

## EXAMINING NEIGHBOURHOOD EFFECTS IN REGIONAL INEQUALITIES OF HUNGARY: A GIS-BASED APPROACH FROM TOPOLOGICAL RELATIONS TO NEIGHBOURHOOD HETEROGENEITY

JAKOBI ÁKOS<sup>1</sup>

**ABSTRACT** – There is an increasing number of popular works, which explain regional inequalities of the economy and the society with spatial location. Techniques like spatial autocorrelation or spatial clustering became basic methods of mainstream examinations. These methods put geographical location, or more precisely the neighbourhood effects, in the focus of the researches. The typical research questions originate from Waldo Tobler’s classical law of geography, namely that “everything is related to everything else, but closer things are more closely related”. According to the research assumptions, the regions that are neighbours in geographical space are usually similar to each other also in socio-economic sense. For the examination of neighbourhood effects, GIS (Geographical Information Systems) may serve as a quite useful tool. By the application of GIS programmes, the topological relation of the examined objects can be straightforwardly defined. Additionally results of examinations with different methodological approach can be also easily compared. This paper, on the one hand, introduces both different methods of defining neighbourhood relations and the application of GIS in determining contiguity. On the other hand, it reflects how varied or same consequences can be drawn by the application of these techniques. In the analysis of neighbourhood effects, the most deterministic factors of the Hungarian regional inequalities are examined: the level of personal income and the unemployment rate. The examination points out where stable social and regional clusters were formed in Hungary, in the mirror of the different methodological approaches.

**Keywords:** methodology, GIS, topological relations, neighbourhood, spatial autocorrelation, Hungary

### INTRODUCTION

Spatial models are usually taking the advantage of the general characteristics of GIS techniques, namely that they are able to simplify the complicated reality of social spatial processes in a selective, inspiring and suggestive way. Models can be image-like, analogous or abstract, the last two can be found in many examinations of regional science. In the creation of spatial models, the application of GIS techniques or the usage of GIS-based spatial methods has become a widespread and unquestionably useful tool. The role of GIS is specifically important in the integration of topological information in a spatial model, for example, or by the application of certain spatial filter methods, which require the knowledge of describing the spatial location of objects.

Contiguity models or neighbourhood-relation models can be interpreted also as topology-based models, where objects, geographically connected and close to each other, are analysed concerning the assumed similarity of their socio-economic attributes. The essence of neighbourhood examinations is to examine empirically the thought that objects neighbouring in space are located also next to each other in socio-economic sense. It is about comparing location information or topology with socio-economic datasets.

While previously, neighbourhood information was collected by manual measures and interpretations of maps, nowadays, GIS programs make it possible to solve these tasks. It became also

---

<sup>1</sup> Assistant Professor, PhD, Department of Regional Science, Eötvös Loránd University, Budapest, Hungary. Address: Pázmány Péter sétány 1/c., Budapest, Hungary, 1117. E-mail: soka@ludens.elte.hu

easier to make different topology tables (e.g. rook-based or queen-based contiguity matrices, distance-based or K-nearest neighbour calculations, etc.). These techniques became rather useful also for geographic analysis of social and economic phenomena. An outstanding method based on topological information is spatial autocorrelation, which got diffused among researchers rapidly since the beginning of the application of GIS techniques, therefore, nowadays new and newer results appear on the importance of spatial location in social processes.

### **THEORETICAL DEFINITION OF NEIGHBOURHOOD RELATIONS**

It is a crucial question for analyses based on the examination of topological relations of spatial objects, namely which of the separate entities (regions, points, etc.) of a spatial dataset are declared to be contiguous and which are not. Additionally, what kind of algorithm is it based on? The differences in the interpretation of topological relationship may substantially modify the values of the topology tables (neighbourhood matrices, contiguity tables or spatial weight matrices), thus the results of the calculations directly based on the tables would have been modified as well. The system, in which the neighbourhood relations are determined, should be chosen therefore carefully and in a farseeing way.

The application of contiguity models appeared already quite long in professional circles. A wide literature deals with the topic in various approaches, from econometric modelling (e.g. Anselin, et. al. 2004, Getis – Aldstadt, 2004) to specifically GIS-based solutions (e.g. Goodchild, 1987, Wang 2006), however, the number of empirical analyses and local applications (for some Hungarian examinations see Dusek, 2004) have increased only in the last decade, when GIS became broadly and easily accessible for researchers.

It is clear that originated from the diversity of contiguity systems and neighbourhood connections, there can not be named any overall standing solution to define topological relations. A crossroad is faced since the very beginning, when the decision should be made whether it is about examination of point-, region- or line-formations. Definition of topological relation is perhaps the easiest in the case of regional objects (polygons, counties, provinces, etc.), when entities are contiguous, which assort with a direct border (sharing a common border with non-zero length). A rare problem could be only (particularly in case of regular or seemingly regular spatial formations) if two regions are connected not by a well defined borderline, but only in one point (node). These kind of relations can be found by chessboard-like regional formations (see e.g. some counties or states of the USA). Therefore, in order to solve the problem, methods originated from chess can be borrowed by the application of queen- or rook-contiguity (or rarely the bishop-contiguity). Another, however even more infrequent problem, can be the determination of contiguity by adjacent regions with indefinite borders. The real problem is the definition of the location of the border itself that can be done by fuzzy-algorithms (Wang – Hall, 1996), while the determination of contiguity comes only after it.

There is no direct neighbourhood relation between point objects, therefore topology and contiguity should be defined differently this time. The determination of topological relations of distinct points in space (e.g. cities, stations, etc.) can be typically traced back to the term of distance, when neighbouring are points, which are the shortest distance away from each other (nearest neighbours) or closer than a given distance (threshold distance neighbours), or perhaps topological neighbours are all the nearest neighbours, which have a total number of amount less than K (K-nearest neighbours, i.e. three nearest neighbours) (see also Jakobi – Jeney, 2008).

The contiguity relation system of point formations makes is easy to understand the topology of graphs and network formations. Network contiguity, that is probably the easiest to comprehend, exists when there is a direct connection (edge) between two of the nodes. The neighbourhood relations in special cases, namely in directed networks are not always symmetric: sometimes, when an edge leads from i to j, it is not sure whether an edge leads also from j to i.

New and interesting solutions are provided by methods, which apply techniques of defining neighbourhood relations of different types of formations in a combined or improved way. A typical and frequently used method is, for example, the conversion of regional objects to point formations, and

EXAMINING NEIGHBOURHOOD EFFECTS IN REGIONAL INEQUALITIES OF HUNGARY: A GIS-BASED APPROACH FROM TOPOLOGICAL RELATIONS TO NEIGHBOURHOOD HETEROGENEITY

then, the determination of neighbourhood relations within the point set by one of the above-mentioned techniques. Methodological problems or differences can be occurred, however, not in connection with the definition of contiguity, but rather with the interpretation of regional objects as point entities, since it is important to decide what points (with what position) will represent or substitute the original regional unit. Most often, a representative point or more precisely its geographical coordinates can be used to characterise a region (e.g. the coordinates of the capital of a county or region), in other cases for instance, the centroid of the region (polygon object) can be applied as a substitute point. This latter solution, however, may have a specific problem when it is about non-convex shaped territorial objects, where the calculated centroid occasionally may be located outside the original polygon. It happens also frequently, when point formations are transformed to regions in order to determine contiguity relations. In this case regional objects are created by the application of the well-known GIS methods of the extension of point data to polygons, namely by the methodology of the creation of Thiessen- or Voronoi- or Dirichlet-polygons. The created regional objects can be evaluated afterwards by the traditional principles of neighbourhood relations.

**TOPOLOGY IN GEOSPATIAL MODELLING**

While previously, matrices for the examination of topological relations had been created by manual measures or readings from analogous map information, in the age of geographical information systems, the tasks are solved already by software applications. GIS provide different kind of automated solutions and, in addition, it seems to be useful in neighbourhood operations with large calculation needs. In a substantial part of GIS calculations, contiguity matrices do not appear on the user's screen, only in the theoretical background of the processes. In GIS operations, users only have to define the method and the variables of the examination, and then the results can be expected in essence without any intermediate steps. Solutions based on this (e.g. different methods of spatial autocorrelation) often integrate GIS techniques already in the examinations. The methods of spatial autocorrelation belong to contiguity operations, therefore, geospatial programs, modules or extensions are unambiguously popular, being able to handle such tasks. This option is offered, for example, by some special extensions of ArcView, as well as by the programs of SpaceStat or Geoda, which are typically designed for social spatial analyses.

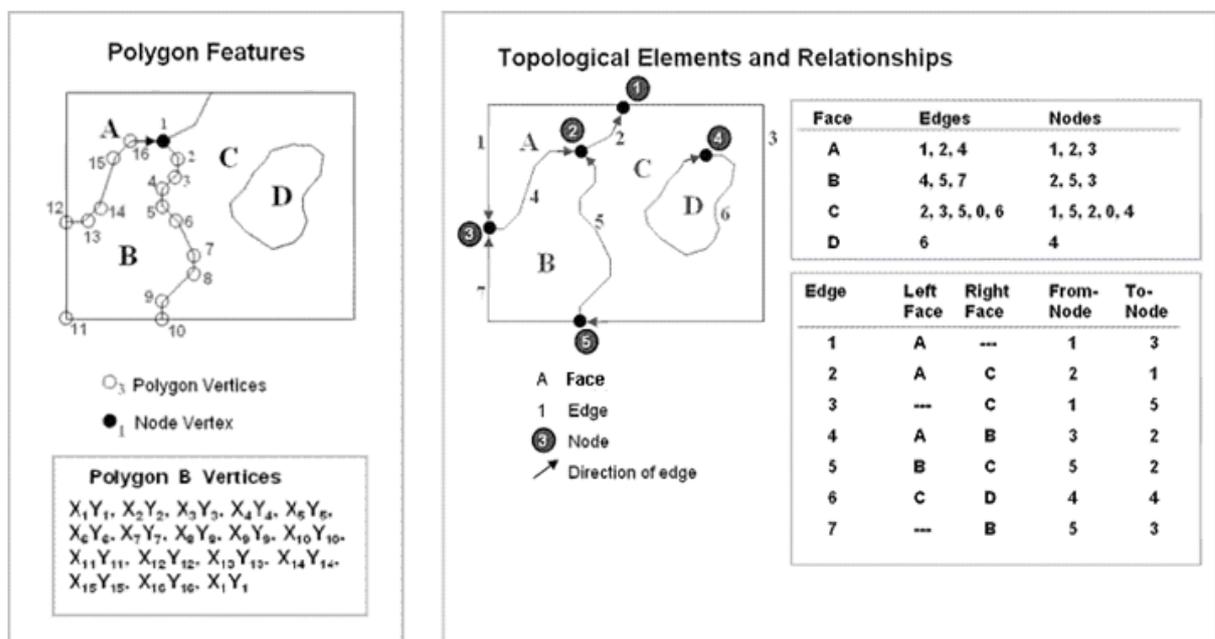
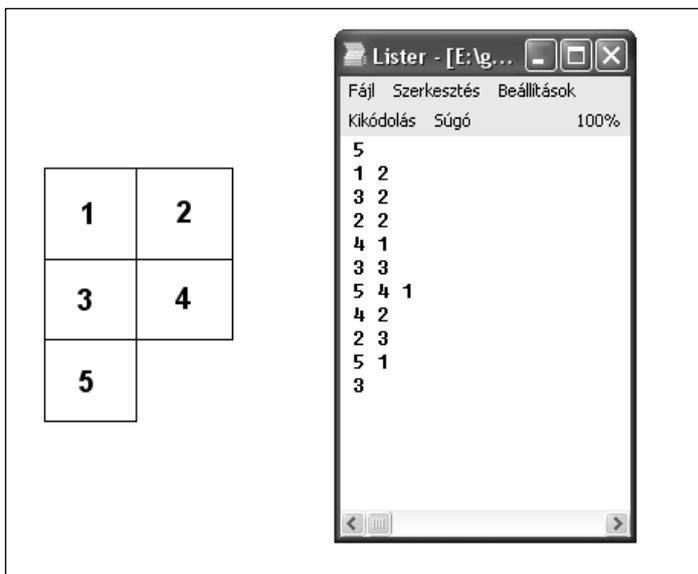


Figure 1. Definition of topological relations (data storage scheme of ArcView)

To manage topological information, the previously mentioned theoretical methods are widely applied in GIS modelling. A substantial question remains, namely how the data system could define and store the topological information in a way able to be examined. The storage of data is quite varied, depending on GIS software environment, mostly being built upon that whether the relations of points, lines or polygon objects should be stored. Most of the applications store topological information in the relation of sequences (vertices), by defining left and right face objects in relation to the directed edges (see Figure 1). They raise, however, difficulty in accessing information, when topology relation matrices themselves are standing in the focus of the examinations.

Geoda GIS spatial analytical software, which is in close connection with ArcView, is able to create codified topological table from latent contiguity information in order to eliminate the above-mentioned problem. Topological tables provided by Geoda are relatively compact; they do not contain any unnecessary information (the not neighbouring relations) (Figure 2). The application and interpretation is simple: the first row of the topological table gives the number of observed objects (5 in this case), the next represents the number of neighbours for object No.1 (2) and the following lines contain the ID-code of the neighbours (3 and 2).



**Figure 2.** Spatial formation of 5 regions and the connecting topological table created by Geoda GIS software

Somewhat different are the distance-based contiguity tables, which indicate pairs of points in each line of the table, as well as the measured distance between the points. The results are presented here also in text files (mentioning only the existing topological relations), which should be transformed if ordinary matrices were used.

The above-mentioned software makes it possible to create topological tables fitting to our previously drawn conceptual frames. Therefore, the method of rook-contiguity, queen-contiguity, threshold distance or k-nearest neighbour models can be used to create spatial weight matrices.

Models, which are basically relied on topological and location information, are well known for many topics in regional science, most of them can be alloyed with geoinformatics.

Known models are, for example, the different spatial autocorrelation models, the spatial filters, the spatial moving average models, the interpolated surface models, the trend surface models, the different kind of gravity models, as well as the spatial flow models. These relatively assorted techniques are all possible to integrate geoinformatic solutions; moreover, sometimes it is preferable to apply GIS techniques during model experiments.

For the evaluation of the similarity of close or adjacent spatial units, or for the examination of regional differences, or rather for the detection of proximity and spatial interaction, it is particularly worth to apply GIS-supported spatial autocorrelation analysis on socio-economic data. Based on the theory of the spatial shift or offset of data, it is examined how similar the values of regions are compared to their spatial neighbours. Local autocorrelation is hereby able to display regional macro-tendencies and also the specificities of the spatial structure of certain social indicators. The topology of the spatial units can be retrieved from the GIS database, while the autocorrelation coefficient could be determined by a simple spatial equation (Cliff – Ord, 1969, Bivand, 2010).

### GIS-BASED ANALYSIS OF NEIGHBOURHOOD HETEROGENEITY IN HUNGARY

The social and economic processes in recent years have induced a significant change in regional inequalities of Hungary. However, it is still a question whether the progressions are leading towards a new spatial structure of the country or are letting the differences stay unchanged. Dynamism and stability can be examined in parallel with the evaluation of time series of socioeconomic data; meanwhile, it should be determined whether steadiness or changes are in majority. Publications dealing with the regional development structure of Hungary all mentioned that after the political and economic transformation in 1990, the large scale spatial structure of Hungary started to change (Enyedi, 2004, Kovács, 2004), but in a couple of years it became stabilised, and other obviously slower processes started to take shape (Jakobi, 2004). Nowadays, the question arises again, whether the last two decades were significantly stable regarding patterns of regional differences, or measurable changes were or are taking place. In our research design, the question was led further to see the importance of neighbourhood effects, namely the spatial location happened to be unchanged or was altered in the last period. This was revealed in the determination of stable social and regional clusters in Hungary by the application of spatial autocorrelation models on time series of some socio-economic factors and naturally with the assistance of GIS applications.

GIS-based rook-contiguity topological tables and global, as well as local indices of spatial autocorrelation were used to determine the changes of the spatial patterns of some key factors in Hungary. Firstly, the tendencies of spatial clustering of *personal income per capita* were analysed, based on annual data of the Hungarian Tax and Financial Control Administration (APEH). The global spatial autocorrelation outcomes reflected a slight but significant rise in the Moran's I indices in the 90s, and a relative stability in the period afterwards (Figure 3.). The overall increase of the index, from a level of 0.4 in 1990 to a level of 0.6 after 2000, underlines the statement that neighbourhood effects have become more and more important in time.

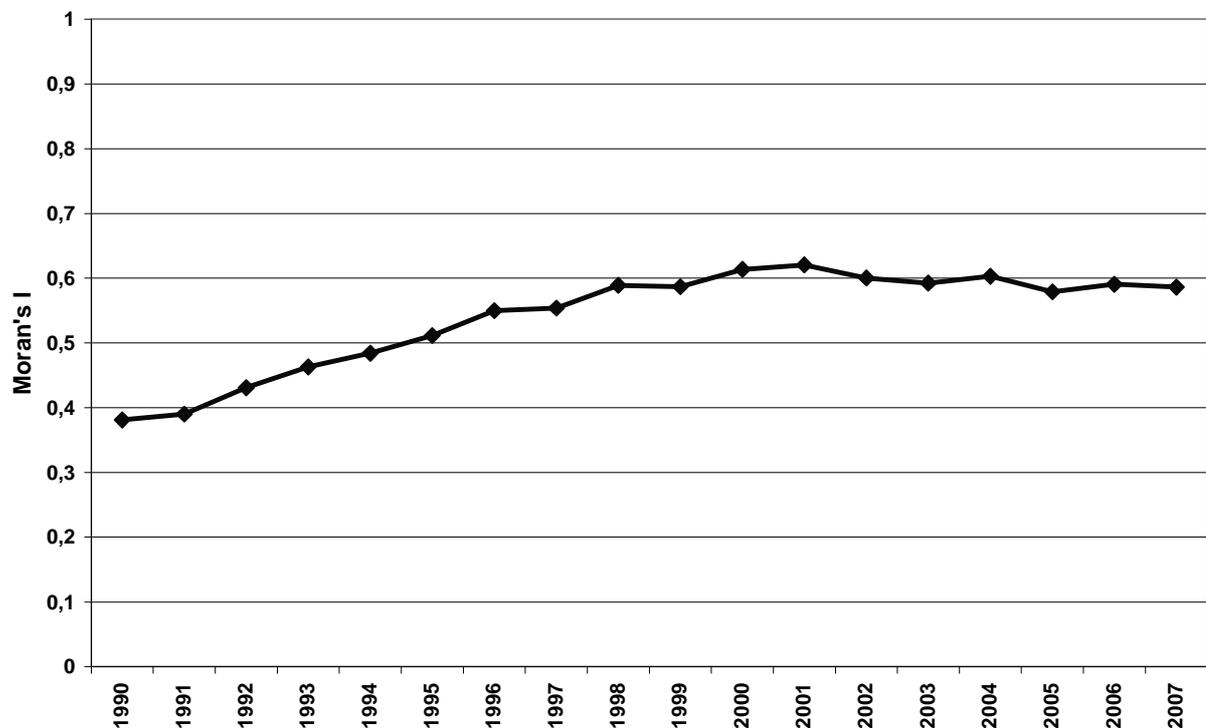
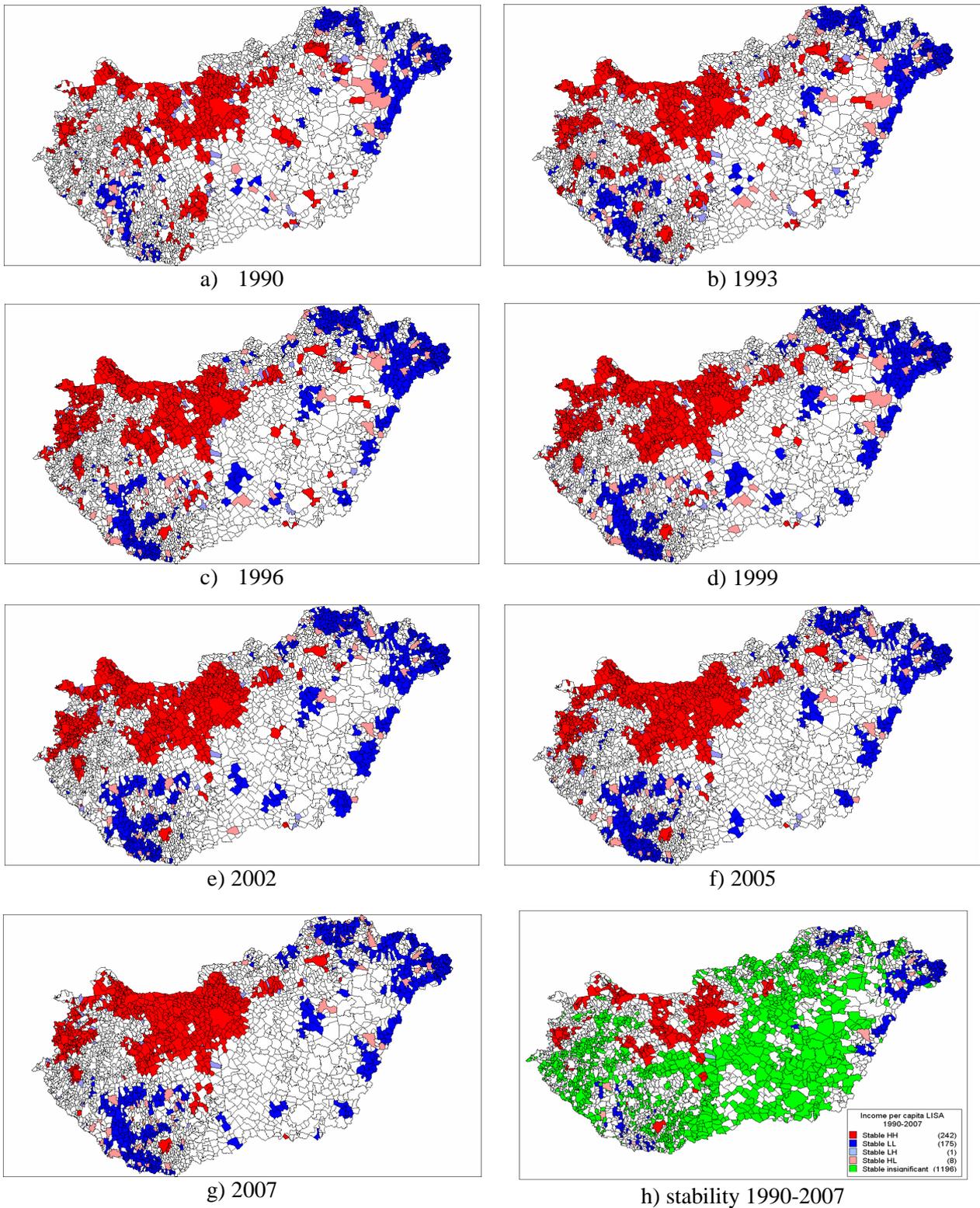


Figure 3. Change of Moran-I index of income per capita (LAU 1, 1990-2007)



**Figure 4.** Patterns and stability of local autocorrelation of income per capita in Hungary (LAU 1)

Note (definition of colours): Red – high local, high neighbour; Pink – high local, low neighbour; Light blue – low local, high neighbour; Dark blue – low local, low neighbour; White – insignificant. (Green on the stability map - stable insignificant).

EXAMINING NEIGHBOURHOOD EFFECTS IN REGIONAL INEQUALITIES OF HUNGARY: A GIS-BASED APPROACH FROM TOPOLOGICAL RELATIONS TO NEIGHBOURHOOD HETEROGENEITY

The spatial clustering of personal income per capita values has significantly increased in time, which was observed in the phenomenon that spatial units (in our examinations LAU 1 spatial entities, namely Hungarian settlements) neighbouring each other became more and more similar to the others during the observed period. Therefore, global autocorrelation values represent overall tendencies in the changes of the importance of topological relations.

The maps in Figure 4 show the spatial patterns and stability of local autocorrelation of income per capita in Hungary from 1990 to 2007, in selected years (LISA – Local indicator of spatial association). At a first sight, the maps seem to be similar, which may allow us to draw the conclusion that the spatial structure was mainly unchanged in the last two decades. The most developed, high income settlements are still located in the north-western part of Hungary and around Budapest, while the lowest income areas can be found primarily in the eastern peripheries of the country. However, this is only at a bird’s-eye view since there were significant and strong processes of clustering in the observed period. To see it more precisely, at the beginning of the 90’s, the spatial clusters of regions with high local and high neighbour values (represented in red colour) or the ones with low local and low neighbour values (dark blue colour) appeared on the maps not as concentrated as in 2005 or in 2007. By the end of the period, the number of settlements – which could have been significantly classified in the high-high or in the low-low groups – had stably and substantially increased. This obviously leads to the conclusion that the local environment of a settlement has an increasing effect on the locality itself, in other words, neighbourhood effects have increased, while neighbourhood heterogeneity has seemingly decreased in time.

By comparing all the maps in the examined period, the real stable segments of the country were determined, namely the settlements that were in the same category on all maps during the period (Figure 4h). The results of the analysis – also by using GIS – show that only 242 settlements had high-high values, and 175 had low-low values during the entire period. They are the real stable parts of the country, with favourable or steadily unfavourable income values both at the observed location and in the neighbourhood. Others were changing categories over time (white colour) or turned out to be constantly insignificant in all calculations (green colour).

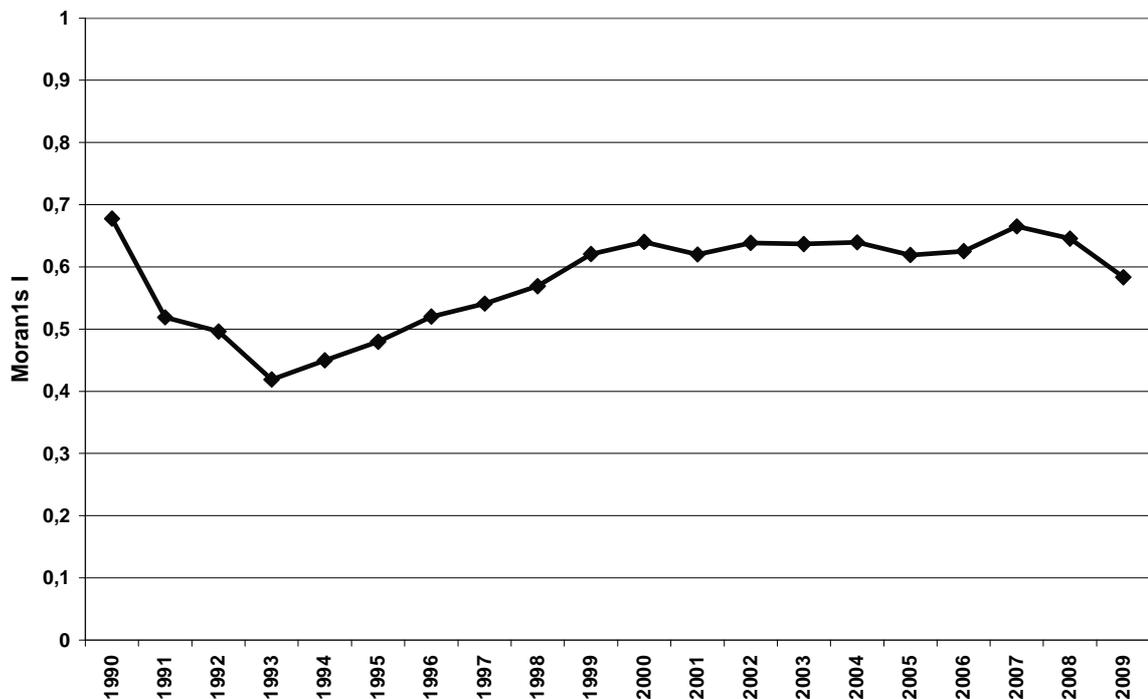
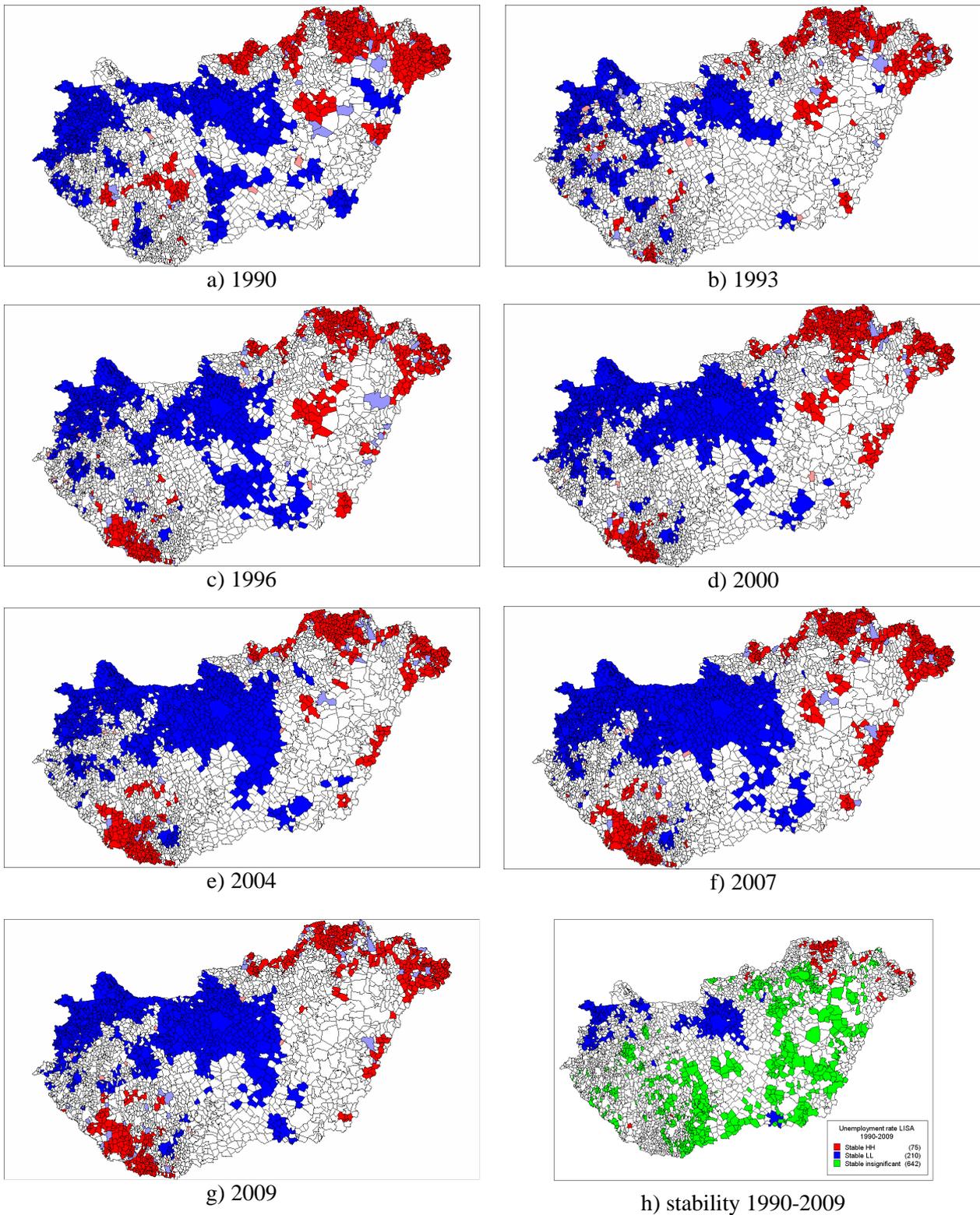


Figure 5. Change of Moran-I index of the unemployment rate (LAU 1, 1990-2009)



**Figure 6.** Patterns and stability of local autocorrelation of unemployment rate in Hungary (LAU 1)

Note (definition of colours): *Red* – high local, high neighbour; *Pink* – high local, low neighbour; *Light blue* – low local, high neighbour; *Dark blue* – low local, low neighbour; *White* – insignificant. (*Green* on the stability map - stable insignificant).

## EXAMINING NEIGHBOURHOOD EFFECTS IN REGIONAL INEQUALITIES OF HUNGARY: A GIS-BASED APPROACH FROM TOPOLOGICAL RELATIONS TO NEIGHBOURHOOD HETEROGENEITY

The second indicator applied in the present assessment was *the rate of unemployment* provided by the Public Employment Service of Hungary (ÁFSZ). The dataset was available for a longer period, from 1990 to 2009, which made it possible to reveal some effects of the last years' financial and economic crisis, as well. The outcomes of the global Moran's I analyses mirrored some significant changes in the importance of neighbourhood effects, however, somewhat differing from the income results (Figure 5). In the beginning of the 90s, when unemployment rapidly increased in almost all parts of the country due to the so-called transformation recession, the importance of neighbourhood effects decreased noticeably. At that time, it was less important for settlements what happens in their neighbourhood, since it was basically independent from the topological relation. After 1993, the tendency has changed and the importance of neighbourhood relations slowly and significantly started to increase. In other words, the processes of spatial clustering began to improve, and the neighbourhood effects started to impinge on spatial distribution of unemployment. The value of the global Moran's I index remained around 0.65 between 2000 and 2008, when a remarkable decrease appeared. The effects of the global economic recession also influenced the labour market of Hungary in 2008-2009, so that the importance of topological relations rapidly decreased in the coming years.

Observing the maps of spatial clustering of unemployment in the selected years, the above-mentioned tendencies can be indirectly traced (Figure 6). However, the local autocorrelation patterns give more details on the processes behind. In the first couple of years, spatial clusters (high-high and low-low areas) were located in many parts of the country, although some major cores had already appeared. By the end of the decade, clusters became even more compact, reflecting that the spatial structure of unemployment became relatively stable in geographical terms. This compact character seems to be slightly in dissolution from 2007 to 2009, primarily because the number of settlements with insignificant neighbour effects – mostly in the western part of the country – began to increase. Nevertheless, the geographical stability of the favourable and unfavourable areas is still unquestionable, although to a bit smaller extent than in the case of income data. There were 210 settlements in the country, which had always low unemployment rates, both locally and in their neighbourhood (Figure 6h). These areas can be determined very well on the map around Budapest, in Szeged, and in a well-defined cluster in Northwest Hungary. On the other hand, 75 settlements were found where the unemployment rate was high all the time, both locally and in the neighbourhood. These stable clusters are basically located in the North-Eastern peripheral regions of the country.

Until now, it was already known that the maps of unemployment reflect just the opposite spatial pattern than those of the income maps, but the temporal analysis of spatial autocorrelation indices and patterns revealed that there are slight differences between the variables of income and unemployment in the extent and the allocation of constantly stable clustering areas. However, all the results reflected that the neighbourhood heterogeneity is changing in time, and geographical location or topological environment has an explanatory role of the regional inequalities of Hungary.

## REFERENCES

- ANSELIN, L., FLORAX, R., REY, S.J. (Eds.) (2004), *Advances in Spatial Econometrics: Methodology, Tools, and Applications*. Springer-Verlag, Berlin.
- BIVAND, R. (2010), *The Problem of Spatial Autocorrelation: forty years on*, Available at: <http://cran.r-project.org/web/packages/spdep/vignettes/CO69.pdf> (accessed 15 January 2011.)
- CLIFF, A.D., ORD, J.K. (1969), *The Problem of Spatial Autocorrelation*, in: *Papers in Regional Science*, 1, pp. 25–55.
- DUSEK, T. (2004), *A területi elemzések alapjai* [Fundamentals of regional analysis], Regionális Tudományi Tanulmányok, 10. ELTE Regionális Földrajzi Tanszék – MTA-ELTE Regionális Tudományi Kutatócsoport, Budapest.

- ENYEDI, GY. (2004), *Processes of regional development in Hungary*, in: Gy. Enyedi, I. Tózsá (eds.) *The Region. Regional Development, Policy, Administration, E-Government*. Akadémiai Kiadó, Budapest, pp. 21–34.
- GETIS, A., ALDSTADT, J. (2004), *Constructing the spatial weight matrix: using a local statistic*. *Geographical Analysis*, 36, pp. 90-104.
- GOODCHILD, M. (1987), *A spatial analytical perspective on geographical information systems*. *International Journal of Geographical Information Systems*, 1, pp. 327-334.
- JAKOBI, Á. (2004), *Traditional and New Causes of Regional Inequalities in Hungary*, in: Hacker, R. S. – Johansson, B. – Karlsson, C. (eds.) *Emerging Market Economies and European Economic Integration*, Edward Elgar Publishing Ltd., Cheltenham, UK, pp. 160-185.
- JAKOBI, Á., JENEY L. (2008), *A szomszédsági mátrix problematikájának megoldási lehetőségei – elmélet és gyakorlat* [Possible solutions of the neighbourhood matrix problem – theory and practice], *Területi Statisztika*, Vol. 10 (47.), 3, pp. 295-304.
- KOVÁCS, Z. (2004), *Socio-economic transition and regional differentiation in Hungary*, in: *Hungarian Geographical Bulletin*, Vol. LIII, No. 1-2, pp. 33-49.
- WANG, F. (2006), *Quantitative methods and applications in GIS*, Boca Raton, Taylor and Francis Group.
- WANG, F., HALL, G.B. (1996), *Fuzzy representation of geographical boundaries in GIS*, in: *International Journal of GIS*, 10 (5.), pp. 573–590.