TECHNICAL REPORT

ESPON

JULY 2014

Local data in M4D:

- LAU2 and Very Important Geographical Objects (VIGO)
- Delineating an alternative geometry at local scale

CONTENT

This technical report describes the methodology used to delineate alternative geometries at local scale and explores the conceptual potential of the VIGO.
The delineation of an alternative geometry at local scale is the output of 4 methods tested and described in this paper. In the case o f VIGO, exploring the relation between the geographical objects and the LAU2 is based on a typology of possible scientific approaches that combines different tools of spatial and geostatistical analysis.

23 pages

ESPON M4D -
MULTI DIMENSIONAL DATABASE DESIGN & DEVELOPMENT

EUROPEAN UNION
Part-financed by the European Regional Development Fund
INVESTING IN YOUR FUTURE

LIST OF AUTHORS

Alexandru Rusu, UAIC, CUGUAT-TIGRIS, Iasi, Romania

Alexandru Banica, UAIC, CUGUAT-TIGRIS, Iasi, Romania

Octavian Groza, UAIC, CUGUAT-TIGRIS, Iasi, Romania

Daniel Tudora, UAIC, CUGUAT-TIGRIS, Iasi, Romania

Contact

alexrusucuguat@yahoo.com

UAIC, CUGUAT-TIGRIS, Iasi, Romania

Tel. (+ 40) 0742 42 20 58

TABLE OF CONTENT

Introduction

The collection of data regarding a very sensible geographical frame (the LAU2 geometry) is often blocked by two problems - the coherence of the geographical dataset we use (the basemap) and the access to data. If the creation of an alternative geometry was imagined as a possible solution to overcome the first issue, the VIGO helped as to present the theoretical and methodological ways by which one can create indicators from almost pure spatial data. The scientific stakes of this technical report are related to the exploratory phase needed to better understand how the territorial knowledge might be created for the basis of the administrative system.

The delineation of the alternative geometry intended to create what we address, in this paper, as *pseudo-LAU1*, a kind of missing link between the local level (LAU2) and the NUTS3 level of political and administrative decision. In many cases, the scientific approach faced a common geographical topic - the problem of regionalization or district creation. Starting with an expert opinion solution and finishing with the elaboration of a GIS tool that enables the solving of the districting problems, in specific cases, we reviewed the most used methods to bypass the aggregation of LAU2 in pseudo-LAU1.

The relation between the VIGO and the LAU2 is the second topic of this technical report. Taking into account the fact that these geographical objects are legion in the territory, a synthesis table that describes how is possible to deal with them is the starting point of the second part. The mentioned table combines methodological tools that mobilizes concepts such as the potential of spatial interaction, the cumulated population by distance towards the nearest VIGO or qualitative typologies derived from time-distances. In some cases, the analysis and the working flows were transformed in algorithms and GIS tools.

The expertise provided in the exploration of these two topics has a triple output:

- the methodological description of the scientific approach;
- the indicators created for each problem solved;
- synthetic GIS tools that may be used in order to reproduce t he general approach, in different geographical contexts.

1 Delineating an alternative local geometry for selected countries in the Eastern Europe

1. Theoretical background

1.1 Introduction to the problem

The creation of an alternative geometry for selected countries from the Eastern Europe represents a classical problem of regionalization. In spatial analysis this problem intersects two important topics: *how to merge specific spatial units in large ones*, respecting a topological rule (must have a common limit) and *how to ensure that the regional frame created is homogeneous*. The algorithm that solves this double problem must deal with the following questions:

a) how to define a homogeneous or a heterogeneous region ?
b) how to measure the degree of regional homogeneity ?
c) how can one find an optimal number of regions to create?
d) how to ensure a common size (surface) to the regions that will be created?

a) the first question might look simple, but the answer is quite a complex one. Generally, the homogeneity of a region is derived from the values of the indicators that describe the spatial units composing that region (e.g. the LAU2 incorporated in a NUTS3). With only one indicator (density of population, for example), we assume that a region is a homogeneous when the values are similar. The problem is more complicated when we evaluate the homogeneity of a region, starting with n indicators. As a matter of fact, dealing with only one indicator is a particular case of the problem. The core of the issue is the multivariate analysis of the regional degree of internal similarity (homogeneity) and this aspect is linked with the second question.

b) measuring the internal similarity demands the creation of similarity index. This index can be expressed as the absolute difference between the values of the indicator (s) that one uses in the evaluation. These absolute differences can be squared and they will maximize the dissimilarities. In the same logic, these differences can be weighted, in order to obtain a common frame of reference for the values. Assessing the overall homogeneity is a matter of calculation of one classical geo-statistical analysis tool - the territorial auto-correlation indicator.

c) the basic criterion for the creation of an x number of regions is derived from the previous point. Using the territorial auto-correlation index, one can maximize or minimize the number of needed regions for the studied are. This index is similar to the classical Pearson coefficient of correlation and it takes values ranged from -1 (total heterogeneity) to 1 (total homogeneity). Comparing the values of the index for a different number of regions is the way to decide how many units one will create.

d) probably the most difficult question to ask. The answer depends on the study area and on the methodology of region creation. Some methodologies (hierarchical clustering with spatial constraint) will provide regions with different sizes, even if statistically sound and validated. This inequality of size is due to the nature of the indicators used in the classification and regionalization.

As we already mentioned, the creation of an alternative geometry represents a classical problem of regionalization. In the early years of quantitative geography the topic is widely debated (Haggett, P., 1965), but the technological limitations allowed little progress in the implementation of advanced algorithms. This topic of regionalization presented interest both for planning professionals and for policy-makers and the preoccupations on finding a good tool of districting continued. Nowadays we assist to many attempts in defining a methodology for districting or regionalization, the scientific production being polarized by some key terms - optimization, design, model, approach, planning etc.

Fig.1 Word cloud based on a sample of 200 scientific articles dealing with the districting problem. Source of data: www.science-direct. com. Made with www.worldle.com.

1.2 Work done

The solutions we proposed are derived from some recent studies that emphasize the potential utility of the regionalization for decision-makers and policy designers (districting and re-districting problems, gerrymandering, planning regions). In the scientific literature, the most cited methodology to solve this regionalization problem is the hierarchical clustering approach with spatial constraints. This method weights the similarity matrix with a spatial matrix (distance, contiguity, k-number of neighbors etc.) and it will implement the classical algorithm by merging the individuals in respect to a statistical criterion (minimum linkage, complete linkage, Ward's criterion etc.).

However, the hierarchical clustering approach is just a way to define regions or districts in a given space. In the elaboration of an alternative geometry based on the LAU2 frame, we have tested several methods to delineate the pseudo-LAU1 regions. ***The first method*** explored was based on a selection of aggregation centers (expert opinion) that will eventually be used to coagulate the pseudo-LAU1. The approach is limited and might receive much criticism. First of all, it is a method that can be applied only on a reduce area and it will engage all the expertise of a geographer. Secondly, the expert opinion method is unfriendly because it multiplies the quantity of information needed in order to delineate the pseudo-LAU1 frame, due to the fact that the method can only be applied NUTS3 by NUTS3. In the case of an ESPON state with numerous NUTS3 (France), it implies an iteration of the method more than 90 times, with no warranty that the final result is scientifically sound.

Expert oppinion -1	P-median solution - 2	Hierarchical clustering with spatial constraints - 3	Pseudo K-mean clustering with spatial constraints - 4
Difficult to evaluate.	Tested on a large number of countries;it provided a first version of the alternative geometry.	Implemented in RedCap (www.spatialdatamining.org). Not reliable for large datasets such as the French LAU2.	Solution proposed by our approach.

Tab. 1 Methods tested for the alternative geometry delineation

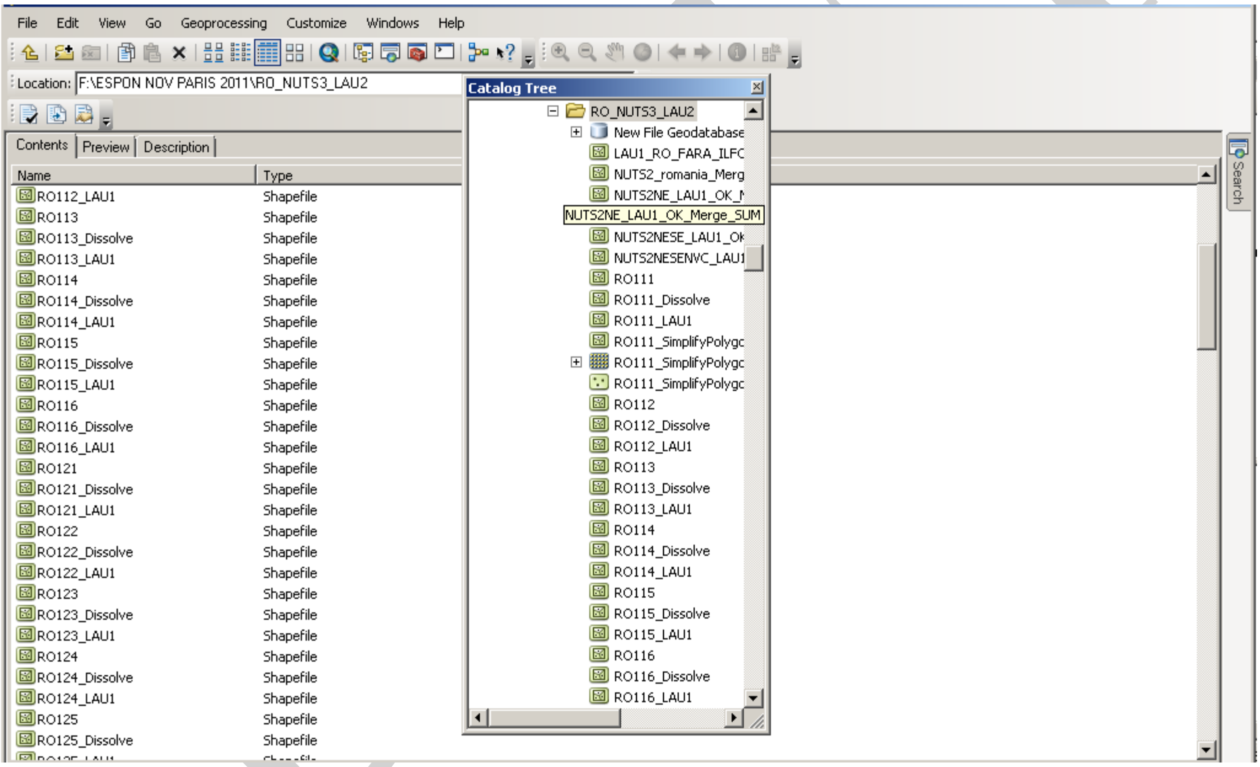

Fig. 2 Output of the method no.1 - case study : Romania

The image describes the quantity of information stocked after applying the expert opinion method on a country with 42 NUTS3 (Romania). It's just a part of the information we stocked because the implementation of the method mobilizes also point spatial structures (the centroids of the LAU2 used for the aggregation). As this first attempt to elaborate an alternative geometry lacks scientific consistency, it was abandoned.

The ***second method*** we explored is based on the location-allocation model. This model responds to one simple question: how to locate n facilities in a region, avoiding the competition between them? When we translate this question for the problem we seek to solve, we obtain this version: how to select a proper number of centers that will aggregate the LAU2 in pseudo LAU1, so that the centers are placed in a regular lattice that ensures equal areas for the regions we create?

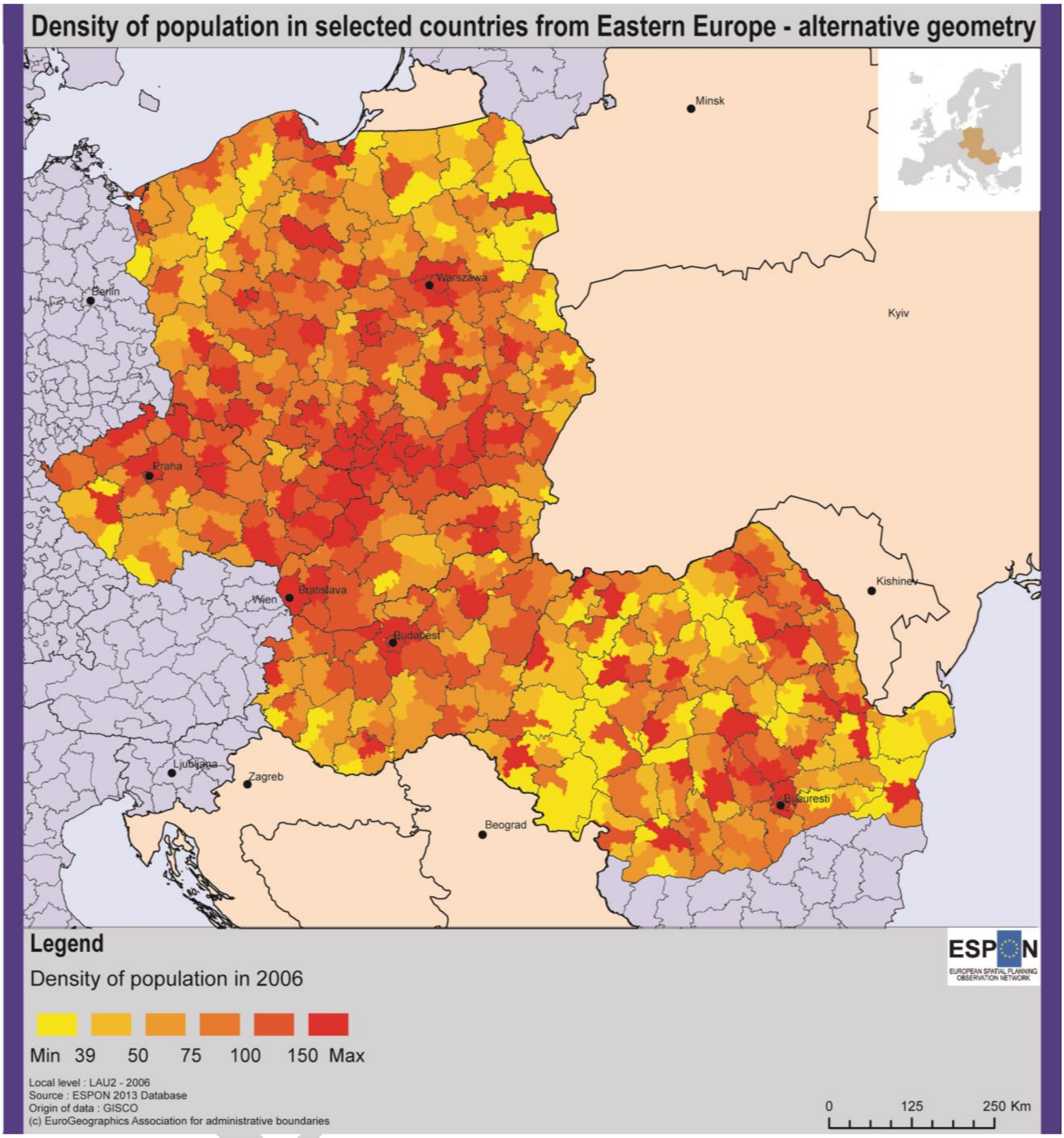

Fig. 3 Alternative geometry derived by the second method

Using a p-median algorithm that interrogated an origin-destination distance matrix between the LAU2 from selected countries calculated in the road network, we have managed to obtain an alternative geometry that covers a large number of ESPON states. In the illustration we present the results for only five Eastern European countries and a possible use of this output, when mapping a simple indicator such as the density of population. In the implementation of the algorithm we have forced the aggregation to take place only within the NUTS3 limits, multiplying the time needed for calculus, but obtaining better results. There are several steps to be followed in order to create this alternative geometry, using the location-allocation method:

1) a surface criterion must be retained in order to have a comparable geometry in the studied space. We have made an option for pseudo-LAU1 similar in size to the German NUTS3.
2) The LAU2 polygons were converted to centroids and for each centroid the location in the road network was calculated.
3) the boundaries of the NUTS3 were extracted and used as spatial limits for the selection of the candidate centers of LAU2 aggregation.
4) the method (location-allocation with p-median algorithm) was implemented and iterated for the selected countries.
5) once the candidate centers were obtained, the LAU2 were aggregated in pseudo LAU1. The aggregation criterion is the shortest distance between an LAU2 and a candidate center, using the road network.
6) the internal limits of the LAU2 were dissolved and a pseudo-LAU1 frame was obtained.

The main limitation of this approach is related to the use of a network in the attempt to identify the candidate centers. In this case, for some countries of the ESPON space it would be impossible to delineate a pseudo-LAU1 geometry for a simple reason: we lack a proper network. Moreover, the solution is based on the identification of the most "central" candidate points in a given region, making this methodology dependent on the morphology of the network. We consider that this aspect should be avoid.

A ***third solution*** explored is the hierarchical clustering with spatial constraints. Very promising at the beginning, it proved to be a deception for our purpose, once tested. Despite the elegant method that weights a similarity matrix with a spatial relations one, the results are unstable and unfitted to our goal.

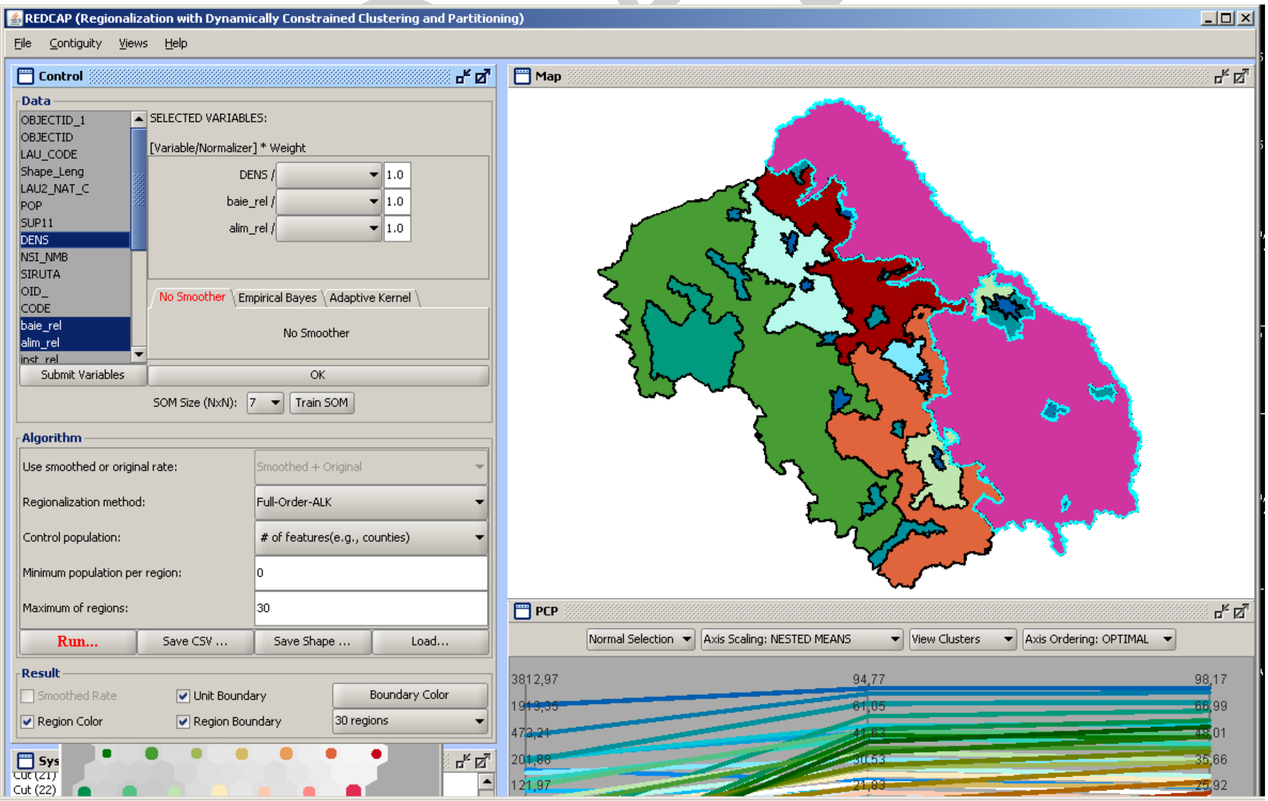

Fig. 4 Hierarchical clustering with spatial constraint - unstable results

Applying this third solution will generally provide regions with unequal area, like in the illustration above. From a scientific point of view, the method itself is very robust, but it is just not appropriate for our intention. It functions just like a common hierarchical clustering technique, excepting the fact that the similarity matrix of the individuals is weighted by a binary contiguity matrix. In this case, only the values of the neighbor individuals are maintained. A first search for the most similar pair of neighbors provides an initial merge of spatial units. The search for a minimum similarity is iterated and a new merge is provided. When there are no more pairs of individuals left in the similarity matrix, the search stops and a cluster dendrogram is provided. In a spatial context of autocorrelation, the merging iteration will naturally produce regions with unequal area. The boundaries of the regions that this method creates overlay the major territorial discontinuities in the geographical repartition of the indicators. Tested on several study areas and with different indicators, the hierarchical clustering algorithm failed to fulfill our expectations - the creation of an alternative geometry of pseudo-LAU1, covering selected countries from the Eastern Europe.

The **last method (4th)** we tested is a combination between the network analysis techniques and the exploration of a similarity/dissimilarity matrix between the LAU2, in a given space. As a matter of fact, it represents a methodological cocktail of tools that were applied during for the second and third algorithms of districting. The starting point for this approach is derived from an analogy between the statistics and the spatial analysis.

Statistics	**Spatial Analysis**
Mean	Mean Center
Weighted mean	Weighted mean center
Standard deviation	Standard distance
Regression	GWR
Hierarchical clustering	Hierarchical clustering with spatial constraint
K-mean classification	? P-median approach on a similarity graph
Factorial analysis	?

Tab. 2 Analogy between two categories of tools : statistical tools and geo-statistical instruments

A large number of tools common in the statistical analysis of data find an equivalent in the spatial analysis. If there is equivalence between the standard distance and the standard deviation, when calculating dispersion parameters, it might be also equivalence between different classification techniques and the districting algorithms. The most popular method used for data classification is the hierarchical clustering. However, other methods of classification exist and they might be more appropriate in some specific contexts of geographical analysis. One of these methods is the K-mean clustering tool that enables the creation of subsets of information of almost equal size. The tools aggregates the individuals into subsets, using the statistical distance between individuals and their allocation to the closest statistical centroids ("local means"). Its adaptation for geostatistical analysis is rather simple and it is based on the analogy between the steps of implementation:

- individuals to be classified = LAU2 to be regrouped in pseudo-LAU1
- statistical distance = similarity/dissimilarity graph between the LAU2, using contiguity as rule of spatial data control

- statistical centroids = candidate centers for the LAU2 aggregation in pseudo-LAU1

- K-mean data partition = p-median solution of location-allocation

After having established the general frame of the methodology no. 4, we have organized the steps of implementation, as follows:

1. Define the variables needed to create the similarity/dissimilarity matrix. If the other three methods of districting used only one variable, with this tool the multivariate analysis is allowed. The similarity matrix between the LAU2 is build by creating the statistical distance between the individuals. Formalized, the statistical distance can be expressed as :

- Dij = |V1i-V1j|+|V2i-V2j|+..., where :
- i, j => any pair of contiguous LAU2
- V1,V2, Vn => indicators that describes the LAU2.
- High values in the similarity matrix indicates a strong dissimilarity between a pair of LAU2, small values suggest similarity between them. These values will be later used as impedance in a spatial network, in order to identify candidate centers.
- The proper formalization of the similarity index is crucial for the implementation of the method. There are two aspects to take into account:
- 1.1 The mathematical expression we exemplified is just one particular case. Depending on the statistical internal organization of the data, other expressions can be implemented (squared differences, weighted values etc.)
- 1.2 If the purpose of the districting operation is to obtain regions with the same characteristics (homogeneous), the values of the similarity index can be conserved in their normal form. If the intention is to aggregate LAU2 in a pseudo-LAU1 that is heterogeneous, the impedance in the spatial network should be calculated as 1/Dij.
-
  1. If an option is made for contiguity between the LAU2, build a binary spatially weight matrix. Else, choose an alternative spatial weight policy (IDW, for example). The contiguity matrix will contain only values that describe the relation between the LAU2, 1 if neighbors, 0 otherwise. The matrix obtained at the first point is intersected with the spatial weight matrix and all the non neighbors links are eliminated.
-
  1. If an option is made for territorial belonging, eliminate the unnecessary similarity/dissimilarity values. The spatial weight matrix can be filtered with a territorial belonging field: 1 if in the same NUTS3, 0 otherwise. This step will eliminate all the links between pairs of LAU2 that share a NUTS3 common border.
-
  1. Transform the spatially weighted similarity/dissimilarity matrix in a graph. The transformation of the similarity matrix in a spatial structure is possible using the coordinates of the LAU2 centroids.
-
  1. Define the number of needed regions (for example, aggregate the LAU2 in 84 pseudo LAU1). If one will need to create pseudo-LAU1 of an x size (area), it is trivial to calculate the needed number of regions by dividing the surface of the studied area with the needed x size.
-
  1. Implement the location-allocation model on this graph, choosing the P-median solution. Once the spatial graph of similarities is build, it is transformed in a network. The distance between any pair of LAU2 is given by the value of the similarity between i and j. Implementing the

calculus algorithm will provide the candidate centers of LAU2 aggregation. Each LAU2 will be allocated to a candidate center, the internal boundaries will be dissolved and a spatial frame of pseudo-LAU1 will be provided.

1. Evaluate the result by a territorial autocorrelation coefficient. This step is necessary in order to ensure some flexibility for the method. Let's assume that we are not very sure how many regions we want. In that case, testing each output for territorial autocorrelation will help a decision to be taken. The test itself is extremely simple, but its implementation in a GIS is problematic. Formalized, the test appears as follows:

- TAC index = 1- (Intra-class variance/Inter-class variance)
- The intra-class variance is defined as the average of the similarities within a pseudo-LAU1. The inter-class variance is provided by the average of the similarities between the LAU2 not belonging to the same pseudo-LAU1. Theoretically, the range of values for this index is situated between -1 and 1. Values closed to 1 indicate pseuo-LAU1 regions characterized by homogeneity in the repartition of the indicators, values closed to -1 will signal pseudo-LAU1 regions that are heterogeneous. The major difficulty is to iterate the test each time an evaluation of the districting spatial frame is needed. The operation multiplies the steps in an eventual GIS tool, but it is strongly advised to verify this index.

2. Test, exemples and implementation

An example with draft maps and illustrations will clarify some of the aspects that were less described in the theoretical background. The tested area is one NUTS2 of Romania (The North-East Region) composed by six NUTS3. The NUTS3 are also divided in 552 LAU2, a number large enough to provide a reliable test. The differences in size for these LAU2 are easy to identify, in the West the LAU2 are situated in a mountain region.

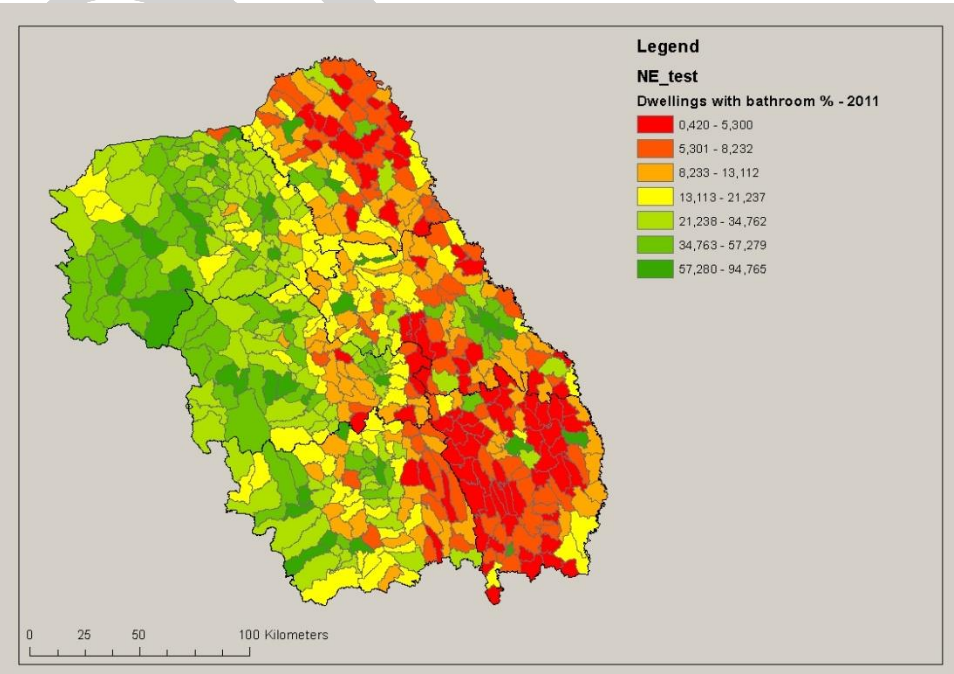

Fig. 5 Cartographic illustration of one indicator used for the similarity graph

For the test of the districting tool, three indicators of territorial endowment were chosen:
a) dwellings with bathrooms inside the house - % of the housing stock
b) access to water pipelines - % of the housing stock
c) the density of population

All the three indicators were extracted from the Romanian Census of 2011 and, despite the academic criticism and the policy-makers reserves, they depict a situation that matches quite well the field reality. The longitudinal opposition in the spatial repartition of the values represents a territorial structure that it is barely complicated by some oasis of urban comfort. We will use these three indicators in order to elaborate a pseudo-LAU1 frame that aggregates the LAU2 in 35 regions. The similarity index was build using the formalization proposed previously and it was calculated in a double form, one serving for the construction of homogeneous pseudo-LAU1, the other for the delineation of heterogeneous pseudo-LAU1.

After this index of similarity was build, the similarity matrix was filtered by a spatial and a territorial constraint. In the first step, we have maintained as useful only the data describing the similarities between each pair of neighbor LAU2. The table was purified again and only the data describing the similarities between each pair of neighbor LAU2 belonging to the same NUTS3 were conserved. Using the latitude and the longitude of the centroids that identify the LAU2, a similarity graph was derived. The graph was transformed in a network dataset that enables classical techniques of network analysis: shortest path, closest facility identification or location-allocation algorithm.

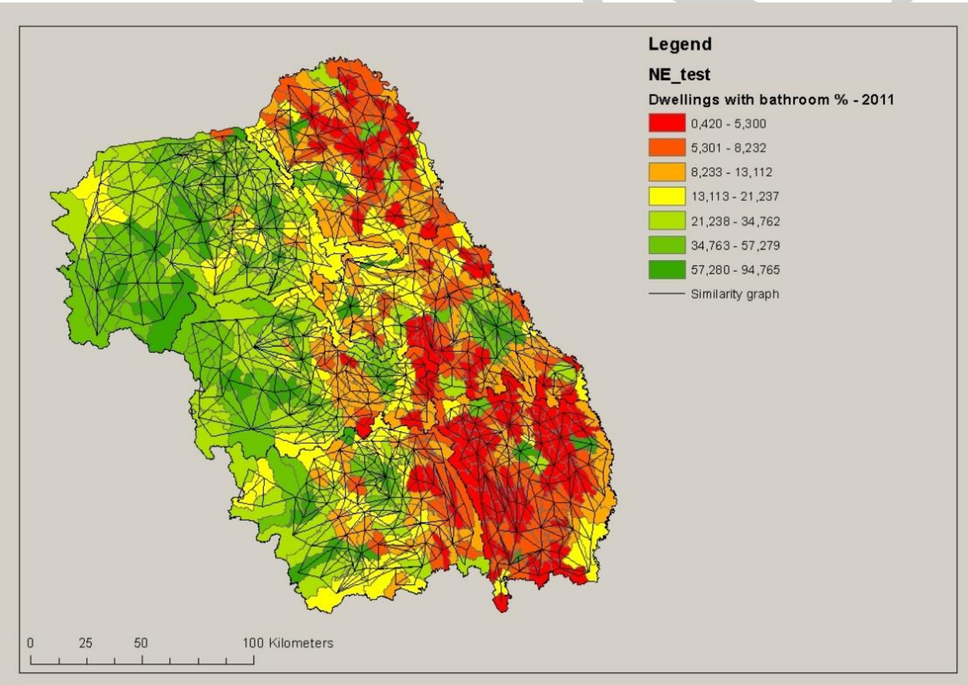

Fig. 6 Similarity graph overlayed on the cartographic illustration of one basic indicator

As it can be observed in the illustration above, the graph looks like an archipelago of links separated by NUTS3 limits. In other terms is a set of sub-graphs, a disconnected network. In average, an LAU2 in the study area has 5.5 neighbors, a value closed to 6 as predicted by the central-place model.

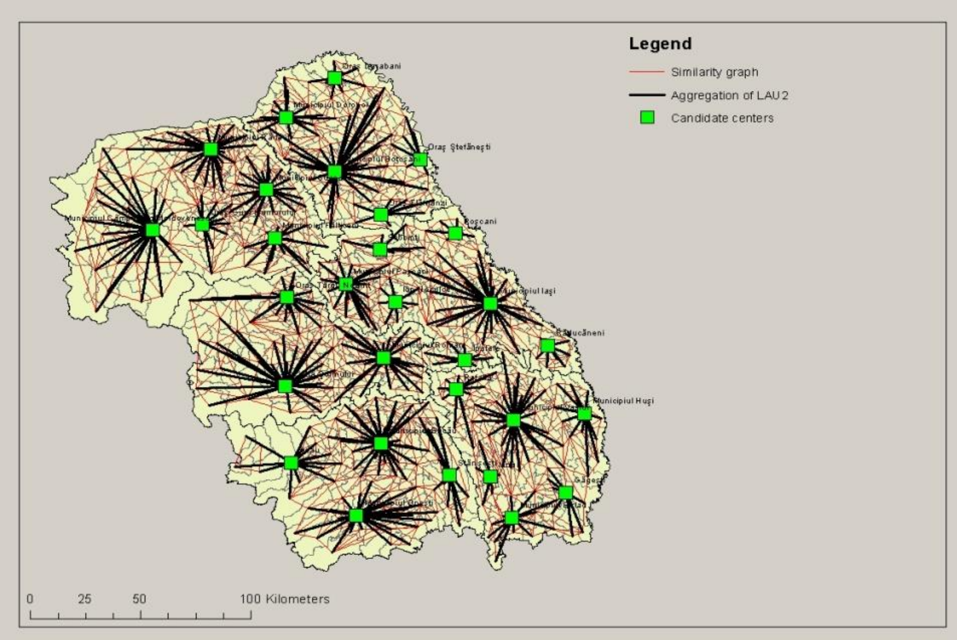

Fig. 7 Candidate centers and aggregation of LAU2

Our intention was to divide the LAU2 frame in 35 pseudo-LAU1 regions and an intermediate step is to select 35 possible candidate centers. These centers were selected using a p-median algorithm that maximizes the dissimilarities between the LAU2, within a pseudo-LAU1 district. As the graph was constrained by the NUTS3 limits, all the allocations were forced to follow this topological rule. In some cases, it will be more interesting not to obey to this constraint, especially when the geographical phenomena are independent to administrative limits.

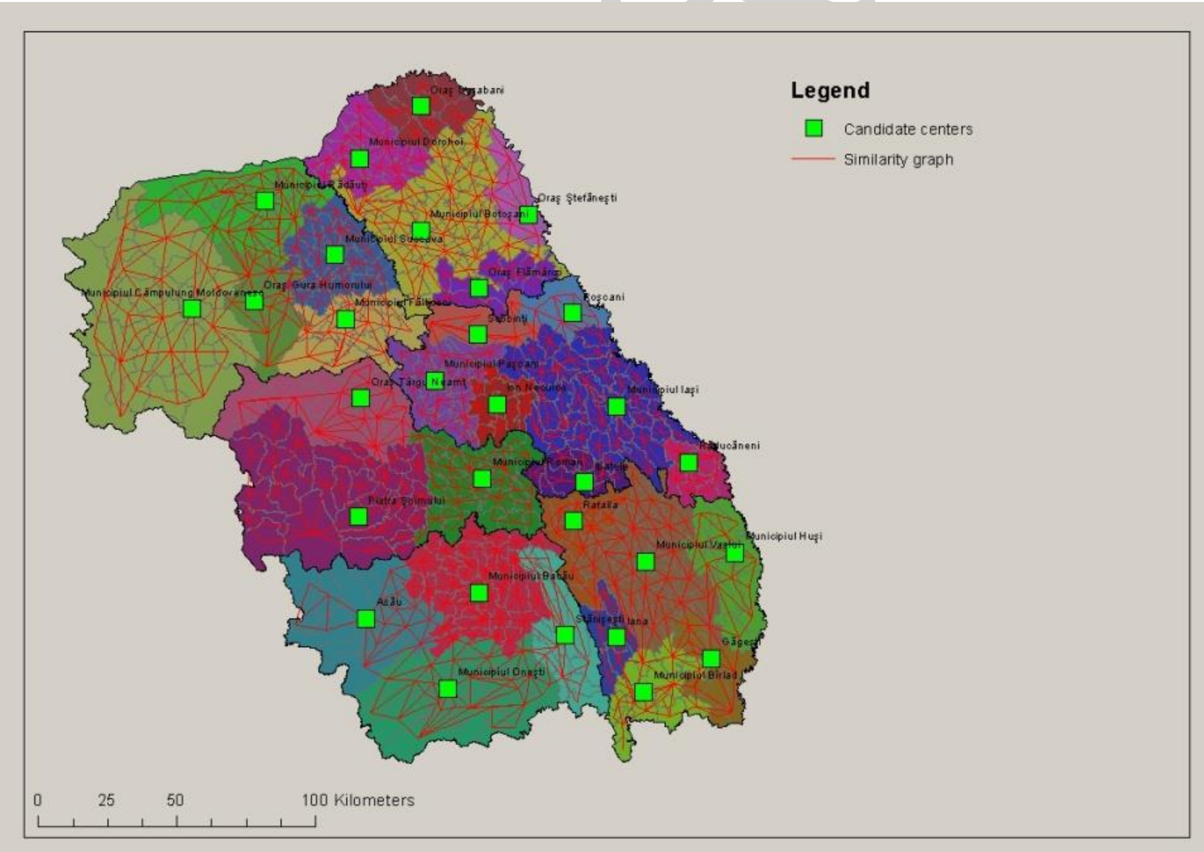

Fig. 8 Pseudo-LAU1 geometry with 35 candidate centers

The identification of the candidate centers allowed us to delineate a territorial frame of 35 pseudo LAU1. The territorial autocorrelation test for this number of districts/regions is -0.68, indicating that we deal with heterogeneous pseudo-LAU1. The candidate centers overlay the urban system and they regroup the rural LAU2 in the proximity.

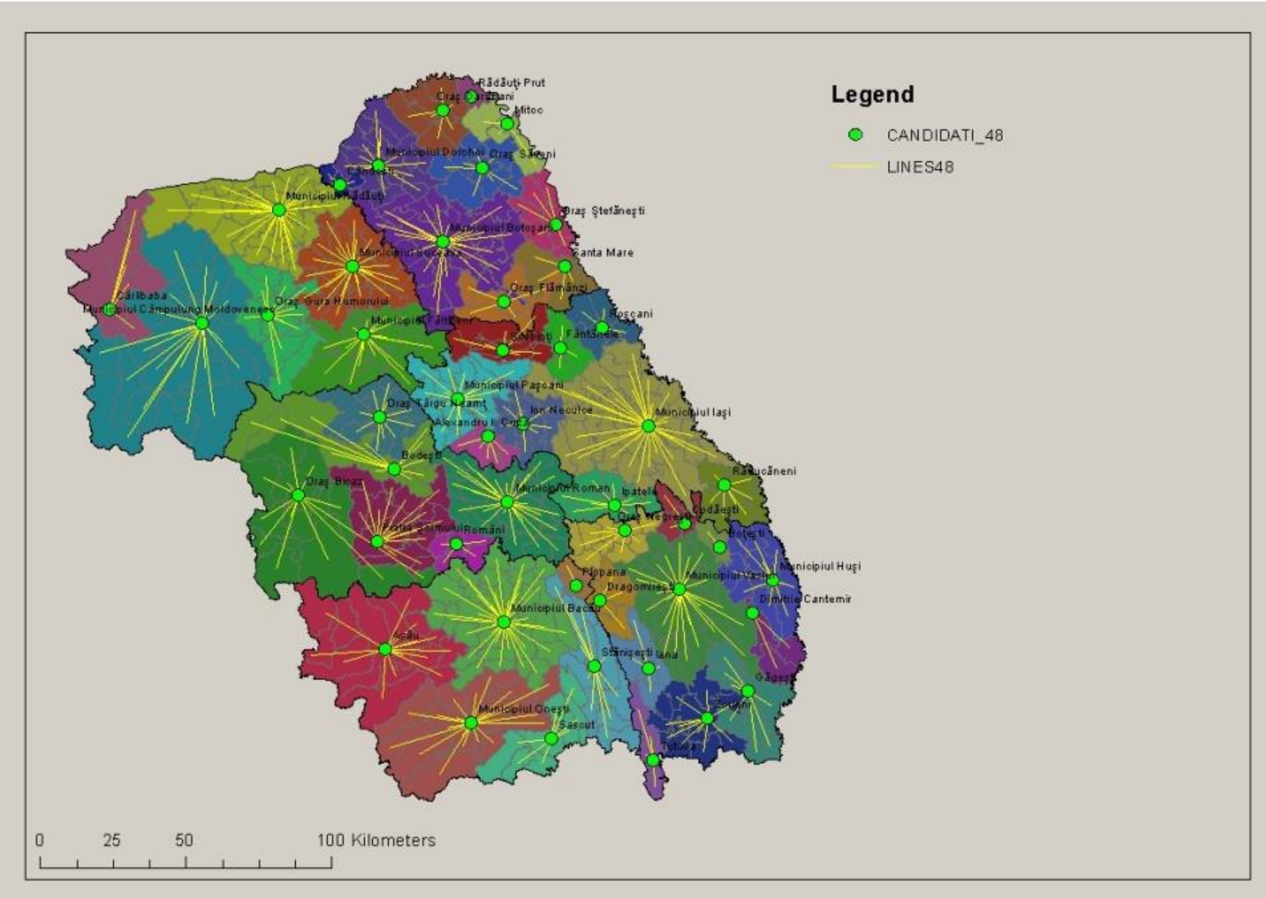

Fig. 9 Pseudo-LAU1 geometry with 48 candidate centers

With 48 districts, after the algorithm was implemented again, the territorial auto-correlation index has a value of -0.92. Some of the new created regions are very small in size and they can be aggregated with larger ones. In this situation, taking into account the values of the TAC index, an option should be made for the second frame of districting, based on 48 candidate centers.

After a first test on the North East Region of Romania, the moment has come to apply the algorithm on a larger area. Poland has an administrative basic frame composed by 2478 LAU2. Using only two land cover indicators, we have managed to apply the algorithm of districting in a reasonable time and the results can be observed in the next illustration.

Problems started to arise when we applied the districting technique to a graph of similarity, at regional scale. Five countries were selected for this exercise (Czech Republic, Hungary, Poland, Romania and Slovakia) and the time of implementation exceeded the *one cup of coffee GIS rule*. This fact is explained by the time needed to interrogate a matrix of 17 862 LAU2 by 17 862 LAU2. Applying the algorithm at national scale is a tactical advantage, especially when the creation of the districts demands a territorial constraint.

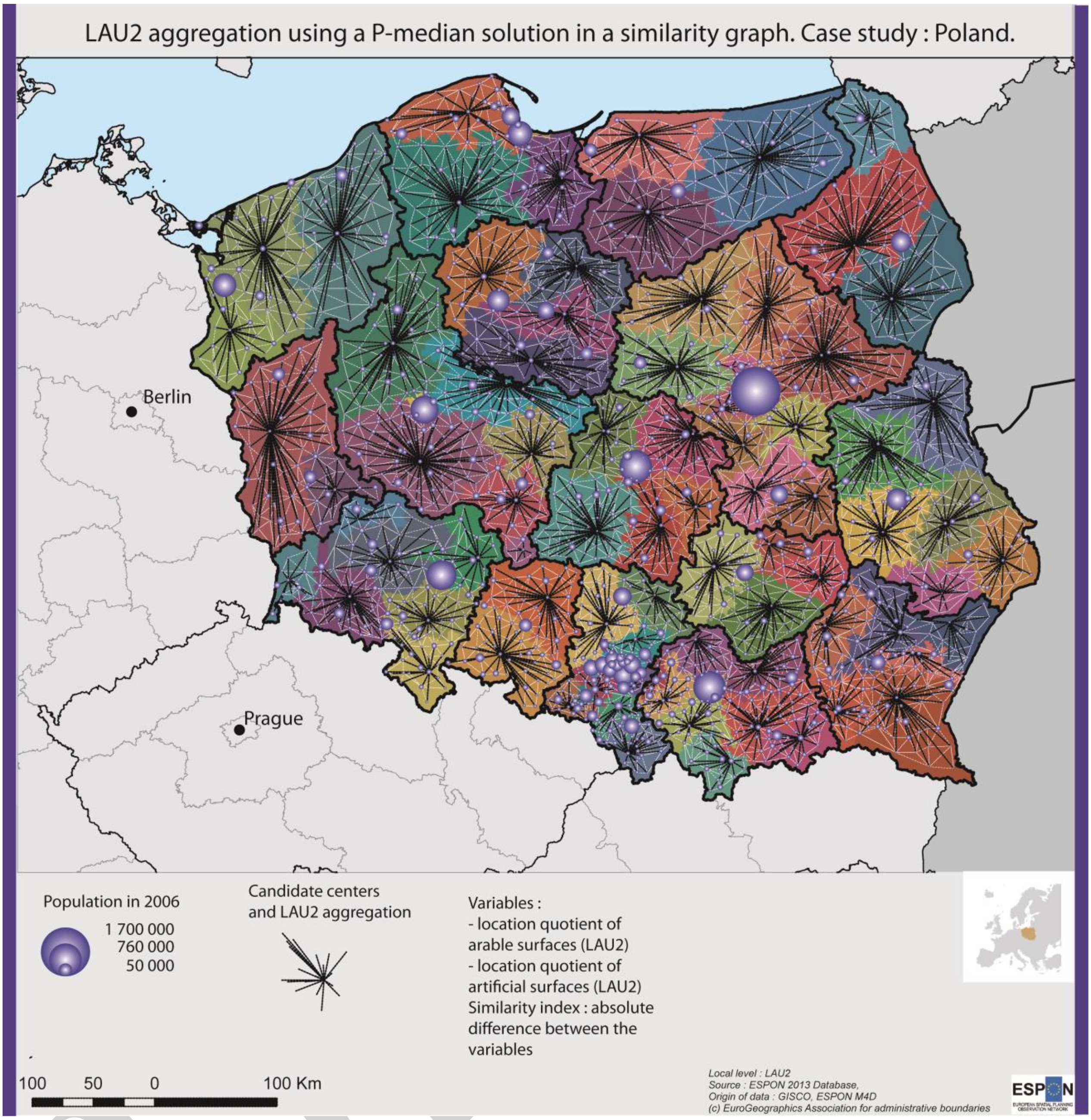

Fig. 9 Delineation of an alternative geometry for Poland

3 Conclusions and policy relevant key findings

The methodological exploration proposed in this technical report describes the algorithm and the methodological steps needed in order to delineate the pseudo-LAU1 frame. After the test of 4 different methods, some conclusions and policy relevant key findings should be pointed. There are three topics that should be stressed out in this final part:

1) Evaluation of the delineation methods. The method number 2 and the method number 4 are the most reliable tools for the districting problem. The method number 2 has the advantage of the speed and can be implemented on a large study area, being independent on descriptive indicators of the LAU2. The method number 4 is more productive in terms of alternative geometries that it can provide. Theoretically, there are at least four cases of figure to explore:

	Territorial constraints (a)	No territorial constraints (b)
Similarity (1) (maximizes the similarity between the LAU2)	It creates *homogeneous* pseudo LAU1 regions, not overlapping the NUTSx limits.	It creates *homogeneous* pseudo LAU1 regions, overlapping the NUTSx limits.
Dissimilarity (2) (maximizes the dissimilarity between the LAU2)	It creates *heterogeneous* pseudo LAU1 regions, not overlapping the NUTSx limits.	It creates *heterogeneous* pseudo LAU1 regions, overlapping the NUTSx limits.

Tab. 3 Crossing the similarity graph with the territorial constraints - possible combinations

However, there are some limits for this algorithm and this inconvenient is linked to the access to the local data. Without harmonized indicators for a large study area, the method number 4 will rely only on variables that are insufficient for the complex description of the local situations.

2) It is normal for the policy designers and decision makers to be skeptical when reading this technical report. Their skepticism is fuelled by some questions that are natural: how to use of this tool of districting/regionalization? The answer is not simple and we are aware that working with the official geometry is problematic enough, why use an alternative one? Used for case studies, the alternative geometry is a powerful tool that enables the mapping of regional specificities in a different way, excluding the common territorial visions and introducing new information.

3) The third conclusion is related to some vocabulary issues. Terms like districts, regions or pseuo-LAU1 overlay semantically and they are used as synonyms for the elements of the alternative geometry. In order to avoid repetition in a text that describes repetitive algorithms and methods, we have made an use of all them. In the same logic, terms like homogeneous or heterogeneous are used for their geo-statistical sense, without any policy connotation.

2 Exploring the relation between the LAU2 frame and the very important geographical objects

1. Theoretical background

It is not common to start a theoretical background with an image, but for the illustration of the complexity of the VIGO we will not obey to this rule. The draft map presents a collection or a sample of geographical objects in Poland. From road signs to hospitals and universities, all these objects are candidate to be a geographical VIP (very important point). Establishing a hierarchy among them demands a criterion and this criterion must not be fuzzy. From a geographical point of view, the set of criterion must clearly include a spatial dimension like scale, distance or territorial frequency. Depending on the scale of analysis, a school is a VIGO. However, at the zoom-out this quality will shade and other objects will replace it. In that case, in geographical terms the quality of being or not VIGO is dependent on what the French school of spatial analysis calls *jeu d'echelle*.

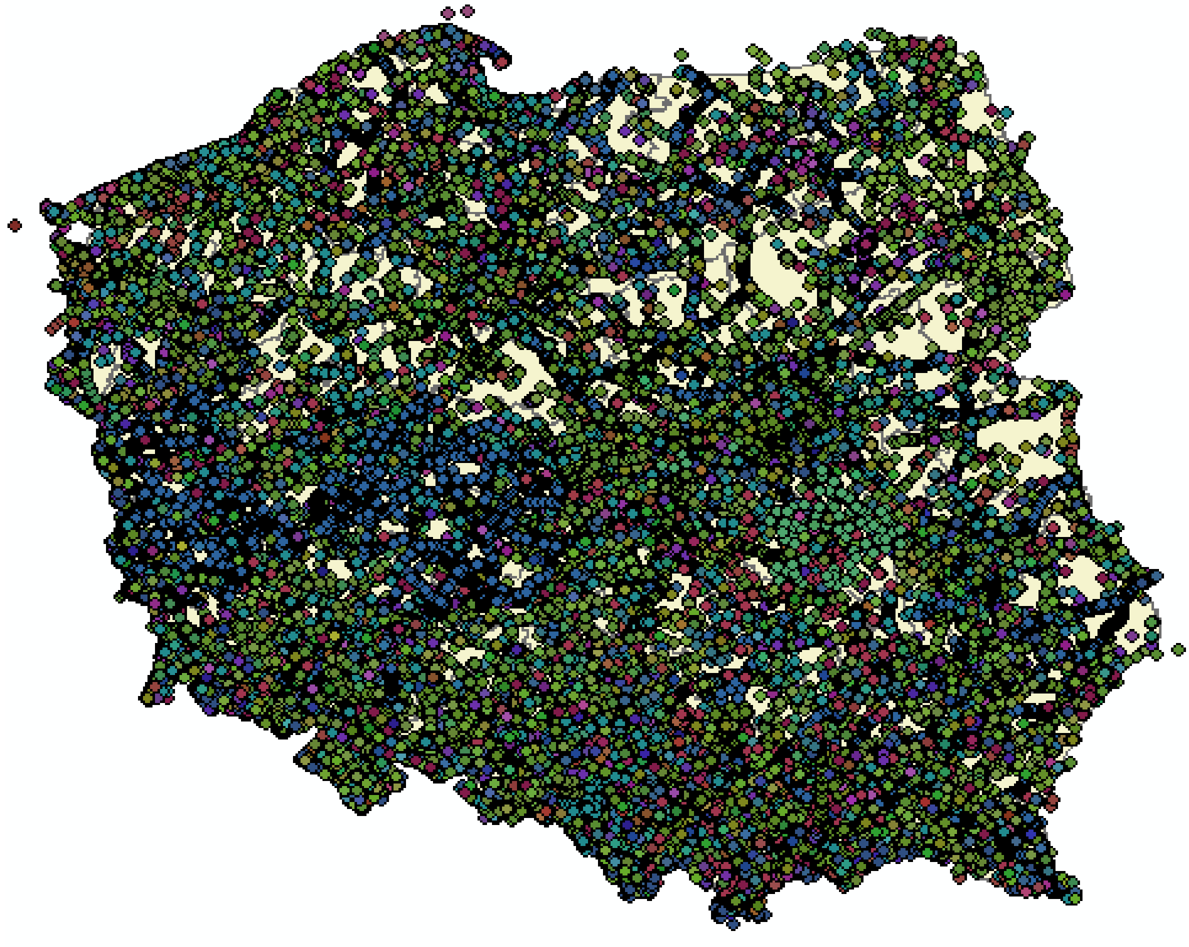

Fig. 10 Points of interest in Poland. Source : www.openstreetmap.org

If the label of VIGO is a matter of scale and as our intention is to observe the relation between them and the LAU2 frame, a better way to diagnose this relation is to focus on the methodological stakes related to this working package. For some methodological issues declined in the table, exploratory solutions were investigated and maps and datasets will be available, as case studies. The table is coded on the rows and columns and multiple issues/solutions were imagined:

·	·	· a	· b	·	· C	· d
·	·	· The VIGO is :	·	· and	· The study area is:	·
· 1	·	· a LAU2	· not a LAU2. It is a distinct geographical object, such as port or an airport.	·	· All the ESPON Space	· not all the ESPON Space. Focus on one country or a set of countries
· 2	· The VIGO has a *mass*	· A classical model of potential of interaction* can be implemented. For example : available urban population in a gaussian span of n km. **a&d&2**	· Huff model. A probabilistic model that describes the potential interaction between j and i.**b&c&1**	·	· Few datasets available from official sources. Reliability of data is uncertain and interfering with the updating problems. **c2**	· The relation between the LAU2 and the natural protected areas (environmental topic) was analyzed for selected countries in Eastern Europe. **b&d&2**
· 3	· The VIGO does not have a *mass***	· Not relevant. If the VIGO is a discrete point spatial pattern, it can be summurized by a LAU2 frame and reduced to the **a2** case.	· Average distance between any j and k nearest i. **b&c&3**	·	·	· The case study of France. A description of the relation between the LAU2 and the distribution of commercial services.**a&d&3**

2 Case studies and implementation

a&d&2: The VIGO is an LAU2, has a mass (demographic mass in 2011) and does not cover all the ESPON space. It is a case study on selected countries in Eastern Europe. The definition of the urban LAU2 respects the national definitions. The values of the model can be optionally constrained by NUTSx borders. The model of potential of interaction can be based on Euclidean distances or network distances. In the last case, a time-impedance may be added. The multiple possibilities complicate the methodology, but will provide at least three local indicators.

b&c&1: The Huff model was implemented in a GIS, in order to evaluate the potential of interaction between a VIGO (airports) and the LAU2 from the ESPON Space. The results were impossible to be calculated for Greece and Bulgaria, due to topological errors or network missing. The model and the associated tool can also be easily implemented to other VIGO analysis. The Huff model is applied since the early years of quantitative geography and it creates iso-lines of equal probability of interaction. The methodological challenges are represented by the translation of the information from the isolines to the LAU2 frame.

b&d&2: The environmental dimension of the relation between the VIGO and the LAU2 was intercepted by this case study. The problem is even more interesting when we take into account the fact that some of the European natural areas are overlapping the LAU2 and that these VIGO need a double spatial approach - as polygons and as points (centroids). A first attempt to solve this problem was limited by the high number of protected areas. Only a reduced amount of LAU2 in the selected countries for the case studies is not covered by the VIGO polygons. Filtering the protected areas dataset is a distinct possibility, in order to evaluate their territorial relation with the VIGO.

b&c&3: In this case study it is analysed the relation between VIGO (highway entries) and the LAU2. The connectability to the major transportation corridors is key information in the evaluation of the territorial competitiveness. It also allows us to compare the LAU2 in different territorial contexts. The application of the method calculates the road-distances between the LAU2 centroids and the highway entries. Extremely dependent on the configuration of the network and on the quality of the road segments, this exercise presents potential in a chronological context.

a&d&3: A case study that provides relevant information on the relation between the everyday life VIGO (commercial services) and the LAU2, at different scales. Using distances and the cumulated population by these VIGO in the proximity, we can easier evaluate their territorial impact as equipments. As study area, we have made an option for France. With more than 36 000 LAU2 is a massive challenge for the evaluation of the relation between the LAU2 and the VIGO. The classification of the LAU2 was inspired from the multi-scalar analysis implemented in some of the ESPON tools (HyperAtlas), but adapted to the nature of the indicators we created and ignoring the territorial belonging of the LAU2.

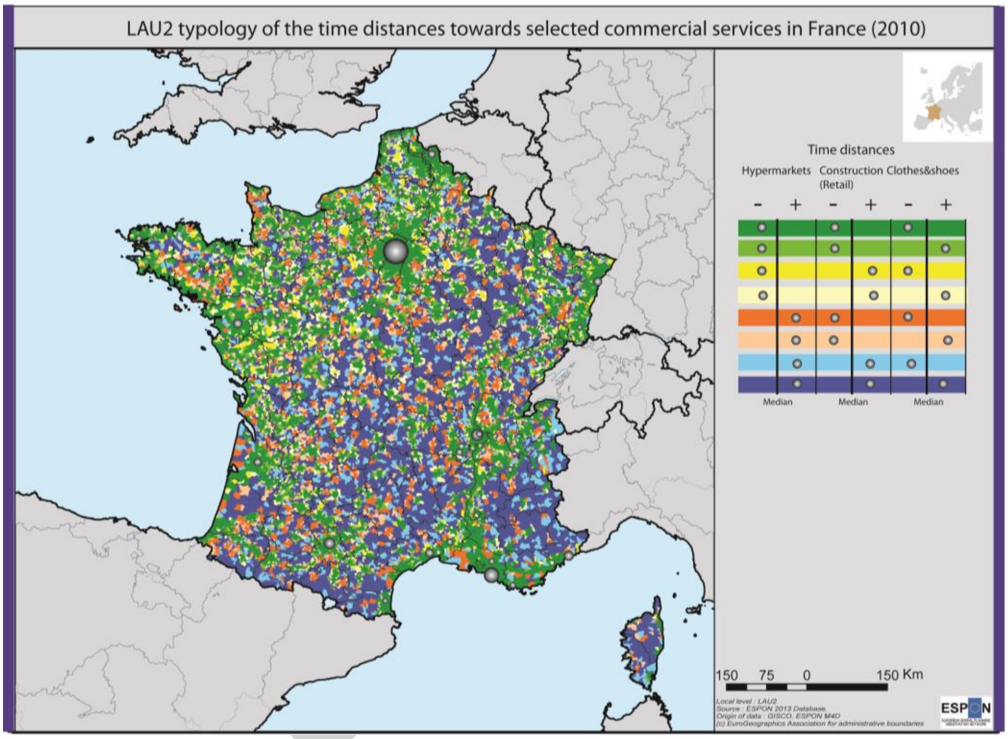

Fig. 11 Illustration of the relation between the VIGO (selected services) and the LAU2 frame in France

For a better illustration of the methodological intentions developed in this technical report, we will insist on the exploration of the relation between the VIGO and the LAU2 geometry, at a macro-regional scale. Our intention was to measure the amount of population available at local level and to relate this demographic mass with the distance towards the closest airport. Cumulating the population by time distance is the solution we proposed because it can provide some new information regarding the territorial role of airports for the areas they serve as transportation facilities.

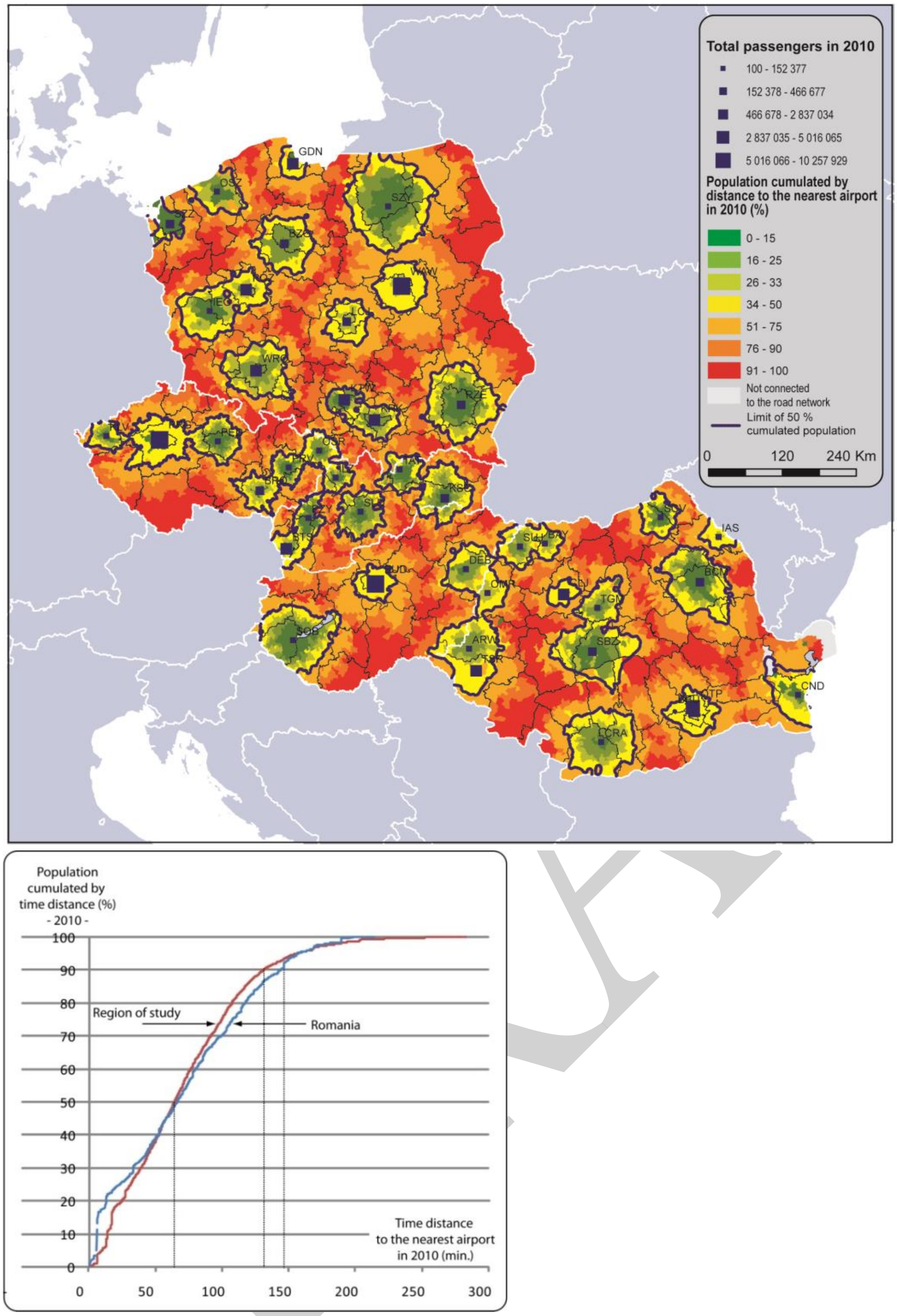

Fig.12 Population cumulated by distance to the closest airport for selected countries in the Eastern Europe.

The last illustration is a case study of spatial interaction potential in Poland. In a general approach of the model, the Gaussian kernel is fixed. In the evaluation of the relation between the LAU2 and the VIGO, we have managed to vary the size of this kernel, according to the mass of the LAU2. This approach provides methods for smoothing the data and for a better interception of the territorial role played by the LAU2.

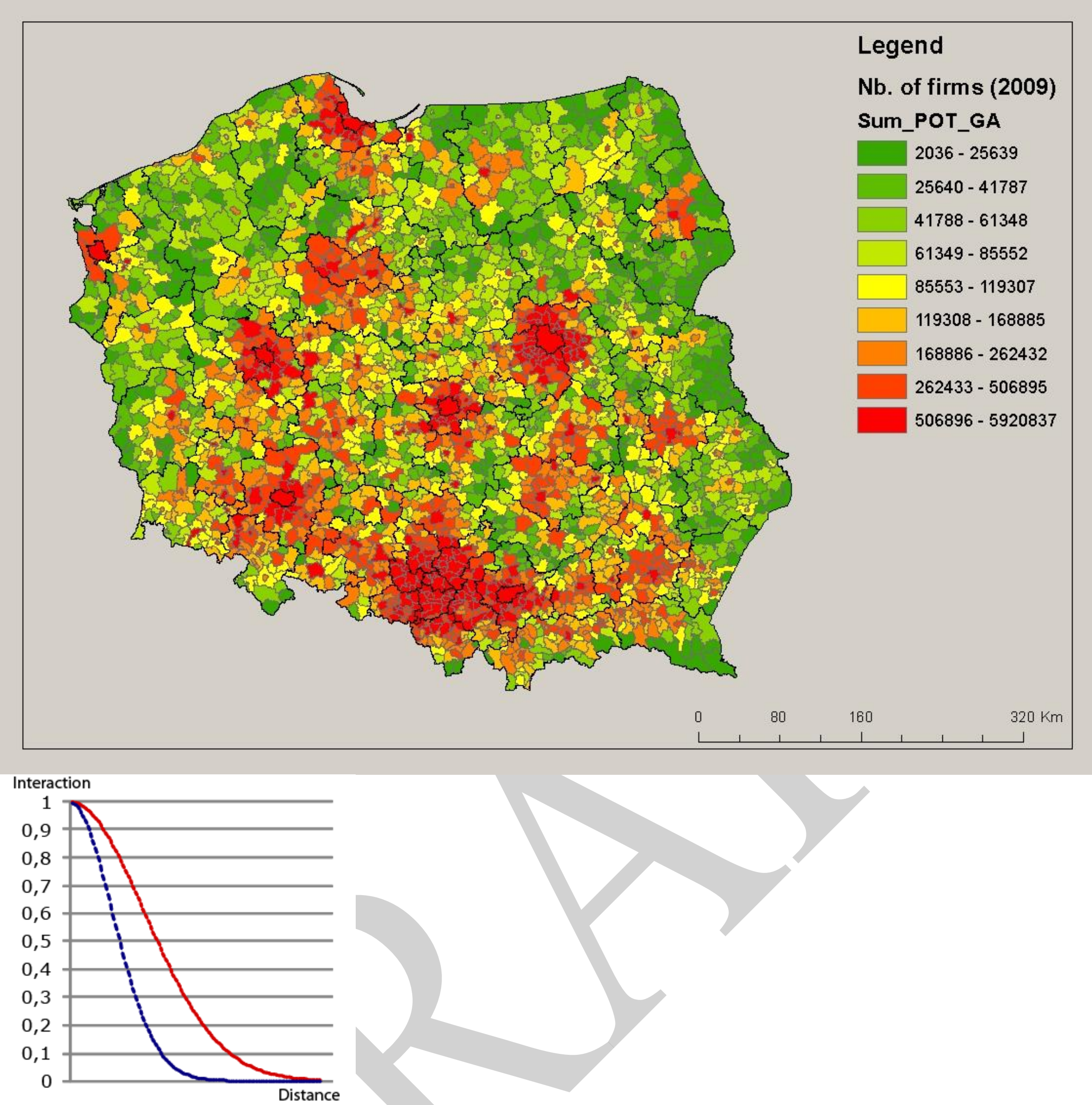

Fig. 13 Output of the potential of spatial interaction model with variable Gaussian span.

3 Conclusions

The major methodological problems to be solved are related to the quantity of information needed to be treated for each analysis. The limits of the analysis are often linked to the software architecture and ignoring them demand fragmented and iterated procedures. Basically, the methodological challenges are reduced to the translation of the geographical models (Huff, potential of interaction, average distance to the k nearest VIGO etc.) in steps of implementation in GIS. It might look a trivial problem, but it is not, especially when we need to weight the VIGO with a mass variable. As in the case of the alternative geometry, tools were designed to accelerate the model's implementation. These tools will be described in a technical report, together with the methodological approach.

Bibliography

Coombes, M., Casado-Díaz, J.M., Martínez-Bernabeu, L., Carausu, F., 2012, Study on comparable Labour Market Areas (Final research report), Framework contract 6001. 2008.001 - 2009.065

Decroly J.M., Grasland C., (1993), "Frontières, systèmes politiques et fécondité en Europe", *Espace Population Sociétés, 2,* pp. 135-152

Guo, D. and H. Wang (2011). "Automatic Region Building for Spatial Analysis", Transactions in GIS, 15(s1), pages 29-45. [Journal Link]

Guo, D. (2008). "Regionalization with Dynamically Constrained Agglomerative Clustering and Partitioning (REDCAP)".International Journal of Geographical Information Science. 22(7), pp. 801-823.

Guo, D. (2009). "Greedy optimization for contiguity constrained hierarchical clustering", The 4th International Workshop on Spatial and Spatiotemporal Data Mining, IEEE International Conference on Data Mining, pp. 591-596.

Haggett, P., Arnold, E., 1965, London, Locational Analysis in Human Geography

Sanders, L., 2001, Paris, Modèles en analyse spatiale, Lavoisier