SPATIAL ANALYSIS OF CURRENT MIGRATION DIFFERENTIATION OF MUNICIPALITIES IN CZECH REPUBLIC

[a]LINDA KREJNÍKOVÁ

Department of Demography and Geodemography, Faculty of Science, Charles University in Prague, Albertov 6, 128 43 Prague 2, Czech Republic
email: [a]linda.krejnikova@gmail.com

Abstract: At present, the spatial dependence of geographic phenomena, where the level of some variable in one region correlates with the level of this variable in the nearby region, is emphasized. The focus of this paper is to evaluate the current state of migration spatial differentiation in the Czech Republic on data for municipalities (1) regardless of the spatial conditionality of this phenomenon using cluster analysis and (2) with respect to the spatial conditionality of the phenomenon using the principles of spatial autocorrelation. The aim is to create and interpret clusters of municipalities according to the similarity of the intensity of migration balance and the intensity of migration turnover for both applied approaches, and finally to compare the outputs of these two approaches.

Keywords: Keywords migration, cluster analysis, spatial autocorrelation.

Introduction

In the last few decades, the issue of spatial autocorrelation is received considerable attention, not only in geography but also in economics, biology, epidemiology, ecology, urban planning and sociology (Getis 2008). Geographers have long recognized the role of distance on spatial phenomena, as evidenced by Tobler's First Law of Geography that says "everything is related to everything else, but near things are more related than distant things" (Tobler 1970). Spatial autocorrelation tests whether the value of observed variable is independent on values of the variable at neighbouring localities (Cliff and Ord 1973). When modelling spatial data, it is necessary to work not only with the characteristics of observed phenomenon but also with the influence of spatial scale to avoid errors in the results interpretation (Anselin 1988).

Pioneers in exploring the spatial autocorrelation were authors Cliff and Ord (1969, 1973, 1981). They also generalized and fully developed the Moran's I (author is Moran 1950) which is currently the most widely used statistics for measuring spatial autocorrelation. In the nineties, the research in the field of spatial autocorrelation focused on local conditions (Getis and Ord 1992) and subsequent local Moran's I statistics, named as LISA – Local Indicators of Spatial Association, which is used to identify possible centres of statistically significant clusters, was proposed (Anselin 1994). In this paper, both mentioned statistics of spatial autocorrelation are used.

The aim of this paper is to conduct a detailed analysis of the current state of migration spatial differentiation in the Czech Republic based on data for the smallest possible territorial units - municipalities. This paper focuses on the identification of spatial units consisting of municipalities with similar migration indicators (net migration intensity per 1000 inhabitants and the intensity of migration turnover per 1000 inhabitants). Quantitative analysis of migration spatial differentiation of the Czech Republic at the municipal level provides a framework for spatial differentiation at the micro level, which can be regarded as necessary to address new research questions as well as for the formulation of local development policies and strategic documents. The paper is divided into two parts: the methodological part, which includes introduction to dataset and its processing and also basic theoretical approaches and the empirical part, focusing on actual quantitative analysis (cluster analysis and analysis of spatial autocorrelation using global and local Moran's I), comparison of two applied approaches and interpretation of results.

1 Data and methodology

Analysis of migration spatial differentiation was made based on data for municipalities of Czech Republic in 2011. Data come from the public database of the Czech Statistical Office and was processed in an IBM SPSS statistical software and geographic ArcGIS software.

Two approaches have been used and then compared in order to describe the migration spatial differentiation. The first is the cluster analysis as a method of multivariate statistical analysis, a second the analysis of spatial autocorrelation. Initially, the variables intensity of migration balance and intensity of migration turnover (per 1,000 mid-year population of 2011) for each municipality in the Czech Republic were created.

Cluster analysis uses tools and methods to detect and create "natural" groups of entities (individuals, objects, phenomena) that occur in the analyzed multidimensional dataset. Detecting and creating clusters is based on grouping elements according to their mutual similarities. Created groups should be as homogeneous as possible, while differences between the groups should be as large as possible (Härdle and Simar 2003; Norušis 2012). Lack of cluster analysis, view of the analysis of geographic data, is its insensitivity to the spatial connections between the observed territorial units (Novák and Netrdová 2011). Cluster analysis was applied to the both monitored variables separately. In both cases, hierarchical clustering was used, type of scale proximity was the Euclidean distance and the clustering algorithm was based on intragroup linkages. Number of clusters was determined on 4.

Analysis of spatial autocorrelation is based on the calculation of a global and local Moran's I criteria. First, the degree of spatial variability (spatial clustering) of surveyed variables using the global Moran's I criteria were measured. It is calculated as

$$I = \frac{n \sum_{i=1}^{n} \sum_{j=1}^{n} w_{ij} (x_i - \bar{x})(x_j - \bar{x})}{\sum_{i=1}^{n} \sum_{j=1}^{n} w_{ij} \sum_{i=1}^{n} (x_i - \bar{x})^2}$$

where I is the global Moran's I, n is the number of surveyed units (municipalities), x_i is the value of the monitored variable in the unit i, xj is the value of the monitored variables in the unit j, x with stripe is the arithmetic average of the monitored variable and w_{ij} indicates the matrix of spatial burdens (Cliff and Ord 1973). Using this matrix, neighbouring units (municipalities) are defined on the basis of established criteria for defining of "neighborhoodness" (the spatial weighing scheme).

In this study, the threshold distance 10 km was chosen as the spatial weighting scheme based on the results of studies already carried out on a similar statistical sample (Blažek and Netrdová 2009; Spurná 2008). Moran's I reaches values from -1 to +1. The negative Moran's I indicates negative spatial autocorrelation, the positive value of Moran's I indicates positive spatial autocorrelation. The closer the value of Moran's I to -1 or +1, the stronger is the spatial autocorrelation.

To identify spatial clusters, local spatial autocorrelation statistics (the local Moran's I) were further calculated as

$$I_i = \frac{(x_i - \bar{x}) \sum_{j=1}^{n} w_{ij} (x_j - \bar{x})}{\frac{\sum_{i=1}^{n} (x_i - \bar{x})^2}{n}}$$

where I_i is the local Moran's I (Anselin 1994). The sum of these local statistics must be equal to the global Moran's I, therefore,

$$I = \sum_{i}^{n} \frac{I_i}{n}$$

The positive value of local Moran's I indicate that the municipality is surrounded by municipalities with similar values of the observed variable. This municipality is then part of the cluster. The negative value of local Moran's I on the other hand means that the municipality is surrounded by municipalities with different values of the monitored variable. Such municipalities are then "outliers". Local Moran's I can be interpreted only in the context of the calculated z-score and P-values. To interpret the clusters of spatial autocorrelation, the types of spatial association (COType) entered into maps of spatial associations are therefore used. Types of spatial association distinguish between statistically significant cluster of high values (high-high clusters), a cluster of low values (low-low clusters), a cluster of outliers in which high values are surrounded particularly by low values (high-low clusters) and a cluster of outliers in which are low values surrounded mainly by high values (low-high clusters). High-high and low-low types of spatial associations indicate positive spatial autocorrelation and vice versa high-low and low-high types the negative spatial autocorrelation (Anselin 1994).

2 Migration spatial differentiation of Czech Republic

2.1 Application I: Cluster analysis

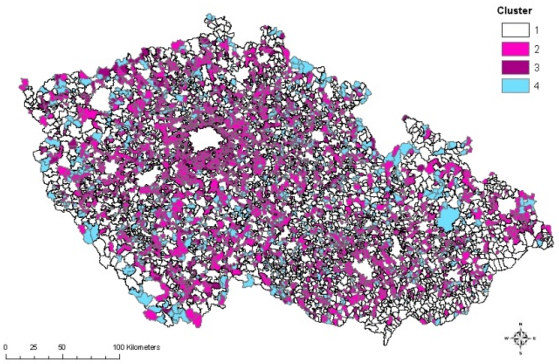
Division of municipalities into individual clusters according to both variables show cartograms in Figure 1 and 2. Descriptive characteristics calculated to better understand the differences between clusters are recorded in Table 1 and 2.

Table 1: Descriptive statistics for clusters of the intensity of migration balance (IMB)

IMB	N	Min	Max	Average	St. dev.	Var.
Cluster 1	2864	-666,67	292,99	-2,29	23,16	536,52
Cluster 2	1946	8,93	243,84	21,28	17,24	297,06
Cluster 3	542	35,26	85,11	50,51	12,67	160,46
Cluster 4	899	-107,14	-11,15	-21,23	9,96	99,12

Own calculations (data source: CSO).

Figure 1: Cartogram of IMB clusters



Own calculations (data source: CSO).

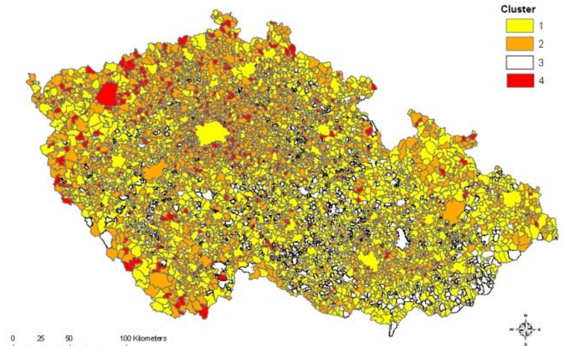
For the IMB, following clusters were created: *Cluster 1* contains 2,964 municipalities incl. Prague, Brno-city, Ostrava, Plzen-city and other large cities. It is a cluster with the second lowest average IMB. The difference between immigrants and emigrants per 1,000 inhabitants is on average slightly below zero in these municipalities. However this cluster has the highest variance, therefore, its internal variability is relatively large. *Cluster 2* contains 1,946 municipalities. It is a cluster with the second highest average IMB per 1,000 inhabitants. The values of IMB are positive in all municipalities within this cluster, therefore these municipalities are migration profitable. The variance of the values of IMB in this cluster is also considerable. *Cluster 3* contains 542 municipalities and it is a cluster with the highest average IMB. Municipalities belonging to this cluster are significantly migration profitable. These municipalities, together with the municipalities of cluster 2, form migration profitable periphery of large cities such as Prague, Brno, Plzen-city and more. *Cluster 4* contains 899 municipalities. It is a cluster with the lowest average IMB. Municipalities belonging to this cluster are migration loss and are widely dispersed in the neighbourhood of the state border. This area is represented mainly by municipalities from this or the first cluster and it is therefore a migration loss or neutral region.

Table 2: Descriptive statistics for clusters of the intensity of migration turnover (IMT)

IMT	N	Min	Max	Average	St. dev.	Var.
Cluster 1	2586	28,82	352,49	46,37	27,57	760,06
Cluster 2	2352	53,12	314,74	70,57	19,22	369,20
Cluster 3	899	0,00	666,67	19,53	23,22	539,00
Cluster 4	414	95,51	133,33	109,95	10,67	113,83

Own calculations (data source: CSO).

Figure 2: Cartogram of IMT clusters



Own calculations (data source: CSO).

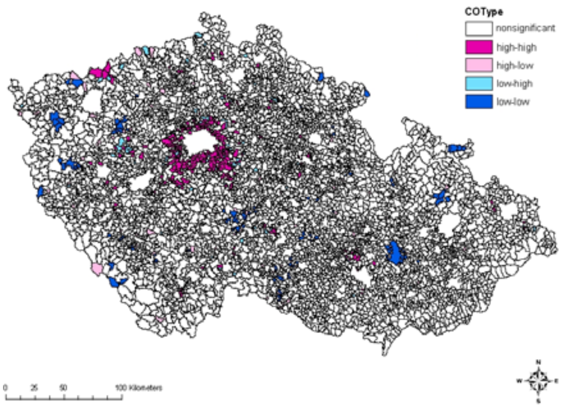
For the IMT, following clusters were created: *Cluster 1* contains 2,586 municipalities incl. Prague, Brno-city and Ostrava. This is a cluster with the second lowest average IMT. It is also a cluster with the highest variance, hence its internal variability is relatively large and municipalities may vary greatly in their values of IMT. *Cluster 2* contains 2,352 municipalities and represents the largest share of the total migration. It is also the cluster with the second highest average IMT per 1,000 inhabitants. *Cluster 3* contains 899 municipalities and it is the cluster with the lowest average IMT. Municipalities belonging to this cluster are significantly migration passive and shapes geographic belt on the border with Slovakia and also on the border between Bohemia and Moravia. *Cluster 4* contains 414 municipalities and it is a cluster with the highest average IMT. Municipalities from this cluster together with the municipalities from the second cluster are migration most active. Geographically form Prague periphery, a strip of municipalities in northwest Bohemia and smaller cluster by the state border in southern Bohemia.

2.2 Application II: Analysis of spatial autocorrelation

For the analysis of spatial autocorrelation, two following null hypotheses were expressed: *H0a: The variable IMB is not spatially autocorrelated* against the alternative hypothesis *H1a: The variable IMB is spatially autocorrelated* and hypothesis *H0b: The variable IMT is not spatially autocorrelated* against the alternative hypothesis *H1b:The variable IMT is spatially autocorrelated*.

Based on the value of global Moran's I, converted to z-scores for the first variable (14.86), the hypothesis H0a was rejected and significant positive spatial autocorrelation (at 1% significance level) of IMB was found. In the case of the second variable, the value of global Moran's I, converted to z-score (45.44), also led to rejection of the null hypothesis H0b and significant positive spatial autocorrelation (at 1% significance level) of IMT was found. Both variables prove statistically significant level of clustering in space. To identify spatial clusters, local Moran's I for each municipality was calculated. In the case of significant value of local Moran's I, types of spatial associations were established. The resulting maps of spatial associations for both variables show Figure 3 and 4.

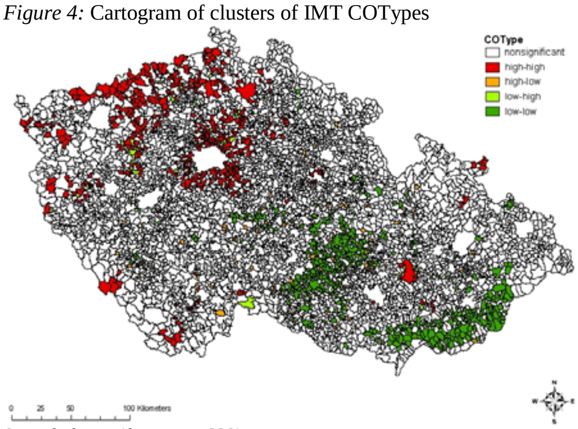
Figure 3: Cartogram of clusters of IMB COTypes



Own calculations (data source: CSO).

In the case of IMB, only clusters of high-high type of spatial association are accumulated, particularly in the adjacent municipalities around Prague. Other types of spatial association prove rather scatter character.

Figure 4: Cartogram of clusters of IMT COTypes



Own calculations (data source: CSO).

Regarding the IMT variable, positively spatially correlated high-high clusters and low-low clusters of municipalities are accumulated. High-high clusters of significantly high values are accumulated around Prague but in addition it creates a belt of northwest Bohemia. Interesting are also low-low clusters of significantly low values that stretch in two northeast-oriented strips on both sides of Brno.

Conclusion

This paper brought the analysis of current state of migration spatial differentiation of municipalities in the Czech Republic with focus on the identification of homogeneous clusters formed by municipalities with similar migration indicators (the intensity of migration balance and the intensity of migration turnover per 1000 inhabitants) using cluster analysis and analysis of spatial autocorrelation.

Cluster analysis led to the formation of four clusters for each variable. After that, the analysis of spatial autocorrelation was performed in two steps. Firstly, using the global Moran's I, significant positive spatial autocorrelation of both variables was demonstrated. Based on the type of spatial association, spatial autocorrelated clusters of municipalities for each of the variable were identified subsequently. Regarding the intensity of migration balance, only municipalities with statistically significant high values (migration profitable municipalities) were accumulated. In case of the intensity of migration turnover, only clusters of significantly high values (high-high types of spatial association) and significantly low values (low-low types of spatial association) were created.

To sum up conclusions of both analyses performed, it can be said that the Czech municipalities are generally more migration active than the Moravian. Municipalities around Prague form the most significant intersection of high values of both monitored variables, and therefore it is migration active and migration profitable region. Using analysis of spatial autocorrelation, it is possible to observe that these municipalities are in the case of both monitored variables positively spatial autocorrelated. Therefore migration profit (positive migration balance) and migration activity (positive migration turnover) of certain municipality in this region are significantly affected by the high profitability and high migration activity of surrounding municipalities. With the contribution of spatial autocorrelation analysis, it is easier to notice clusters of migration most passive municipalities on the border of Bohemia and Moravia and along the border with Slovakia. These municipalities are positively spatial autocorrelated and therefore migration passivity (low migration turnover) of certain municipality in this region is significantly affected by migration passivity of surrounding municipalities. Comparing the two approaches applied, spatial autocorrelation seems to be very appropriate to supplement or clarify the outputs of cluster analysis when modelling migration spatial differentiation of municipalities in the Czech Republic.

Literature:

1. Anselin, L.: Spatial Econometrics: Methods and Models. Kluwer Academic Publishers, Boston, 1988.
2. Anselin, L.: Local indicators of spatial association—LISA. Geographical Analysis, 1995, 27.2: 93–115.
3. Blažek, J.; Netrdová, P.: Can Development Axes be Identified by Socio-economic Variables. The Case of Czechia. Geografie, 2009, 114.4: 245–262.
4. Cliff, A. D.; Ord, J. K.: The Problem of Spatial Autocorrelation. In Scott, A. J. (Ed.): London Papers in Regional Science, Studies in Regional Science. Pion, London, 1969, 25–55.
5. Cliff, A. D., Ord, J. K.: Spatial Autocorrelation, Monographs in Spatial Environmental Systems Analysis. Pion, London, 1973.
6. Cliff, A. D.; Ord, J. K.: Spatial Processes, Models and Applications. Pion, London, 1981.
7. Czech Statistical Office (CSO): Public Database. [Online]. Available from http://vdb.czso.cz/vdbvo/maklist.jsp?filtr_uzemi=on%2C50%2Con&filtr_obdobi=2011&q_text=&kapitola_id=369&vo=null&q_rezim=1 [cit 14. 12. 2012].
8. Getis, A.: A History of the Concept of Spatial Autocorrelation: A Geographer's Perspective. Geographical Analysis, 2008, 40.3: 297-309.
9. Getis, A.; Ord, J. K.: The Analysis of Spatial Association By Use of Distance Statistics. Geographical Analysis, 1992, 24.3: 189-206.
10. Härdle, W.; Simar, L.: Applied Multivariate Statistical Analysis. Springer Verlag, Berlin-Heidelberg-New York, 2003.
11. Moran, P. A. P.: Notes on continuous stochastic phenomena. Biometrika, 1950, 37.1/2: 17-23.
12. Norušis, M. J.: IBM SPSS Statistics 19 Statistical Procedures Companion. Prentice Hall, NJ, 2012.
13. Nosek, V.; Netrdová, P.: Regional and Spatial Concentration of Socio-economic Phenomena: Empirical Evidence from the Czech Republic. Ekonomický časopis, 2010, 04: 344–359.
14. Novák, J.; Netrdová, P.: Prostorové vzorce sociálně-ekonomické diferenciace obcí v České republice. Sociologický časopis, 2011, 04: 717-744.
15. Spurná, P.: Prostorová autokorelace – všudypřítomný jev při analýze prostorových dat? Sociologický časopis, 2008, 04: 767-787.
16. Tobler, W. R.: A Computer Movie Simulating Urban Growth in the Detroit Region. Economic Geography, 1970, 46: 234-240.

Primary Paper Section: D

Secondary Paper Section: AO, BB, DE