38th Congress of the European Regional Science Association

28 August-1 September 1998

Empirical analysis of spatial location of activity: a proposal of using spatial association statistics

Esther Vayá[1]

Research Group: Anàlisi Quantitativa Regional (AQR)

Department of Econometrics, University of Barcelona

Barcelona, Spain

evaya@eco.ub.es

ABSTRACT:

From the paper of Krugman (1991), several works have analysed the consequences of the integration process on the activity location. In this sense, Puga and Venables (1996) among others, have derived a relation between integration and agglomeration with an inverse U-shape. In order to check this hypothesis, it is habitual to compute some of the traditional concentration indexes. However, this kind of indexes considers each region as a isolated economy, without any links among others. This paper proposes the use of some spatial association tests in order to shed light on the consequences of an integration process overcoming the limitations showed by the mentioned indexes (and also the measurements of inequality). First, we discuss, from a theoretical point of view, the utility of these tools to get the goal of the paper. So, we analyse the behaviour of some spatial association statistics for different theoretical spatial distributions of the product. Second, we apply this analysis to the EU case. In this sense, we compare the evolution showed by the GDP pc among the European regions during 1975-1992 with the theoretical distributions defined in the first part in order to know the trend of the spatial distribution of the activity after the deepening in the integration process.

Keywords: Spatial dependence; activity location; concentration/diffusion process; EU regions.

[1]The paper presents results from a project developed in the AQR Reasearch Group.

1. Introduction

In the last years, as a consequence, among other factors, of the apparition of the New Economic Geography (Krugman 1991a, b), the interest in the activity location and its concentration has grown. In this sense, the models included in that subject[1] explain the final location of the industry as a result of the fight between centripetal[2] and centrifugal[3] forces (see Krugman and Venables, 1995, 1996; Venables, 1996). Moreover, an important number of models support the existence of a relation with U-inverted shape between the degree of integration and the level of concentration of the activity.

Regarding this last point, at a empirical level, several papers have analysed the effects of the European integration process on both the location of the industrial sector (Greenaway and Hine,1991; Fluvia and Gual, 1993; Molle, 1996; Brülhart and Torstensson, 1996; Amiti, 1997) and the evolution of regional disparities (Suarez-Villa and Cuadrado-Roura, 1993; Esteban, 1994; Neven and Gouyette, 1994; Armstrong, 1995; Molle and Boeckhout, 1995). In order to get this goal, it is habitual to compute, by one part, traditional indexes of sector concentration and regional specialisation and, by the other, measurements of inequality as the Gini index or the sigma-convergence. However, both groups of indexes show an important drawback: they do not consider explicitly the space. So, although the coefficient of regional specialisation describes the situation of one region *i* (in terms of its productive structure), this measurement does not take into account its spatial location, considering each region as an isolated unit, without any link with its neighbouring regions. Moreover, although the inequality indexes inform about the dispersion of the sample, a same value of those are compatible with different spatial distributions of the income. So, for instance, the Gini index or the sigma-convergence will not detect a situation characterised by a spatial relocation of the most dynamic regions if the dispersion of the variable does not change.

In addition, any of both groups allow us to know if the analysed variable is randomly distributed through the space or, on the contrary, there is a significant trend to clustering of similar or dissimilar values of the variable. In this sense, it is necessary to point out that it is more likely to find a scheme of spatial dependence when the analysis is done for a low level of aggregation (as the regional case). Besides, by means of the mentioned indexes it is not possible to detect neither pockets of regions with a higher or lower value of the variable in comparison with the expected one nor possible non spatial stationarities.

Keeping in mind these ideas, and in order to overcome the drawbacks showed by the traditional indexes, this paper proposes the study of topics related to the location of the activity and the regional disparities from a spatial perspective, applying different spatial dependence and association tests (not used before for a similar goal). This communication is organised as follows. In section two, there is a brief description of the spatial dependence test that will be used in the paper, analysing at the same time their utility to get the goal proposed. In section three, the objective is to check the sensitivity of these contrasts against different spatial distributions of the studied variable. In order to get it, a simulation exercise is done, where firstly we generate different samples of one variable called “level of activity” that incorporate different hypothesis about its spatial distribution. Secondly, we compute the spatial dependence tests for each one. Moreover, and given that we relate the generated spatial distributions with the scenarios expected by the New Economic Geography models before, during and after the integration process, it is established a possible evolution of the spatial tests during the integration. In section four, we apply the proposed analysis to the case of the GDP per inhabitant for the European regions during the period 1975-1992. Finally, in section five we conclude.

2. Spatial dependence tests

2.1. Global spatial dependence test

We consider that there is spatial dependence or spatial autocorrelation when the values observed at one location (for instance, one region) depend on the values observed at its neighbouring regions. This fact means that it is not possible to change the location of the values of one variable without affecting to the information in the sample (Anselin, 1988; Anselin and Florax, 1995). The existence of a scheme of spatial autocorrelation in one variable could be check by means of, among others, the Moran’s I test (Moran, 1948) that has the following expression:

$$I = \frac{N}{S}\frac{\sum_i \sum_h w_{ih} z_i z_h}{\sum_i z_i^2}$$

where N is the number of observations, w_{ih} is the element of the spatial weights matrix W that express the potential interaction between two regions *i* and *h*, S is the sum of all the weights (all the elements in the weights matrix) and z_i represents the normalised value of the analysed variable in region *i*. About the weights w_{ih}, although there is not a consensus about the correct specification of W, it is habitual to use a contiguity criterion, that is, w_{ih} will be 1 if regions *i* and *h* are neighbours and 0 otherwise. For our analysis, we have chosen this type of matrix W, standardising it by rows (so that each row sums 1).

After standardisation of this statistic,[4] a significant and positive value indicates a trend to clustering of similar values of the variable in the space (positive spatial dependence). So, for example, if the variable was the level of activity, a scheme of positive spatial dependence would show that neighbouring regions concentrate similar levels of activity, revealing that the location of this activity follows a clear spatial pattern. On the contrary, when the test is significantly negative, the trend is to clustering dissimilar values in neighbouring locations (negative spatial dependence). Applying it to the previous example, one explanation could be that the strength of the centripetal forces in one region prevents the diffusion of the industry to its neighbours.

The no significance of the Moran's I implies to accept the null hypothesis: the non existence of spatial autocorrelation and, then, the prevalence of a randomly distribution of the variable through the space.

Finally, although we have defined a contiguity matrix, it would be interesting to know whether these schemes of autocorrelation would be the same among non strictly adjacent regions. One way to get it is by means of the spatial correlograms (Cliff and Ord, 1973, 1981), that is, a figure where appear the values of the Moran's I standardised for different orders of contiguity matrix.[5]

2.2. Local spatial association tests

As we can see, the previous statistic of spatial autocorrelation is centred in the study of a scheme of global spatial dependence, taking into account all the regions of the sample. However, it, jointly with other spatial dependence tests, are not sensitive to situations where there are clusters concentrated in specific areas of the analysed territory. In this sense, this kind of statistics are not able to identify specific significant hotspots or macro-regions where the activity is highly concentrated (or the opposite situation) in comparison with its expected value. Considering this, we have computed two tests of local spatial association: the local-Moran's I_i (Anselin 1995) and the G_i^* (Getis and Ord 1992). Both statistics indicate to a what extent each region is surrounded by regions with high or low values of the analysed magnitude. The local-Moran statistic, I_i can be defined as:

$$I_i = \frac{z_i}{\sum_i z_i^2 / N} \sum_{h \in H_i} w_{ih} z_h$$

where z_i is the observation for region *i* in deviations from the mean and H_i is the set of neighbouring regions to region *i*.[6] In this case, a significant positive (negative) result indicates the existence of a cluster of similar (dissimilar) values of the variable surrounding region *i*.

However, in order to know if the clusters of similar values concentrate high or low values of the analysed magnitude, it is necessary to compute also the second statistic of local spatial association. This statistic only can be applied to positive variables and it is defined as:

$$G_i^* = \frac{\sum_{h=1}^{N} w_{ih} x_h}{\sum_{h=1}^{N} x_h}$$

Once standardised (distributing by a normal N(0,1)), a positive significant result indicates the existence of a spatial cluster of high values of the variable, whereas a negative one indicates a clustering of low values.

So, it could be distinguished three possible results (without considering the case of non significance of none of the local test). First, when one region has a positive and significant value of both test, I_i and G_i^*, showing the existence of a cluster of similar and high values of the variable. Second, when one region has a positive significant value of the I_i but a negative G_i^*, showing the presence of a cluster of similar and low levels of activity. The last case is that the I_i is negative and significant, giving information about the existence of a cluster of dissimilar values of the variable because region *i* concentrate high levels of activity meanwhile its neighbours show the opposite situation (or vice versa). In any case, it is necessary to point out that, for instance, a positive and significant value of the I_i in one region implies that region *i*, jointly with its neighbours, are concentrating higher (or lower) levels of activity than the expected ones.

Finally, it is worthily noting that certainly the view of a map that depicts the analysed variable could give us an intuitive idea about, for instance, the existence of a trend to clustering similar levels of activity on the space or the presence of some regions that concentrate high or low levels of this variable (making not necessary the spatial analysis proposed here). However, the information derived from the map is subjective and highly dependent on, for example, the selected number of intervals to depict the variable in the map. On the contrary, the spatial dependence tests give objective and statistic criterions to confirm both the presence of a spatial association scheme in the distribution of the variable and the existence of pockets of non spatial stationary.

3. Generating different spatial distributions of activity

3.1. Hypothesis about the spatial distribution of activity

Once defined the global and local spatial dependence tests, in this section we will analyse their behaviour against different spatial distributions of the level of activity. For that, we have generated one hundred sample for each of the four following hypothesis: (1) *equal distribution;* (2) *core-periphery scheme*; (3) *monocentre*; (4) *polycentre* (all the samples have the same average value). Before explaining these hypothesis, it is necessary to point out that the European regions (specifically, 108 regions) have been used as a spatial units of reference in order to identify the simulated sample with a real territory. Concretely, we have selected a NUTS II level of aggregation for Portugal (5 regions), Spain (17), France (21), Italy (29), Greece (13) and NUTS I level for UK (11), Belgium (3), Holland (4) and German (11). In addition, Denmark, Ireland and Luxembourg are considered as a single region.

Regarding the defined spatial distributions of activity, the first hypothesis, *equal distribution*, describes a situation where the activity is homogeneously distributed through the space as a consequence, among other possible factors, of the existence of a similar regional dotation of production factors, the presence of very high transport costs (as, for instance, in the first moments of an integration process) or the domain of an autarchy situation.

In the second case, we suppose that there is a Core-Periphery scheme with a clear geographical division. So, the Northern regions (regions of France, UK, Germany, Holland, Luxembourg, Belgium and the North of Italy) are concentrating most of the activity,[7] finding the opposite situation in the Southern regions (in fact, the average value of the variable in the South is a 45% of the average in the North).

In the third hypothesis, we have distinguished three cases: (3.1), (3.2) and (3.3). Concretely in case (3.1), again there is an important concentration of activity but, opposite to the previous case, it is located only in a little group of neighbouring regions (16 regions), specifically in some regions of the South of Germany, North of Italy and East of France, being randomly distributed the rest of the activity in the other European regions (the average of the variable in the regions that are not included in the monocentre is a 50% of the average in the monocentre). However, and following Puga (1996) and Puga and Venables (1996), we have incorporated the assumption of a diffusion process of the activity. Specifically, in case (3.2), we have supposed that, as a consequence of the presence of diseconomies of scale in the core regions, in a first step the activity spreads towards the neighbouring regions of the monocentre (the average of the variable in the adjacent regions to the monocentre is a 90% of the average in the monocentre, meanwhile the average of the rest of the regions is 60% of the latter). In a second step (case 3.3), this process affects also to the neighbours of second order of contiguity of the monocentre, being randomly distributed the activity in the rest of the regions. So, a hierarchical spatial distribution appears, existing more concentration in the core but with a decay trend with the distance (the neighbours of second order of contiguity have an average equal to a 75% of the average in the monocentre, while the average for the rest of the regions is the 50% of the latter).

However, and in order to know if the spatial dependence tests are sensitive to both the spatial location of the monocentre and its number of neighbouring regions, we have relocated it. So, first, the whole of the monocentre (again 16 regions) has been situated in France, being absolutely surrounded by regions that are in the sample (in the previous case it does not happen the same). The second hypothesis has been that the monocentre is in a peripheral location, specifically in Portugal and in some of its neighbouring Spanish regions (16 regions also). The main difference of this second hypothesis in comparison with the two previous cases is the small number of neighbours of first and second order of contiguity to the monocentre, leading to that the diffusion process affects only a few spatial units.[8]

In the last case, (4) *polycentre*, we have considered that there is a relevant concentration but, opposite to the previous cases, now the activity is highly located in some regions that are not close but scattered through the territory (the location of these core-regions has been randomly chosen). Moreover, it has been incorporated the hypothesis that, as a consequence of the strong attraction of the core-regions, their adjacent regions show smaller level of activity than the rest of the regions (the adjacent neighbours of each core-region have an average value equal to the 50% of the average of their core-regions, while the rest of the regions have an average equal to the 70% of the latter).

3.2. Results of the simulation exercise

3.2.1 Traditional analysis

An example of each type of spatial distributions is in figures 1.a, 1.e, 2.a, 2.e, 3.a and 3.e. Moreover, an analysis of inequality has been done. [9] For that, we have depicted the density function[10] corresponding to each hypothesis, including the value of the Gini and the Theil indexes.

From figure 1.b (corresponding to the case 1), it could be detected the existence of one only mode with a high concentration of its probability, showing a scarce level of regional dispersion. Absolutely different is the density function for the case 2 (figure 1.f). Now, two different modes are distinguished, showing the existence of a relevant heterogeneity and so high inequality in this distribution.[11] When the third spatial distribution is analysed, the results are summarised in figures 2.b, 2.f and 3.b. As we can see, when there is not a hierarchical distribution around the monocentre, there is one significant mode with an important concentration of the probability that included the most of the regions in the sample. In addition, there is an outstanding tail on the right of this mode, reflecting the existence of a group of few regions with a higher concentration of activity. Once the diffusion process has been triggered, the concentration of probability in the mode seems to decrease, widening the tail detected before. Finally, the density function for the case 4 (figure 3.f) shows an important mass of probability concentrated around the average value of the variable. However, a relevant tail on the left of the mode appears, showing the existence of a group of few regions with low levels of activity.

Regarding the Gini and Theil indexes, the distribution in case 1 has the less values of them, showing the existence of a rather even distribution of the activity. On the contrary, the distribution with a core-periphery scheme shows the highest values of both indexes, reflecting highly significant degree of dispersion. Concerning the case of a monocentre (without and with a hierarchical distribution of the activity) and polycentres, the results of both are very similar among them (although the dispersion is lightly higher in case 4 than in case 3).

As we can see from these results, although the traditional analysis of inequality gives us information about the degree of the dispersion, it does not shed light on the spatial distribution of the variable in each case. So, for instance, from the figure 1.f, we are not able to know if the regions that are included in each mode are close in the space or not. Moreover, meanwhile the value of the indexes for the cases 3 and 4 are very similar (although there are some differences in their density functions), the location of the activity answers to absolutely different criterions in each case: group of neighbouring regions concentrating most of the activity in case 3, and some regions widespread through the space, concentrating most of the activity and attracting the activity of their adjacent regions in case 4. This result is specially important if we consider the different consequences of an improvement in the situation of one region located in the monocentre in case 3 and of one core-region in case 4 over their neighbours respectively.

Taking into account these ideas, the spatial analysis of each of the generated spatial distribution seems to be justify.

3.2.2 Spatial analysis

The results of the Moran'I standardised for weights matrix up to ten order of contiguity are depicted in the spatial correlograms in tables 1, 2 and 3. Furthermore, the results of the local spatial association tests are summarised in different maps in these tables. As we can see, only the regions that have shown a significant (at a 10% level of signification) and positive value of the I_i test are marked, incorporating a positive sign (cluster of similar high levels of activity) or negative (cluster of similar low levels of activity) depending on the sign obtained for the G_i^* test. Besides, we have marked those regions that have a negative and significant value of the I_i test, showing the presence of a cluster of dissimilar levels of activity among this region and its neighbours.

First, as we can see in figure 1.c, the spatial correlogram associated with the hypothesis of even distribution shows the non existence of any scheme of spatial dependence for any order of contiguity (in any case the two discontinuous lines are exceeded). Moreover, the scheme detected at a global level is reproduced at a local level (figure 1.d) since any region has a significant value for the I_i test.[12]

On the contrary, the spatial correlogram for the Core-Periphery hypothesis (figure 1.g), shows a completely different situation: positive and high spatial dependence for the first orders of contiguity (from 1 up to 5) and a clear scheme of negative spatial dependence for the last orders (from order 7), maintaining its intensity despite the distance. At a local level (figure 1.h), some clusters appear, situated in some regions of the South, with lower levels of activity than the expected ones. So, from this result it seems that those regions which are far from the general trend showed by the others will have likely more significant values of the local tests (in this sense it is necessary to point out that 63 regions are considered as Northern regions and 45 as Southern).

In 2.c, 2.g and 3.c appear the spatial correlograms associated with the case 3 where exists a monocentre with a high concentration of activity, without (figure 2.c) and with (figures 2.g and 3.c) a hierarchical distribution of the rest of the activity in the other regions. Two main conclusions could be derived from these figures. First, and opposite to the previous case, the shape of the correlogram is convex for the higher orders of contiguity. That is, a positive and significant spatial dependence scheme appears at the firsts orders of contiguity but, after some orders (where there is not spatial autocorrelation) a negative spatial dependence domains, arriving at a minimum value to disappearing again at the ten order of contiguity (negative spatial dependence decreasing with the distance). So, meanwhile the negative spatial dependence remains (even increases) at the last orders of contiguity when there is a clear North-South division, in case of a monocentre situated in the centre of Europe, a positive spatial autocorrelation wants to reappear among very far regions. The second conclusion is related to the consequences of the diffusion process. So, after comparing the three figures, it seems that as the distribution of activity is more hierarchical, the number of orders of contiguity with a positive spatial dependence increases: up to the third in figures 2.c and 2.g and up to four in the last case. This result is reflecting the bigger similarities among the regions in the monocentre with their neighbours of first and second order of contiguity as the diffusion process advances. In addition, as the distribution is more hierarchical, it is necessary less distance to find a negative and significant spatial dependence scheme among regions.

Regarding the local analysis (figures 2.d, 2.h and 3.d), it seems that, again, those regions that are more different than the general trend have significant values of the local test (remember that only 16 of the 108 regions are in the monocentre). So, as we can see in figure 2.d, only the regions that are included in the monocentre present positive and significant valued of the I_i and G_i tests, showing the existence of clusters of similar and high levels of concentration of activity. Furthermore, it is interesting to note that the diffusion process leads to an increase in the number of clusters with high level of activity, spreading towards the neighbouring regions of the monocentre. However it is necessary to point out that in the case 3, although there is a diffusion process, the regions in the monocentre maintain their relevant role, concentrating most of the activity (even in the case defined in 3.3). For that reason, in the local analysis, the diffusion makes that the clusters detected at the beginning remain. In contrast, the results obtained would be different in case the diffusion process led to the lost of the privileged position of the core-regions (obtaining a more equilibrate distribution of the activity). In this last case, it would be likely that, at a local level, the clusters detected before disappeared after the diffusion process, losing significance also the spatial dependence scheme at a global level.

As it was mentioned before, in order to know the sensitive of the spatial dependence tests to both the location of the monocentre and the number of its neighbours, we have replicated the analysis moving the place of the monocentre, firstly to France and secondly to Portugal and some Spanish regions (for space reasons, we have not included the obtained results). Regarding the global spatial autocorrelation tests, in the first case, although the results and conclusions of the spatial correlogram are similar than those obtained in figures 2.g and 3.d, two main differences are detected. First, the concave (convex) shape at the lower (higher) orders of contiguity is more strong, being larger the values of the Moran's I for each matrix in comparison with the prior case. Moreover, the positive and significant spatial dependence returns at the 10 order of contiguity. In addition, there is a light less number of orders of contiguity with a non significant scheme of spatial dependence, showing more strong changes in the sign of the spatial autocorrelation. This result could be explained by the bigger number of neighbouring regions for superior orders of contiguity when the monocentre is placed in the geographical middle of Europe, doing more dramatic the contrast among regions with high and low concentration of activity.[13] When the monocentre is in one extreme of Europe, two main differences could be drawn. First, meanwhile in previous case the spatial correlogram had a stronger convex shape for the higher orders of contiguity than case 3, now the shape of the spatial correlogram is close to that derived for the Core-Periphery scheme, given that the negative spatial dependence losses intensity very. This result could be explained because the monocentre is not surrounded in the most part of its perimeter by regions included in the sample, appearing a geographical division of the activity. The second difference is related to the consequences of the diffusion process. Due to the scant number of neighbours of first and second order of contiguity of the monocentre, the spread of the activity has few effects on the spatial correlogram.

Regarding the results of the local spatial analysis, the changes in the location of the monocentre are reflecting in similar changes in the location of the clusters detected. So, again the regions that are included in the monocentre in each case show significant values of the local spatial association tests, reflecting the existence of clusters with similar and high concentration of activity. Besides, the diffusion process has the same consequences that those detected in case 3, that is, the spread of the clusters towards the neighbouring regions of the monocentre.

With regard to the case 4, the spatial correlogram associated is depicted in figure 3.g. As we can see, and opposite to the foregoing cases, there is a significant negative spatial dependence among adjacent regions, reflecting the effect of the core-periphery schemes included in this last spatial distribution of the activity (clear trend to clustering dissimilar levels of activity in the space). In addition, the positive and significant spatial autocorrelation detected for the second order of contiguity could be explained because of all the core-regions present similar levels of activity among them. Furthermore, and regarding the local analysis, each of the core-periphery schemes is reflected in the apparition of a cluster with a negative and significant value of the I_i test, showing the opposite situation between each core-region and its neighbouring regions (note the total correspondence between the regions with high levels of activity in figure 3.e and the regions that have a significant I_i test in figure 3.h).

Finally, it is worthily analysing if the degree of standard deviation of the variable, both among the regions included in the same group, *intragroup deviation* (for instance, all the Northern regions among them and all the Southern regions in case 2) and between regions belonging to different groups, *between* group deviation, (for instance, among regions included in the monocentre in case 3.1 and the rest of the European regions), have consequences on the tests. In order to check it, we have replicated the analysis changing the standard deviation between and intragroups. Although for space reason we have not included all of the obtained results, two main conclusions could be derived from them. First, as the deviation intragroup increases (that is, less homogenous are the regions inside each group), lesser is the probability of obtaining a significant spatial dependence scheme. Second, as the groups are more different between them (that is, more heterogeneous are the groups of regions among them), more intensive is the signification of the spatial dependence scheme detected.

3.3. Temporal evolution of the spatial dependence tests

Once analysed the spatial correlograms and the local spatial tests for each spatial distribution, it could be interesting to discuss what behaviour would be expected in case that the integration process follows the U-inverted shape proposed by some models of New Economic Geography.

So, in the firsts moments of the integration, when it was null or very slight, it could be expected a spatial correlogram and a local clusters similar than those depicted in figures 1.c and 1.d, showing neither significant spatial dependence scheme for any order of contiguity (because the activity would be dispersed through the space) nor any region (or few regions) with a significant value of the I_i test.

After, as the integration process was enhanced, it would be expected that the centripetal forces prevailed over the centrifugal, making more advantageous the concentration of the activity in the space. In this case, the spatial correlogram would look like that in figure 1.g (reflecting a clear core-periphery scheme), to 2.c (concentration of activity in some close regions that have a bigger market, better dotations of infrastructures of transports, better human capital, ...) or, for instance, to 3.g (some core-periphery schemes widespread through the space). In any case (with the exception of 3.g), we could expected the appearance of positive and significant spatial dependence scheme at the first orders of contiguity and negative spatial autocorrelation among far regions, showing differences in the spatial distribution of the activity among the European regions. In addition, a local analysis should show the existence of some clusters with a higher (or lower) concentration of the activity than the expected one (or, clusters with a negative sign of the I_i test if the situation was similar than the case 4). However, it should be kept in mind that so that a positive spatial dependence appears it is necessary that the activity is concentrated in some regions close in the territory and that concentration is enough relevant.[14]

Finally, if as a consequence of an enhancement of the integration, a diffusion process was triggered, it could be expected a reshape of the spatial correlograms to those in figures 2.g or 3.c. In the extreme, if the final result of the integration was the return of a rather even distribution of the activity among the European regions, the scheme of spatial autocorrelation should disappear (reappearance of figures 1.c and 1.d). The same result would get it in case that the randomness domains the distribution of the activity at the end of the integration.

However, and before finishing this section, it is necessary to note three points. First, and specially in case 3, we have supposed that the concentration process is supraregional in such a way that some neighbouring regions present advantages and attract the activity towards them. In the same way, we assume that the diffusion process is also supraregional, affecting to the neighbouring regions of first and second order of contiguity of the monocentre. Although the level of aggregation used in our sample (NUTS II) is enough low to assume the previous hypothesis, however, in case that the concentration process appeared only at a local scale (similar than the industrial districts), a regional spatial analysis would not be able to pick, for instance, a diffusion process or relocations of the core.

Second, it is worthily noting that, in order to get a good description of the spatial distribution of one variable, both the global and the local spatial analysis are necessary. This fact will be specially important when, for instance, there has been a relocation of the core (thinking in case that the monocentre moves from French regions to the core defined in case 3). In this case, the spatial correlogram of both situations would be very similar, leading to reject the existence of significant changes in the spatial distribution of the activity (the traditional indexes of inequality would also not change if this relocation had not effect on the dispersion of the variable). Then, the computation of the local spatial association tests would give us more accurate explanation of the real changes happened.

Finally, it is important to note that this kind of analysis does not substitute in none case neither the traditional analysis of industrial concentration nor the indexes of regional disparity. On the contrary, this spatial analysis has to be a complement of those.

4. Empirical evidence

Once analysed the behaviour of the global and local spatial dependence tests against different spatial distributions of a generated variable called "level of activity", in this section we will apply the analysis proposed to the GDP per capita (p.c.) for the same European regions described in the foregoing section for the period 1975-1992.[15] We have obtained the figures from the REGIO database (EUROSTAT), jointly with the base elaborated by the ESOC-Lab at the LSE.

Before the spatial analysis, we have analysed the degree of inequality in this variable, computing the sigma convergence and depicting the density function for two years 1975 and 1992 (in this case, we have computed the variable in deviations respect the European average). After comparing the figures 4.b and 4.f, it seems that the concentration has increased, with a single mode located between the 80% and the 100% of the average of the European levels. Moreover, there has been a slowly disappearance of the mass of probability in the left tail, corresponding to the less favoured levels. This fact is reflecting the approximation of the poorest regions to the medium levels of product and the decrease in the polarisation that there was at the beginning of the period (caused by an important number of regions with very low levels of GDP p.c.).

The analysis of the sigma-convergence shows similar results (and similar than those obtained by several studies of convergence in Europe), that is, a decrease of the dispersion during the seventies and a stagnation since then.[16]

However, with this traditional analysis it is not possible to know if the product is randomly distributed through the space or, on the contrary, if there are homogenous behaviours among neighbouring regions; if there are no spatial stationarities in this variable; if, although the dispersion has decreased, there have been relocations of the core-regions or, for instance, if there are some regions that concentrate excessive incomes in comparison with their neighbouring regions.

In order to answer the prior questions, a spatial analysis has been done. The results are summarised in figures 4.c and 4.d for the year 1975 and in 4.g and 4.h for 1992. From the spatial correlograms, it could be drawn three conclusions. First, there is a significant and positive spatial dependence detected among neighbouring regions during all the period, making evident the existence of a trend to clustering of similar levels of GDP p.c. in the space. Second, the positive spatial autocorrelation is detected up to the fifth order of contiguity meanwhile there is a negative significant spatial dependence for higher orders, similar than the spatial correlogram in 1.g, associated with a Core-Periphery scheme (although at the end of the period it is like the monocentre's spatial correlogram in figure 2.c). From these results, there is a strong evidence about a no randomly spatial distribution of the product among the European regions, existing a relevant heterogeneity on it. So, there are some factors or advantages that exceed of the regional frontiers, leading that the activity is concentrated in some neighbouring regions. The last conclusion is related to the evolution of the spatial distribution. After comparing the figures 4.c and 4.g, it seems that there is a persistence in the spatial configuration of the product, without evidence of a diffusion process towards the peripheral regions.

At a local spatial analysis it could be drawn that in 1975 there are clusters of low values of GDP p.c. in the almost totality of the Greek regions and those located on the South of Italy (remembering that it means that each region detected, jointly with its neighbours, concentrate significant lower levels of GDP p.c. than the expected ones). Moreover, the Portuguese regions jointly with some Southern Spanish regions (plus both Castillas) concentrate low values of this variable. A similar scenario is detected at the end of the period widening lightly the clusters of low level towards the Northeast of Spain and the South of Italy. However, it seems that there are significant differences in the location of the clusters of high levels of product during the analysed period. In 1975 there was a core located specifically in the North of France, in some neighbouring Italian regions, regions in the Northeast and in the West of Germany and in some regions of Belgium and Holland. On the contrary, in 1992 it is detected a relocation of the most dynamic regions of Europe. So, most of the clusters detected at the beginning of the period in Holland and France disappear at the end, meanwhile the number of clusters with high levels of product increases in the South of Germany and in their neighbouring regions in the North of Italy (leading to increase, in this last country, the duality among Northern and Southern regions). Finally, it is necessary to comment the clusters of dissimilar levels of product. In this sense, in 1975 the Belgian region of Walloon, that concentrated low levels of product, it was surrounded by Belgian, French and German regions that showed high levels of GDP p.c. At the end of the period, the number of clusters of dissimilar values is widening to Cerdeña (with high levels of product) and to Lisbon and Madrid (capital of countries that concentrate most of the activity in comparison to their neighbours).

5. Final remarks and conclusions

This communication has pursued as a main goal to propose the use of some spatial dependence tests in order to carry out a spatial analysis that complement the habitual tools used in studies related to the location of the activity or the level of regional disparities. In this sense, we have presented global and local spatial association statistics used in the paper that consider, at the same time, the behaviour detected in one region and in its neighbours (defining as neighbouring regions those that have potential interactions among them).

Specifically, and in order to analyse the sensitivity of the spatial tests to several spatial distributions of one variable, it has been done a simulation exercise. In this case, and after generating several series that incorporate different hypothesis about the spatial distribution of one generic variable called "level of activity", we have computed the mentioned tests. Furthermore, and under the assumption of a relation with a U-inverted shape between integration and concentration, it has been derived a theoretical evolution of the analysed tests during an integration process.

Finally, we have applied the spatial analysis to the GDP p.c. in the European regions during the period 1975-1992. From the obtained results, it could not get evidence about a significant diffusion process of the product through the European regions. On the contrary, despite the relocation of the core-regions observed, it has been detected a outstanding persistence in the schemes of location of the activity, showing the predominance of the centripetal versus centrifugal forces even at the end of the period.

Final notes

[1]See Ottaviano and Puga (1997) for a survey of this kind of models.

[2]Internal returns to scale and external economies derived from the existence, for instance, of backward and forward linkages (among others).

[3] For example, diseconomies of scale or the high competition in products and factors that exists in places that concentrate most of the activity.

[4] In this case, the Moran's I standardised follows a normal distribution N(0,1).

[5] One region h would be second order contiguous to another region i if it is first order contiguous to another unit that itself is first order contiguous to the region h.

[6] Again, the standardised test is distributed by a normal N(0,1).

[7] We are assuming that the Northern regions have a bigger market and better infrastructures and human capital dotations.

[8] The monocentre in case 3 has 15 and 17 first and second order of contiguity neighbours respectively. The same happens when the monocentre is in France. However, when the monocentre is in Portugal and Spain, it has only 4 first and 4 second order of contiguity neighbouring regions.

[9] Although the generic variable that we have called "level of activity" could be associated both with, for instance, the employment in the industrial sector (following a traditional analysis of location) or the product per inhabitant (following a study of the level of regional disparities), in this case we have linked the created variable more to the last concept.

[10] The density function is useful to get some quickly idea about the dispersion of one variable. So, higher concentration of the probability, lower inequality (convergence).

[11] Moreover, it is necessary to take into account that the shape of the density function depends on the characteristics of the variable as, for instance, the number of the observations of the sample.
[12] Although we have not incorporated the results, another spatial distribution has been analysed: an uneven distribution of activity. So, from the case 1, we have increased the standard deviation up to obtain a variable with high degree of inequality but without any spatial scheme incorporated. In this case, the results obtained for the global and local spatial tests are identical than those derived for the case of an homogeneous distribution: no spatial dependence for any order of contiguity and none spatial cluster. This is an example of the necessity to compute both the traditional analysis of inequality and the spatial analysis proposed in order to obtained a completed vision of the distribution of the variable: only the computation of the inequality indexes or the view of the density function can give us information about the high degree of dispersion in this second situation.
[13] For instance, the monocentre in case 3 has 15, 17, 13 and 8 neighbours of first, second, third and fourth order of contiguity respectively. When the monocentre is in France, the number of neighbours are 15, 17, 22 and 12 respectively. This fact is result of the location of the monocentre: in the second case, it is in the middle of Europe and absolutely surrounded by regions included in the sample. In case 3, there are some regions that have not neighbours in a part of their perimeter.
[14] On the contrary, if only appeared some regions (sparked through the space) with a slight concentration of activity, it would be possible that the null hypothesis of no spatial autocorrelation would be not rejected.
[15] In previous works, we have done a simple spatial analysis of the GDP per worker, the growth rates of the GDP p.c. and GDP per worker (López-Bazo et al, 1997) and the industrial employment (Vayá, 1997) for the European regions for the period 1975-1992. Specifically, the results and questions derived from these works led us to study in depth the information supply for the spatial dependence tests by means of the simulation exercise done in this communication.
[16] This conclusion has been obtained after comparing the density function and the sigma convergence for 1975, 1985 and 1992. Specifically, the values of the sigma-convergence are: 0,42 for 1975, 0,34 for 1985 and 0,34 for 1992.

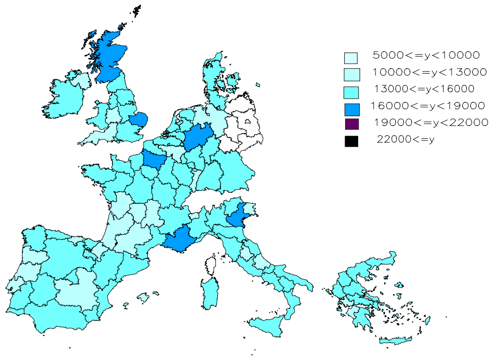
References

Amiti, M., 1997. Specialization patterns in Europe.Discussion Paper Nº363,Centre for Economic Performance,LSE.
Anselin, L., 1988. *Spatial econometrics: methods and models*. Kluwer Academic Publishers, The Netherlands.
Anselin, L., 1995. Local indicators of spatial association-LISA. Geographical Analysis 27, 93-115.
Anselin, L. and R. Florax, 1995. *New directions in Spatial Econometrics.* Springer-Verlag, Berlin.
Armstrong, H.W., 1995. Cross-sectional analysis of the regional growth process, in *Convergence and Divergence among European Regions*, H.W. Armstrong and R.W. Vickerman (eds), Pion Ltd, London : 40-65.
Brülhart, M. and J. Torstesson, 1996. Regional integraction, scale economies and industry location. Discussion Paper nº. 1435, Centre for Economic Policy Research.
Cliff, A. and Ord, J., 1973. *Spatial Autocorrelation*. London, Pion.
Cliff, A. and Ord, J., 1981. *Spatial process. Models and applications*, Pion Ltd., London.
Esteban, J., 1994, La desigualdad interregional en Europa y en España: descripción y análisis, in *Crecimiento y Convergencia Regional en España y Europa*, vol.2, IAE, Barelona, pp. 13-84.
Fluvià, M. and J. Gual, 1993. Comercio Internacional y Desarrollo Regional en el Marco de la Integración Económica Europea. In *Crecimiento y Convergencia Regional en España y Europa*. Vol.2, pp. 85-123. Instituto de Análisis Economico, CSIC.
Getis, A. and J.Ord, 1992. The analysis of spatial association by use of distance statistics. Geographical Analysis 24, 189-206.
Greenaway, D.and R.C.Hine, 1991, Intra-industry specialization and adjustment in the European economic space. Journal of Common Market Studies, 29(6), 603-22.
Krugman, P., 1991a. Increasing returns and economic geography. Journal of Political Economy 99, 484-499.
Krugman, P., 1991b. *Geography and Trade*. MIT Press, Cambridge MA.
Krugman, P. and A. Venables, 1995. Globalization and the inequality of nations. Quarterly Journal of Economics 110, 857-880.
Krugman, P. and A. Venables, 1996. Integration, specialization and adjustment. European Economic Review 40, 959-967.
López-Bazo, E.; Vayá, E.; Mora, A.J.; Suriñach, J., 1997. Convergencia regional en la Unión Europea ante el nuevo entorno económico, *ICE Revista de Economía*, 762, 25-41.
Molle, W., 1996. The regional economic structure of the European Union: an analysis of long term developments.

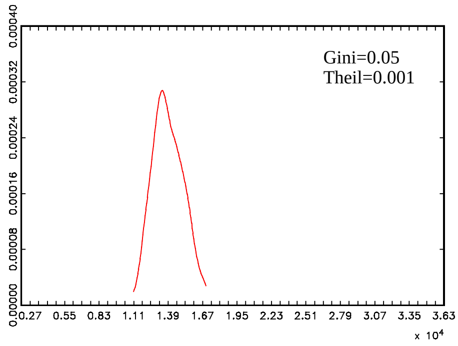
Regional growth and regional policy within the framework of european integration. Karin Peschel (eds), Physica-Verlag.
Molle, W.and Boeckhout, S, 1995: Economic disparity under conditions of integration-A long term view of the European case, *Papers in Regional Science*, 74:105-123.
Moran, P., 1948. The interpretation of statistical maps. Journal of the Royal Statistical Society B 10, 243-251.
Neven, D.J. and Gouyette, C., 1994: Regional convergence in the European Community, *CEPR Discussion Paper*. Series 914, London.
Ottaviano, G. and D. Puga, 1997. Agglomeration in the global economy: A survey of the 'new economic geography'. Discussion Paper Nº 356, Centre for Economic Performance, LSE.
Puga, D., 1996. The rise and fall of regional inequalities. Discussion Paper Nº 314, Centre for Economic Performance, LSE.
Puga, D. and A.J. Venables, 1996. The spread of industry: spatial agglomeration in economic development. Journal of the Japanese and International Economics 10, 440-464.
Suarez-Villa, L and J.R. Cuadrado, 1993. Regional Economic Integration and the evolution of Disparities. Papers in Regional Science 72 (4), 369-387.
Vayá, E., 1997. Spatial association in the employment location among the European Regions, Paper presented in the 37th European Congress, European Regional Science Association.
Venables, A., 1996. Equilibrium locations of vertically linked industries. International Economic Review 37, 341-359.

Table 1. Spatial analysis for cases 1 and 2

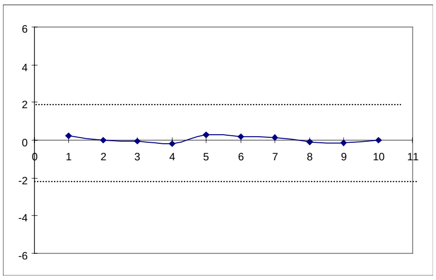
1.a) Spatial distribution, homogenous distribution (case 1)



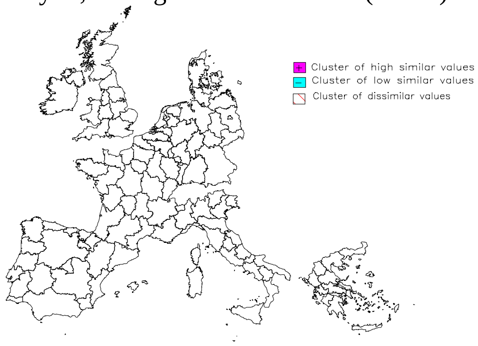
1.b) Density function, homogenous distribution (case 1)



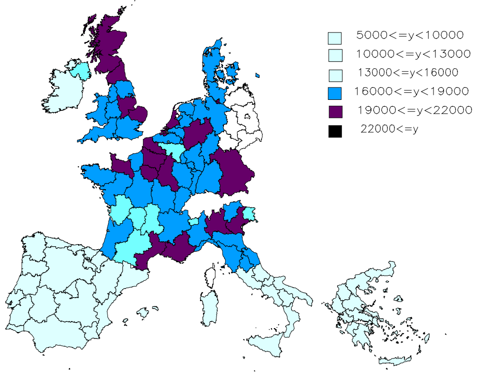
1.c) Spatial correlogram, homogenous distribution (case 1)



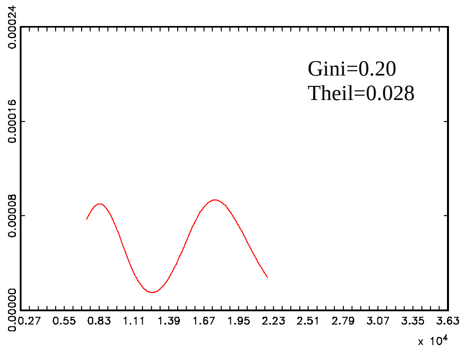
1.d) Local analysis, homogenous distribution (case 1)



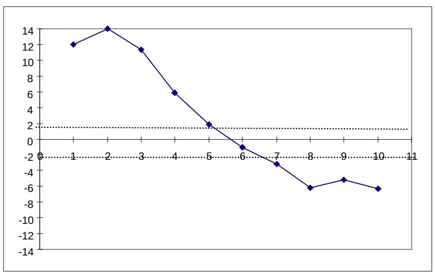
1.e)Spatial distribution, Core-Periphery distribution (case 2)



1.f) Density function, Core-Periphery distribution (case 2)



1.g)Spatial correlogram,Core-Periphery distribution(case 2)



1.h) Local analysis, Core-Periphery distribution (case 2)

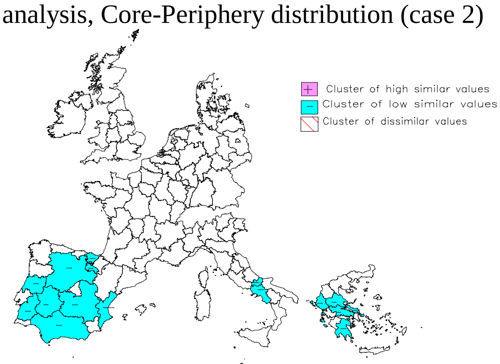
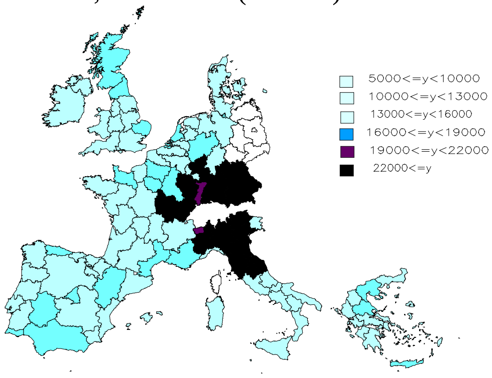
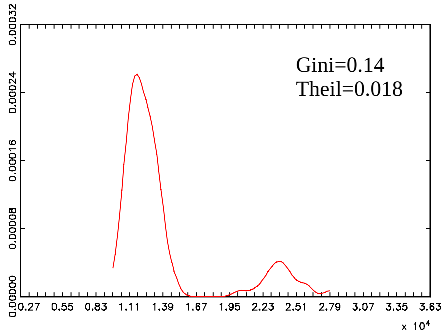


Table 2. Spatial analysis for cases 3.1 and 3.2

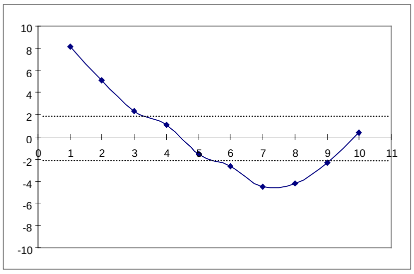
2.a)Spatial distribution, mococentre (case 3.1)



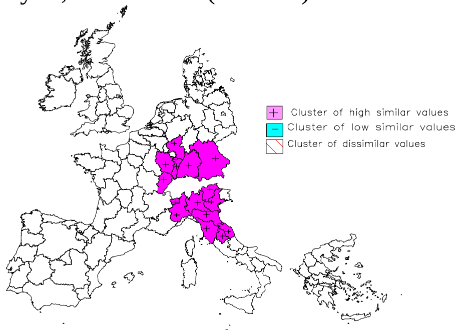
2.b) Density function, monocentre (case 3.1)



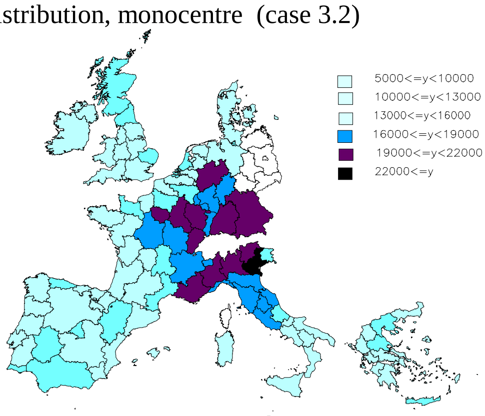
2.c) Spatial correlogram, monocentre (case 3.1)



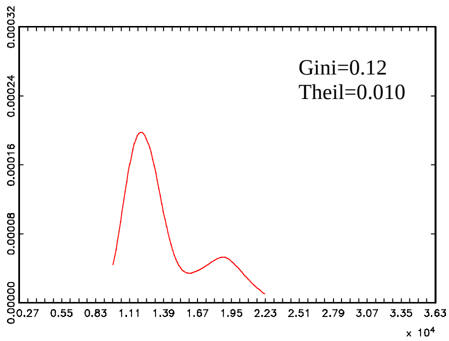
2.d) Local analysis, monocentre (case 3.1)



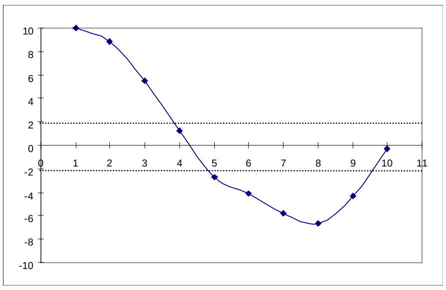
2.e). Spatial distribution, monocentre (case 3.2)



2.f) Density function, monocentre (case 3.2)



2.g) Spatial correlogram, monocentre (case 3.2)



2.h Local analysis, monocentre (case 3.2)

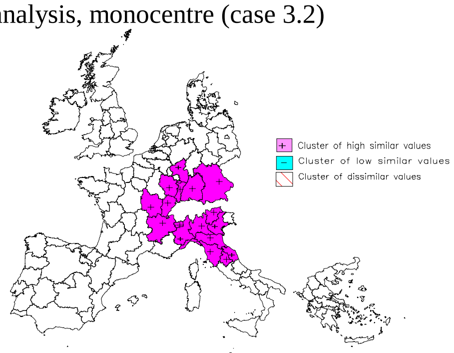
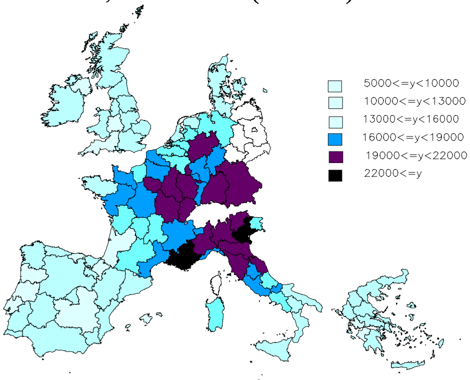
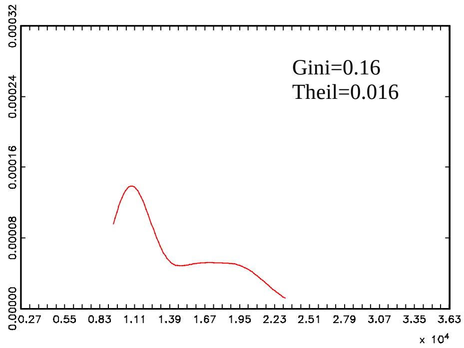


Table 3. Spatial analysis for cases 3.3 and 4

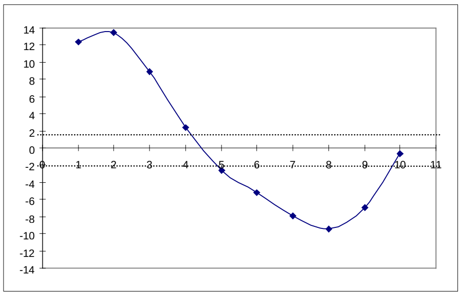
3.a). Spatial distribution, monocentre (case 3.3)



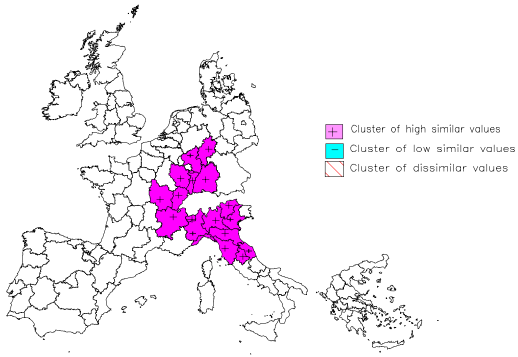
3.b) Density function, monocentre (case 3.3)



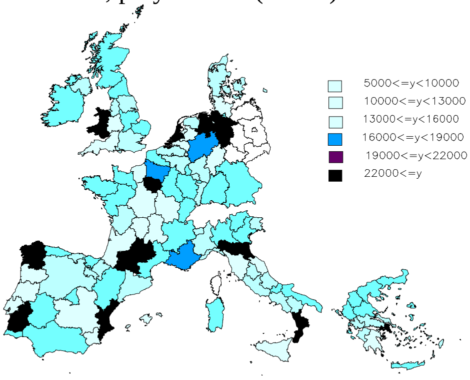
3.c) Spatial correlogram, monocentre (case 3.3)



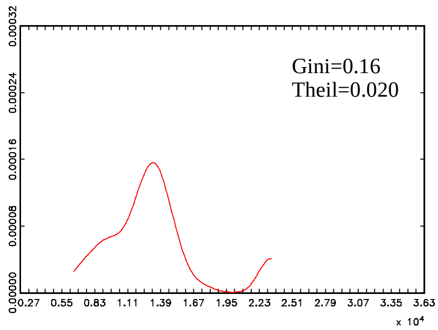
3.d) Local analysis, monocentre (case 3.3)



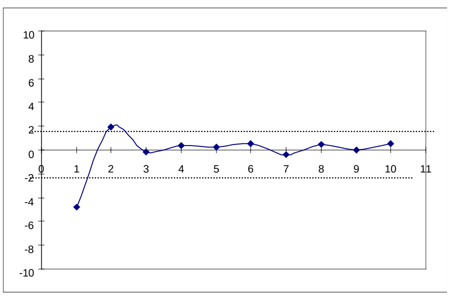
3.e). Spatial distribution, polycentres (case 4)



3.f) Density function, polycentres (case 4)



3.g) Spatial correlogram, polycentres (case 4)



3.h) Local analysis, polycentres (case 4)

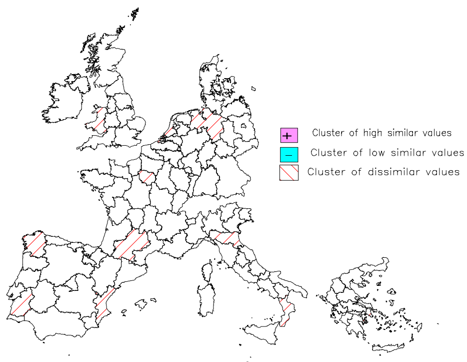
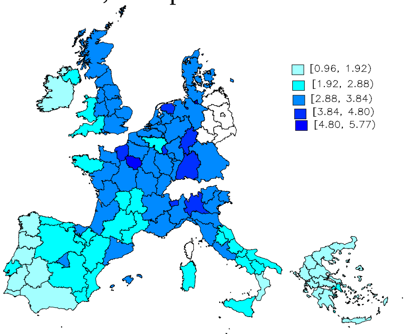
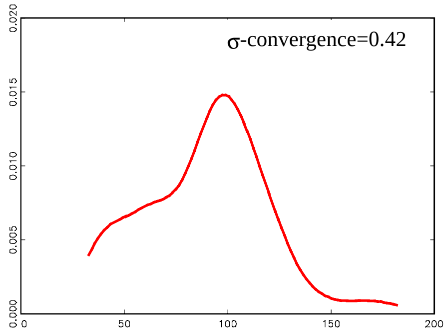


Table 4. Spatial analysis of he GDP p.c.

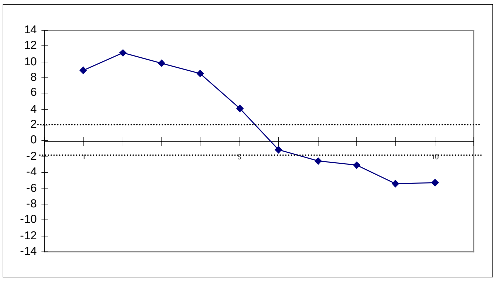
4.a). Spatial distribution,GDP p.c. 1975



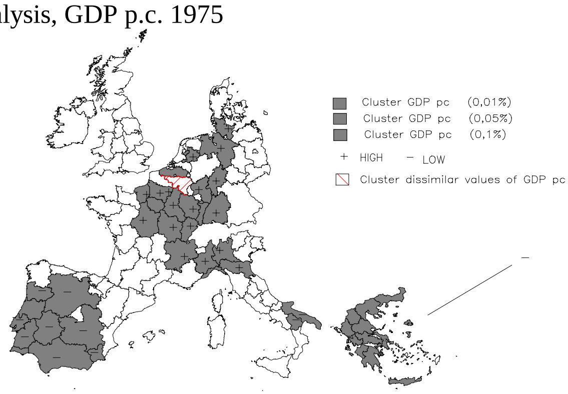
4.b) Density function,GDP p.c. 1975



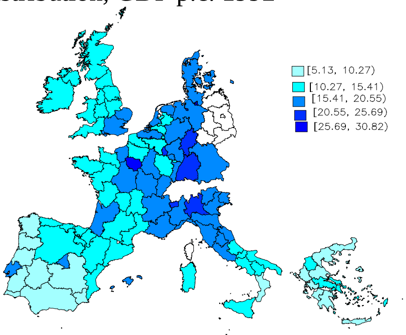
4.c) Spatial correlogram, GDP p.c. 1975



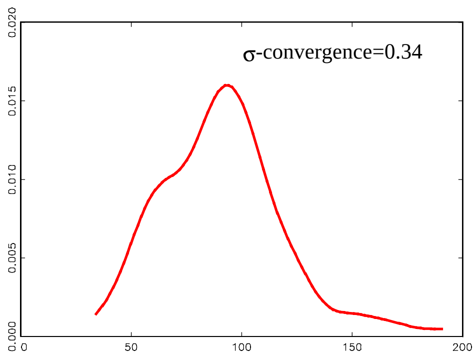
4.d) Local analysis, GDP p.c. 1975



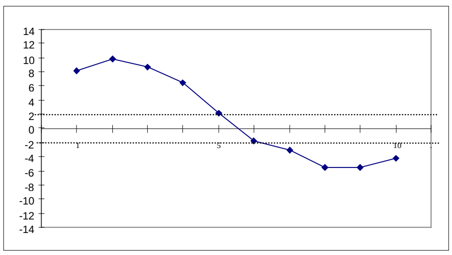
4.e). Spatial distribution, GDP p.c. 1992



4.f) Density function, GDP p.c. 1992



4.g) Spatial correlogram, GDP p.c. 1992



4.h) Local analysis, GDP p.c. 1992