples of the SHDI-based multi-level zoning system at three levels of data aggregation. Figure 3.a represents the 50_level SHDI-based zoning system, which contains 50 polygons; Figure 3.b and Figure 3.c represent the 150_level zoning system and the 800_level zoning system, containing 150 and 800 polygons separately. Figure 3.d shows the original census districts, containing 5000 polygons. In order to discuss how the SHDI-based multi-level zoning system influences the cluster results of the deprivation index, deprivation values

were used as the univariate for the local Moran I analysis within the SHDI-based multi-level zoning system (ANSELIN et al. 2006, Fotheringham & Rogerson 2008). The autocorrelation results (Figure 3.a to 3.d) of the deprivation index lead to similar spatial clusters in Low-Low area distributions, which means that the low deprivation areas are mainly surrounded by the areas with a low degree of landscape diversity characteristics. As shown in Figures 3.a to 3.d, the Low-Low distribution areas, which indicate the low deprivation areas are surround by low deprivation areas, did not have significant changes in the old city center and new city center of Quito. but these areas gradually become dispersive with decreasing neighborhood size in other locations in Quito However, the High-High distributions areas with high deprivation values are surrounded by areas with high deprivation values, gradually appear with a decrease in the neighborhood size, and reach the maximum at the original zoning system at the level of 5000. These areas are mainly located at the south-east and north-west border of Quito. This can demonstrate that the zoning effects exert an obvious influence on the delineation of areas with a high deprivation index.

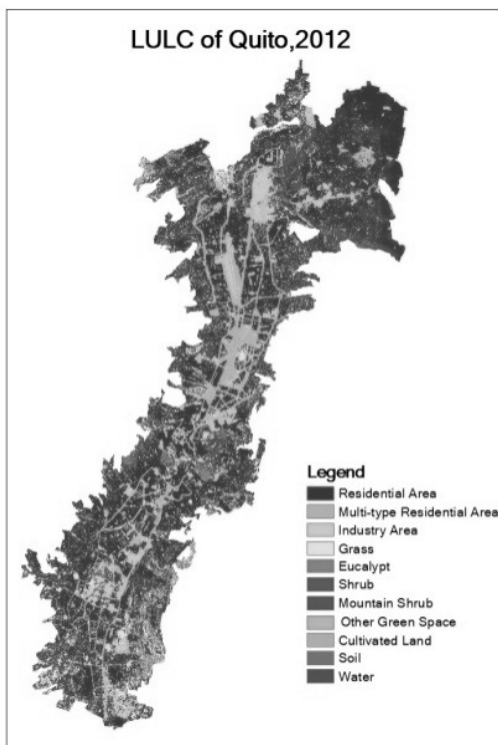


Fig. 2: Land use and land cover of Quito

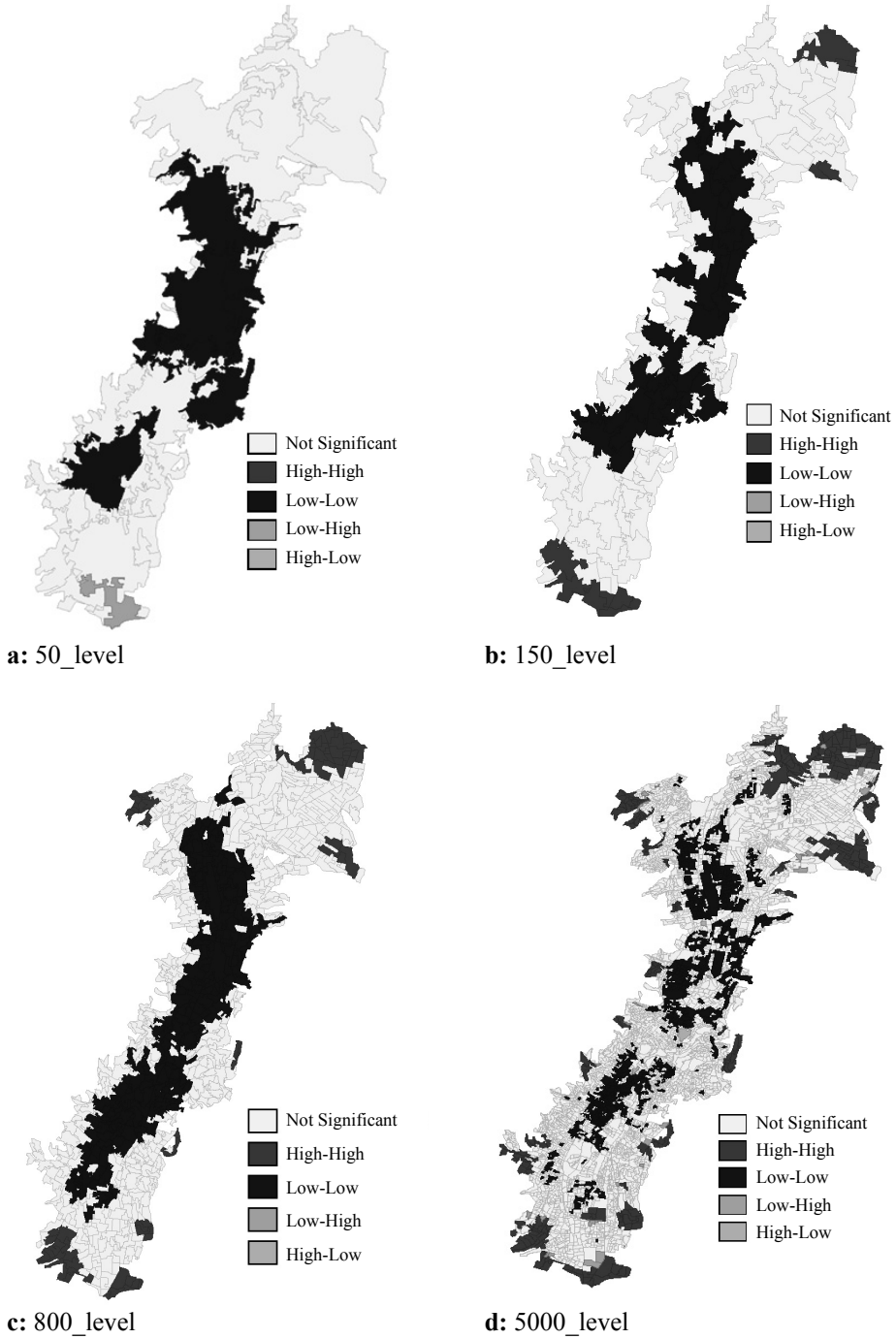


Fig. 3: Autocorrelation of deprivation index based on SHDI based multi-level zoning system

In addition, the autocorrelation of the SHDI and deprivation index in these four levels also clearly show the neighborhood effect on deprivation delineation. As shown in Figure 4, there are also significant changes in the autocorrelations between deprivation and SHDI from the 50_level of the SHDI based zoning system to the 5000_level of the original zoning system. Especially at the 50_level of the deprivation-based multi-level zoning system, the Low-High distribution areas, which mean the low poverty areas are surrounded by areas with a high degree of landscape diversity, are mainly located in the urban central axis of Quito. However, these areas become smaller and more scattered as the zoning level increases. The areas with high levels of poverty (including the High-Low areas and High-High areas) are mainly located in the north-east and south-west parts of Quito in the original zoning system, and the south-western areas demonstrate higher diversity characteristics than the north-eastern areas. verty areas decrease as the neighborhood size increases, and almost disappear at the 50_level of the SHDI based zoning system (Figure 4.a). Therefore, not only the deprivation estimation but also the ecological association between the deprivation index and landscape heterogeneity features is influenced by the neighborhood effect. The delineation of poverty areas would become complex if it is required to take the urban spatial featuThe pores into consideration.

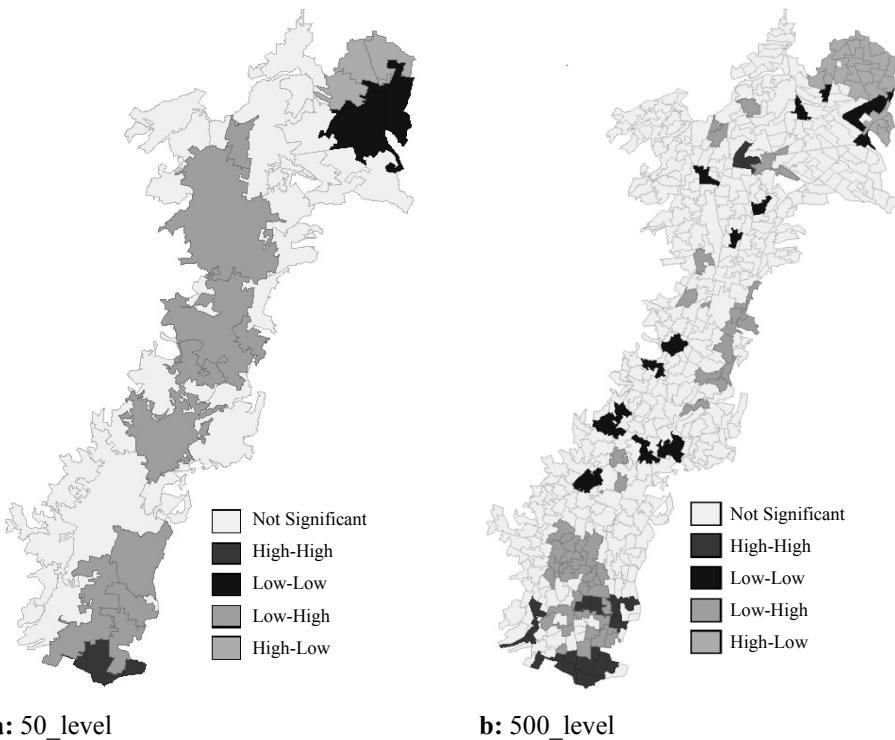


Fig. 4: Autocorrelation of deprivation and SHDI within the SHDI based multi-level zoning system

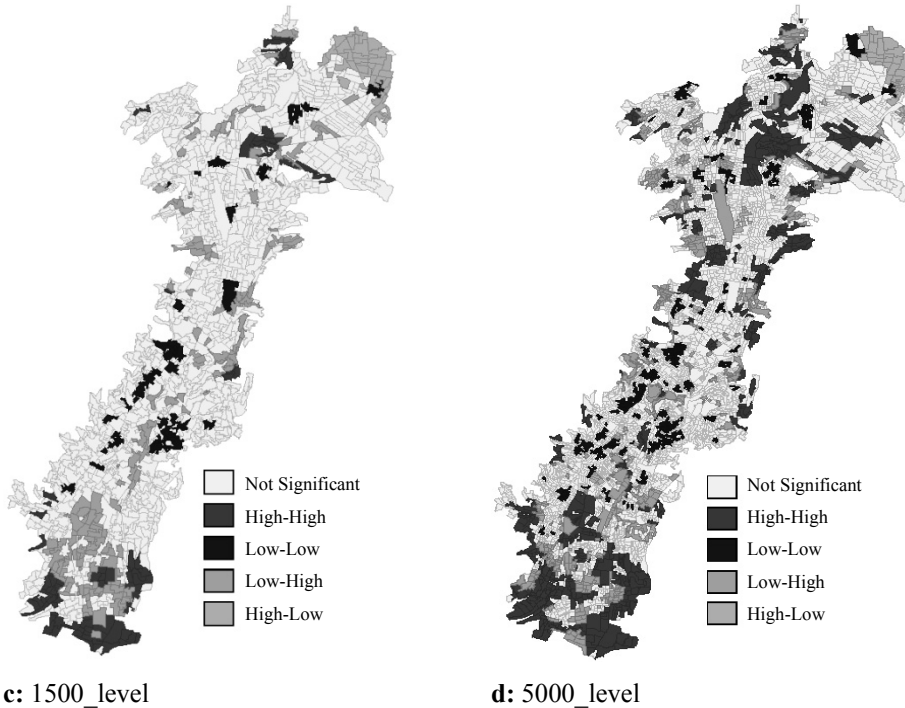


Fig. 4 (continued)

4 Conclusion

The AZTool, which optimized the deprivation index criteria, was used to evaluate the impact of different zoning sizes on urban poverty area delineation. According to the autocorrelation analysis of unique Local Moran I (including Deprivation index) and bivariate Local Moran I (including Deprivation index with Shannon's diversity index) in the Shannon's diversity based multi-level zoning system the urban corridor of Quito, which is the main region of the old city center and new city center, constitutes the areas with a low degree of deprivation, and is surrounded by high values of Shannon's diversity index at the 50_level of the SHDI based zoning system. However, such distributions gradually disappear when the aggregation level of the neighborhood increases. In contrast, the border areas of Quito only show the high poverty degree surrounded by high Shannon's diversity values at the original 5000_level zoning system. Therefore, due to the neighborhood effect, it is still very difficult to delineate the poverty areas through combining the landscape heterogeneity features and social-economic factors. The evaluation of the degree of deprivation was not only influenced by the evaluation criteria, but also by the boundary definitions. In addition, although autocorrelation statistic results demonstrate that different poverty levels of Quito in multi-level zoning systems allow the visualization of the spatial structure of the landscape system property, the causal ecological association between social deprivation and other urban spatial factors need further research to clarify the potential influence of landscape heterogeneity features in neighborhood structures related to socio-economic process.

Acknowledgements

The authors gratefully acknowledge the use of the AZTool software, which is copyright David Martin, Samantha Cockings and the University of Southampton. We would also like to thank the Ecuador's Ministry of Environment for providing us with the 2012 Rapid Eye imagery for our study area. We also thank the Ecuador's Ministry of Health and the Ecuador's Ministry of Education for providing the geo-referenced health services and geo-referenced educational services of the City of Quito.

References

- ANSELIN, L., SYABRI, I. & KHO, Y. (2006), GeoDa: An Introduction to Spatial Data Analysis. *Geographical Analysis*, 38 (1), 5-22.
- COCKINGS, S. & MARTIN, D. (2005), Zone Design for Environment and Health Studies Using Pre-Aggregated Data. *Social Science & Medicine, Burning Issues – Selected papers from the 10th International Symposium in Medical Geography, Manchester 2003*, 60 (12), 2729-2742.
- FLOWERDEW, R., MANLEY, D. J. & SABEL, C. E. (2008), Neighbourhood Effects on Health: Does It Matter Where You Draw the Boundaries? *Social Science & Medicine*, 66 (6), 1241-1255.
- FOTHERINGHAM, A. S. & ROGERSON, P. A. (2008), *The SAGE Handbook of Spatial Analysis*. SAGE.
- HACKER, K. P., SETO, K. C., COSTA, F., CORBURN, J., REIS, M. G., KO, A. I. & DIUK-WASSER, M. A. (2013), Urban Slum Structure: Integrating Socioeconomic and Land Cover Data to Model Slum Evolution in Salvador, Brazil. *International Journal of Health Geographics*, 12 (1), 45.
- LONGLEY, P. A. & BATTY, M. (1996), *Spatial Analysis: Modelling in a GIS Environment*. John Wiley & Sons.
- MALCZEWSKI, J. (1999), *GIS and Multicriteria Decision Analysis*. John Wiley & Sons.
- MIDEROS, A. (2012), Ecuador: defining and measuring multidimensional poverty, 2006-2010. *Cepal Review*, 108, 49-67
- MARTIN, D. (2003), Extending the Automated Zoning Procedure to Reconcile Incompatible Zoning Systems. *International Journal of Geographical Information Science*, 17 (2), 181-196.
- MCGARIGAL, K., CUSHMAN, S. A., NEEL, M. C. & ENE, E. (2002), FRAGSTATS: Spatial Pattern Analysis Program for Categorical Maps.
- MCGARIGAL, K. (2006), Landscape Pattern Metrics. In: *Encyclopedia of Environmetrics*. John Wiley & Sons, Ltd.
- OECD (2008), *Handbook on constructing composite indicators: methodology and user guide*. 160 p.
- OPENSHAW, S. (1977), A Geographical Solution to Scale and Aggregation Problems in Region-Building, Partitioning and Spatial Modelling. *Transactions of the Institute of British Geographers, New Series*, 2 (4): 459-472.
- OPENSHAW, S. (1984), Ecological Fallacies and the Analysis of Areal Census Data. *Environment & Planning A* 16 (1), 17-31.

- PAMPALON, P., PAMEL, D., GAMACHE, P. & RAYMOND, G. (2009), A deprivation index for health planning in Canada. *Chronic Diseases in Canada*, 29 (4), 178-191.
- RAMÍREZ, R. (2012), La vida (buena) como riqueza de los pueblos. *Hacia una socio ecología política del tiempo*. Economía e Investigación IAEN.
- TAPSALL, B., MILENOV, P. & TASDEMIR, K. (2010), Analysis of RapidEye Imagery for Annual Landcover Mapping as an Aid to European Union (EU) Common Agricultural Policy.
- TURNER, M. G., O'NEILL, R. V., GARDNER, R. H. & MILNE, B. T. (1989), Effects of Changing Spatial Scale on the Analysis of Landscape Pattern. *Landscape Ecology*.
- WAGNER, H. H. & FORTIN, M.-J. (2005), Spatial Analysis of Landscapes: Concepts and Statistics. *Ecology*, 86 (8), 1975-1987.
- WU, J., JELINSKI, D. E., LUCK, M. & TUELLER, P. T. (2000), Multiscale Analysis of Landscape Heterogeneity: Scale Variance and Pattern Metrics. *Geographic Information Sciences*, 6 (1), 6-19.