Duque, Juan Carlos; Ramos, Raúl

Conference Paper

Design of homogenous territorial units: a methodological proposal

44th Congress of the European Regional Science Association: "Regions and Fiscal Federalism", 25th - 29th August 2004, Porto, Portugal

Provided in Cooperation with:
European Regional Science Association (ERSA)

Suggested Citation: Duque, Juan Carlos; Ramos, Raúl (2004) : Design of homogenous territorial units: a methodological proposal, 44th Congress of the European Regional Science Association: "Regions and Fiscal Federalism", 25th - 29th August 2004, Porto, Portugal, European Regional Science Association (ERSA), Louvain-la-Neuve

This Version is available at:
https://hdl.handle.net/10419/116928

44th Congress of the European Regional Science Association
Porto, 25th-29th August 2004

DESIGN OF HOMOGENOUS TERRITORIAL UNITS: A METHODOLOGICAL PROPOSAL

Juan Carlos Duque, Raúl Ramos

Summary:

One of the main questions to solve when analysing geographically added information consists of the design of territorial units adjusted to the objectives of the study. In fact, in those cases where territorial information is aggregated, ad-hoc criteria are usually applied as there are not regionalization methods flexible enough. Moreover, and without taking into account the aggregation method applied, there is an implicit risk that is known in the literature as Modifiable Areal Unit Problem (MAUP) (Openshaw, 1984). This problem is related with the high sensitivity of statistical and econometric results to different aggregations of geographical data, which can negatively affect the robustness of the analysis.

In this paper, an optimization model is proposed with the aim of identifying homogenous territorial units related with the analyzed phenomena. This model seeks to reduce some disadvantages found in previous works about automated regionalisation tools. In particular, the model not only considers the characteristics of each element to group but also, the relationships among them, trying to avoid the MAUP. An algoritm, known as RASS (*Regionalization Algorithm with Selective Search*) it also proposed in order to obtain faster results from the model. The obtained results permit to affirm that the proposed methodology is able to identify a great variety of territorial configurations, taking into account the contiguity constraint among the different elements to be grouped.

Keywords: Zone design, Modifiable Areal Unit Problem, Optimisation, Contiguity constraint.

JEL codes: R22, R12, C61

Author:

Name: Juan Carlos Duque
Center: Grup d'Anàlisi Quantitativa Regional (Universitat de Barcelona)
Postal address: Espai de Recerca en Economia.
Avda. Diagonal 690
08034 Barcelona
Telephone: +34 +934037043
Fax: +34 +934021821
E-mail: jduque@ub.edu

Name: Raúl Ramos
Center: Grup d'Anàlisi Quantitativa Regional (Universitat de Barcelona)
Postal address: Espai de Recerca en Economia.
Avda. Diagonal 690
08034 Barcelona
Telephone: +34 +934021984
E-mail: rramos@ub.edu

1. INTRODUCTION AND OBJECTIVES

The interest for geographical information technologies has considerably increased during the last three decades. Nowadays, geographical information is no more exclusive of government and public administrations (in the areas of planning, demography and topography) thanks to the development of computer tools, in software and hardware, that have made possible to use this information in firms and in academic areas.

This kind of statistical information is usually published at different territorial levels with the aim of providing information of interest for all the potential users. When using this information, researchers have two different alternatives to define the basic territorial units that will be used in the study: first, to use geographical units designed following normative criteria (the officially established territorial units such that towns, provinces, etc.) or, second, to apply an analytical criteria to design geographical units directly related with the analysed phenomena.

"Normative regions are the expression of a political will; their limits are fixed according to the tasks allocated to the territorial communities, to the sizes of population necessary to carry out these tasks efficiently and economically, or according to historical, cultural and other factors. Whereas analytical (or functional) regions are defined according to analytical requirements: functional regions are formed by zones grouped together using geographical criteria (e.g., altitude or type of soil) or/and using socio-economic criteria (e.g., homogeneity, complementarity or polarity of regional economies)" (Eurostat, 2004).

The majority of empirical studies tend to use geographical units based on normative criteria for several reasons: this type of units are officially established, they have been traditionally used in other studies, its use makes comparison of results easier and can be less criticized. But at the same time, in those studies using this type of units an "Achilles' heel" can exist if they are very restrictive or inappropriate for the considered problem. For example, if we are analysing phenomena as regional effects of monetary and fiscal policy, how will the results be affected if the aggregated areas[1] in

[1] In this paper, we will use the term "area" to denote the smallest territorial unit. The aggregation of areas will form a "region" and the aggregation of regions will cover the whole considered territory.

each region are heterogeneous? can those results change if the areas are redefined in a way that each region contains similar areas?.

The above mentioned situation could be improved through the use of regionalisation processes to design geographical units, based on analytical criteria, by aggregating geographical units of small size[2], but without arriving at the upper level, or combining information from different levels[3]. In this context, the design of analytical geographical units should consider the following three fundamental aspects:

a. *Geographical contiguity*: The aggregation of areas into regions such that the areas assigned to a region must be internally connected or contiguous.
b. *Equality*: In some cases, it is important that designed regions are "equal" in terms of some variable (for example population, size, presence of infrastructures, etc).
c. *Interaction between areas*: Some variables do not exactly define geographical characteristics that can be used to aggregate the different areas, but perhaps they describe some kind of interactions among them (for example, distance, time, number or trips between areas, etc). These variables can also be used as interaction variables using some dissimilarity measure between areas in terms of socio-economic characteristics. The objective in this kind of regionalisation process is that areas belonging to the same region are as homogeneous as possible with regard to the specified attribute(s).

Unfortunately, in most cases, the aggregation of territorial information is usually done using "*ad-hoc*" criteria due to the lack of regionalisation methods with enough flexibility. In fact, most of these methods have been developed to deal with very particular regionalisation problems, so when applied in other contexts the results could

---

[2] Apart from aspects such as the statistical secret or other legislation about the treatment of statistical data, according to Wise *et al*. (1997), this kind of territorial units are designed in such a way as to be above minimum population or household thresholds, to reduce the effect of outliers when aggregating data or to reduce possible inexactities in the data, and to simplify information requirements for calculations or to facilitate its visualisation and interpretations in maps.

[3] See, for example, Albert *et al*. (2003), who analyse the spatial distribution of economic activity using information with different levels of regional aggregation, NUTS III for Spain and France and NUTS II for the rest of countries, with the objective "using similar territorial units". López-Bazo *et al.* (1999) analyse inequalities and regional convergence at the European level in terms of GDP per capita using a database for 143 regions using NUTS II data for Belgium, Denmark, Germany, Greece, Spain, France, Italy, Netherlands and Portugal, and NUTS-I for the United Kingdom, Ireland and Luxemburg with the objective of ensuring the comparability of geographical units.

be very restrictive or inappropriate for the considered problem. However, and with independence of the applied territorial aggregation method, there is an implicit risk, known in the literature as "Modifiable Areal Unit Problem" (Openshaw, 1984), and related with the sensitivity of the results to the aggregation of geographical data and its consequences on the analysis.

The main objective in this paper is to implement a new automated regionalisation tool to design homogeneous geographical units directly related with the analysed phenomena that overcomes some of the disadvantages of available methodologies.

Thus, the specific objectives are:

a. To formulate the regionalisation problem as a linear optimisation model where it can be taken into account not only the areal characteristics but also their non metric relationships and their contiguity relationships.
b. To propose a heuristic model that allow to solve bigger regionalisation problems, incorporating in its search procedure the own characteristics of a regionalisation process.
c. To compare, in terms of homogeneity degree, the analytical regions designed by applying the regionalisation model proposed in this paper with another regionalisation method based on normative criterion. To due this comparison provincial time series of unemployment rate in Spain will be used.

The paper is organised in the following sections: in section 2 the literature about the different regionalisation methods are briefly summarised; in section 3 the proposed lineal optimisation model for automated regionalisation is described; section 4 introduces an algorithm to deal with more complex regionalisation problems, and, last, the most relevant conclusions of the paper are presented in section 5.

2. REVISION OF THE LITERATURE

In this section the most relevant methodologies for territorial aggregation will be briefly summarised. This summary will be focused on those methodologies with a

higher impact in the specialised literature and on those ones that have been tested satisfactorily in real problems.

Most of these methodologies use techniques based on cluster analysis[4]. In this context, the problem of aggregation of spatial data is considered as a particular case of clustering where geographical contiguity among the elements to be grouped should be considered. This particular case of clustering methods is usually known as contiguity-constrained clustering or simply regionalisation problem. A detailed summary of these aggregation methodologies can be found in Gordon (1999) and for the case of constrained clustering in Fisher (1980), Murtagh (1985) and Gordon (1996).

Thus, regionalisation algorithms can be categorized under three methodological strategies: two-stages aggregation; the inclusion of geographical information in the set of classification variables; and, the use of additional instruments to control for the geographical contiguity constraint.

2.1. Two stages aggregation.

This strategy consists of splitting the aggregation process in two stages. The first stage consists of applying a conventional clustering model without take into account the contiguity constraint, and, in a second stage, the clusters are revised in terms of geographical contiguity. With this methodology, if the areas included in the same cluster are geographically disconnected, those areas are defined as different regions (Ohsumi, 1984).

Two conventional clustering algorithms can be used in this context: hierarchical or partitioning.

2.1.1. Hierarchical algorithms.

They are usually applied when the researcher is interested in obtain a hierarchical and nested classification (for every scale levels), that is usually summarised using dendograms[5]. The main disadvantage of using hierarchical clustering algorithms,

[4] Multivariate statistical tool widely used to classify elements in terms of their similarities or dissimilarities (Jobson, 1991).

[5] Graphical representation of the solutions of hierarchical cluster (Gordon, 1996).

without considering the high computational requirements (Wise *et al*., 1997), is the high probability of obtaining local optimum due to the fact that once two elements have been grouped in an aggregation level, they would not return to be evaluated independently in higher aggregation levels (Semple and Green, 1984). On the other hand, the main advantage that should be highlighted is that there is no need to specify initial partitions to apply the algorithm (Macmillan and Pierce, 1994).

2.1.2. Partitioning algorithms.

More used in regionalisation processes is the K-means clustering procedure, which belongs to partitioning clustering category, this iterative technique consists of selecting from elements to be grouped, a predetermined number of *k* elements that will act as centroids (the same number as groups to be formed). Then, each of the other elements is assigned to the closest centroid.

The aggregation process is based on minimizing some measure of dissimilarity among elements to aggregate in each cluster. This dissimilarity measure is usually calculated as the squared Euclidean distance from the centroid of the cluster[6], see equation 2.1.

$$\sum_{m \in c} \sum_{i=1}^{N} \left(X_{im} - \overline{X}_{ic} \right)^2 \qquad (2 \cdot 1)$$

Where X_{im} denotes the value of variable *i* (*i*=1..*N*) for observation *m* (*m*=1..*M*), and $\overline{X}_{ic}$ is the centroid of the cluster *c* to which observation *m* is assigned or the average X_i for all the observations in cluster *c*.

K-means algorithm is based on an iterative process where initial centroids are explicitly or randomly assigned and the other elements are assigned to the nearest centroid. After this initial assignation, initial centroids are reassigned in order to minimize the squared Euclidean distance. The iterative process is terminated if there is not any change that would improve the actual solution.

---

[6] A detailed summary of these aggregation methodologies can be found in Gordon (1999) and for the case of constrained clustering in Fisher (1980), Murtagh (1985) and Gordon (1996).

It is important to note that the final solutions obtained by applying K-means algorithm depend on the starting point (the initial centroids designation). This fact makes quite difficult to obtain a global optimum solution.

Finally, when K-means algorithm is applied in a two stages regionalisation process, it will be possible that the required number of regions to design will be not necessarily equal to the value given to parameter *k* as areas belonging to the same cluster have to be counted as different regions if they are not contiguous. So, different proofs have to be done with different values of *k* (lower than the number of desired regions), until contiguous regions are obtained. In some cases could be impossible to obtain the desired number of contiguous regions.

Among the advantages of two stages aggregation methodology, Openshaw and Wymer (1995) highlight that the homogeneity of the defined regions is guaranteed by the first stage. Moreover, this methodology can also be useful as a way to obtain evidence of spatial dependence among the elements. However, taking into account the objectives of the regionalisation process, the fact that the number of groups depends on the degree of spatial dependence[7] and not on the researcher can be an important problem.

2.2. Inclusion of geographical information as classification variables.

The second strategy consists of including as classification variables the geographical coordinates of centroids representing the areas to be grouped (Perruchet, 1983, Webster and Burrough, 1972). In this strategy, as a way to force the geographical contiguity, the geographical coordinates are included in the calculation of dissimilarities between areas and, next, conventional classification algorithms are applied.

This kind of approach has been implemented in the SAGE system (*Spatial Analysis in a GIS Environment*) (Haining *et al*., 1996). In its regionalisation algorithm, this system uses an objective function formed by three components, the first controls the intra-group variance taking into account the non spatial attributes, the second, as geographical component, includes the sum of the distances from areal centroids to the cluster centroids in order to force geographical contiguity, and the third component is a

[7] When the spatial dependence is higher (lower) there will be a trend towards the creation of less (more) regions.

deviation measure between the regional value of an attribute and its average value. A different weight is assigned to each of these components in the objective function in order to obtain a unique value to minimise. The regionalisation procedure is based on a partitioning algorithm K-means (Andemberg, 1973).

Calciu (1996) uses the same territorial aggregation strategy, referring to it as "*contrainte spatiale implicite"* (implicit spatial constraint), which incorporates as geographical variables the Cartesian coordinates, conveniently transformed, of the points representing each area. This author is in favour of applying a hierarchical classification algorithm, where the inclusion of the coordinates permits to obtain an improved geographical continuity, although it implies some lost in terms of intragroups homogeneity in relation to the case where the hierarchical algorithm is applied without considering these geographical variables.

The main inconvenient associated to this methodology are the difficulty of treating simultaneously variables expressed in different measure units and the definition of objective weights for each of the variables, specially the geographical ones as the weights should be strong enough to guarantee that geographical contiguous regions are formed (Wise *et al*., 1997).

Another disadvantage is that the final solution can change depending on the applied method to localise the centroid that represents each of the areas to be grouped, especially in those cases where the areas are considerably big (Horn, 1995, Martin *et al.*, 2001).

2.3. Additional instruments to control for the continuity restriction.

The last, but perhaps the most used strategy to solve territorial aggregation problems, consists of controlling the geographical contiguity constraint using additional instruments as the contact matrix or its corresponding contiguity graph. Contact matrix is a binary matrix with elements c_{ij}, where c_{ij} takes value 1 if areas i and j share a border; and 0 otherwise. In the contiguity graph the areas to be grouped are represented as nodes and arcs represent the adjacency relationship between them[8].

[8] For a more detailed description of the methods for the elaboration of this kind of graphs, see Gordon (1996, 1999).

The elements above are used to adapting conventional clustering algorithms, hierarchical or partitioning, with the objective of respecting the continuity constraint.

The main problem with adapted hierarchical algorithms in the context of regionalisation processes is that there can be breaks in monotonicity among elements. This problem is known as reversals: the distance between two objects can be higher than the distance between the union of this object with a third one (Calciu, 1996, Gordon 1996, Ferligoj and Batagelj, 1982). It makes difficult the interpretation of classification.

In adapted partitioning algorithms, contact matrices or contiguity graphs have mainly been applied into two different methodologies: mathematical programming and iterative algorithms.

Regarding to mathematical programming, Macmillan and Pierce (1994) define the regionalisation problem as an optimisation problem where, given a predetermined number of groups to form, the solution will define the optimum territorial aggregation. The proposed solution by these authors to ensure the geographical continuity consists of exponenciating the contact matrix, taking into account that for the formation of a region with n continuous areas is necessary that the $(n\text{-}1)^{th}$ power of the contact matrix does not contain null elements. This solution implies that the feasible space defined by the constraints is non-convex and, as a result, the objective function is likely to get trapped in a local optimal solution.

Cutting algorithms for graph partitionig are another way to see the regionalisation problem from a mathematical programming point of view. In these models, the contiguity graph has associated in their arcs a value of dissimilarity between areas, i.e. $G=(V,E)$, with a weight function $w : E \rightarrow N$.

The cutting algorithms looks for a partition of the node set V into k disjoint sets $F=\{C_1, C_2,..., C_k\}$ where k is integer and $k \in [2..|V]$. Thus, in a regionalization process, the idea could be to maximice the isolation between groups, so the objective in a "maximum k-cut" is to maximice the sum of the weight of the edges between the disjoint sets, i.e.:

$$\sum_{i=1}^{k-1} \sum_{j=i+1}^{k} \sum_{\substack{v_1 \in C_i \\ v_2 \in C_j}} w(\{v_1, v_2\}) \qquad (2\cdot2)$$

Where v_1 and v_2 are the endpoints of an arc[9].

Another method, cited by Neves *et al.* (2001), consists of the reduction of the contiguity graph (G=(V,E)) where each arc has associated a value of dissimilarity between areas (weight function $w : E \rightarrow N$). The reduction makes a progressive elimination of arcs until a minimum spanning tree is obtained. The main point of this representation is that the elimination of one arc at a time implies the partition of the graph in intraconnected, but not interconnected, subgroups (Ahuja *et al*., 1993).

One disadvantage of the regionalisation methodologies modelling the dissimilarity relationships using the arcs of the contiguity graph is related with the fact that an important number of dissimilarity relationships between areas that are not contiguous are not being considered.

Taking into account that the resolution of this kind of problems using conventional optimisation methods is extremely complex[10], other methodologies have been developed in the field of regionalisation that have been very effective in those cases where the number of elements to group is very high. Among these different solutions, the algorithms known as *Iterative Relocation Algorithms* have been widely analysed. These methods try to find the best regional configuration using as a starting point a non-optimal configuration[11] and, next, different movements of areas between regions are done with the objective of improving the objective function. Ferligoj and Batagelj (1982) provide different iterative reallocation algorithms that allow moving an area to a different region only if contiguity constrains are satisfied.

Algorithms such as the *Automatic Zoning Procedure* (AZP) (Openshaw, 1977), the *Land Allocation Problem* (Benabdallah and Wright, 1992), the *Redistricting Problem* (Macmillan and Pierce 1994) and the *Regional Partitioning Problem* (Horn, 1995) have been used in the literature related with the particular case of splitting a country in administrative areas or electoral districts such that the final regionalisation minimises the effects of the Modifiable Areal Unit Problem (MAUP)[12].

---

[9] A compendium of models related to network design can be found in Crescenzi and Kann (2004).

[10] Openshaw (1984) calculated that to aggregate 1,000 areas in 20 regions there are 101,260 different solutions. For more information about combinatorial problems, see Aarts and Lenstra (1997).

[11] Different alternatives to determine the initial solution can be found in Wise et al. (1997).

[12] Openshaw defined the problem of the Modifiable Areal Unit Problem (MAUP) as a potential source of error that can affect the results of those studies based in geographical aggregated information as these results could vary in function of the configuration of this aggregation. The MAUP is related with two

Iterative Relocation Algorithms have been improved using heuristics that permit a better search among the different feasible solutions and to avoid the risk of getting trapped into a local optimum. The most used heuristics in this context are the *Simulated Annealing* (AZP-SA) and the *Tabu Search Algorithm*[13,14] (AZP-TABU), proposed by Openshaw and Rao (1995), and the *Anneal Redistricting Algorithm* proposed by Macmillan and Pierce (1994).

The methodologies of constrained clustering where additional instruments are included, have as a common characteristic that the relationships between the areas to group are symmetric. In this sense, Ferligoj and Batagelj (1983) have developed agglomerative algorithms where asymmetric relationships can be considered.

All the methods presented above are "supervised" models, which means that the researcher knows *a priori* the data structure of the analysed phenomenon. But there are other unsupervised models that can be useful when the researcher wants to analyse a big amount of data and there is not enough information of the factors that can affect the system. In these cases, one possibility consists of applying a non-parametric analysis of data that will permit to find the patterns and relationships among the considered elements. One of the most known applications of these methods in the field of regionalisation is *Self Organization Maps* (SOM) proposed by Kohonen (1984). There is no consensus among researchers about the validity of this methodology, originally developed in the field of artificial intelligence, due to the lack of a theoretical basis that difficult the interpretation of the results (Openshaw, 1992).

A summary of the different methodologies in this section can be found in table 2.1.

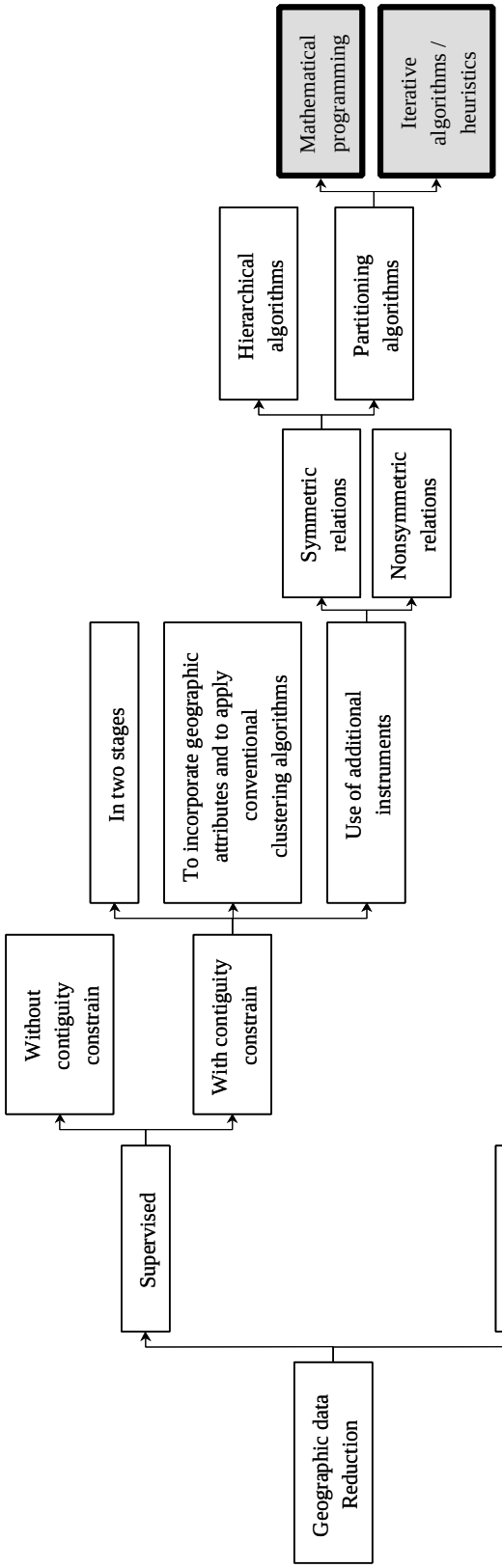
---

different problems regarding the analysis of spatial data: the problem of scale, related with the desired number of regions, and the problem of aggregation, related with the configuration of small areas inside bigger areas. For more information, see Openshaw (1977), Openshaw and Taylor (1981), and in an econometric context, see Fotheringham and Wong (1991) and Amrhein and Flowerdew (1992).

[13] The *Simulated Annealing* was proposed as an optimisation procedure by Kirkpatrick *et al.* (1983) and first time applied in the *Redistricting Problem* by Browdy (1990).

[14] For more information about the *Tabu Search Algorithm*, see Glover (1977, 1989, 1990).

Table 2.1. Summary of the different available methodologies for the reduction of geographical data.



Source: Own elaboration.

3. A LINEAR OPTIMISATION MODEL FOR THE DESIGN OF HOMOGENEOUS TERRITORIAL UNITS

In this section, the regionalisation problem is formulated as a linear optimisation model that allows the design of regions taking into account not only the characteristics of the areas but also their relationships. The possibility of treating the regionalisation problem as a linear model implies that, by its mathematical properties, the feasible region is convex and, as a result, it is possible to find the optimal solution. Another advantages of this kind of formulation are that it is easy to implement in a great variety of commercial software without paying a high price for it, and flexibility when some changes or additional constraints are needed.

Before introducing the mathematical formalisation of the model, its main characteristics and assumptions will be mentioned.

3.1. Model description.

3.1.1. Representation of the geographical set.

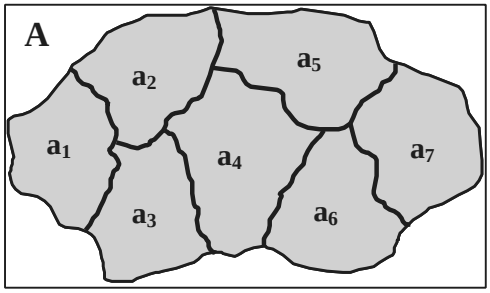
The starting point of any regionalisation process is the identification of the territory to regionalise. As an example, Figure 3.1 shows a territory that could be regionalised. It is composed by a finite number (*n*) of geographical areas of smaller size that form a geographical contiguous $\mathbf{A} = \{a_1, a_2, a_3, \ldots, a_n\}$.

Once the territory of interest has been defined, the next step consists of simplifying the previously defined geographical set in a way that each of the considered elements (*n* areas) and their neighbourhood relationships could be easily represented. This simplification can be done using a graph formed by *n* nodes, each of them representing one of the considered areas, and arcs that represent the geographical contiguity among them.

There are different methods in order to make this kind of simplification. We have selected the most general one, the *Delaunay Triangulation* (DT) (Aurenhammer, 1991). With this method, each arc relates those areas with a common border. One of the main advantages of this method is that the location of the point representing each of the areas does not affect the result of the graph. Other methods, such as the *Gabriel Graph* (Matula and Sokal, 1980), the *Relative Neighbourhood Graph* (Toussaint, 1980) or the

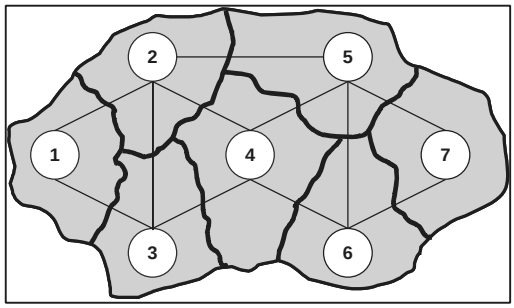
Minimum Spanning Tree (Graham and Hell, 1985) are particular cases of DT and results can be different depending on the location of the areal centroids. Figure 3.2 shows the DT graph of the territory considered in the example.

Figure 3.1. Group of areas that form the territory to regionalise



Source: Own elaboration.

Figure 3.2. Delaunay Triangulation (DT)



Source: Own elaboration

3.1.2. Relationships between the elements to be grouped.

The next step consists of the consideration of the relationships between the different areas (or nodes of the graph). The consideration of these relationships is one of the more relevant elements in the regionalisation process proposed in this section, as its consideration allows to take into account the interactions between. For example, if the objective of the study is to build regions with a similar population in order to establish proper comparisons, it will be helpful to consider also information on dissimilarities regarding other socio-economic variables in order to obtain more homogenous regions.

These relationships are incorporated in the model through a squared and symmetric matrix D_{ij} (i = 1, 2, ..., n and j = 1, 2, ... , n) where d_{ij} contains a dissimilarity measure between every couple of areas i, j.

The selected function to calculate dissimilarities between couples of areas should satisfy the following properties:

$$d_{ij} = d_{ji} \qquad \forall i, \forall j = 1,...,n \tag{3·1}$$

$$d_{ij} \geq 0, \ \left(d_{ij} = 0 \text{ if } i = j\right) \qquad \forall i, \forall j = 1,...,n \tag{3·2}$$

These properties imply that the function should not be metric (it does not have to satisfy the triangular inequality[15]):

$$d_{ij} \leq d_{ik} + d_{kj} \quad \forall i, \forall j, \forall k = 1,...,n \qquad (3\cdot3)$$

The possibility of using distance functions that should not be necessarily metric can be understood as a relaxation of the hypothesis used in the regionalisation models based on centroids where the rest of areas are assigned to each region depending on their proximity. When metric distance functions are used, the centroid-based approach ensures that the final solution will satisfy the geographical continuity constrain.

3.1.3. Strategy for the configuration of regions.

Once we have information about the territorial configuration and the relationships between the different areas, the next step consists of grouping the *n* areas $\{a_1, a_2, ... , a_n\}$ into *m* non-empty sets or regions {1, 2, ... , *m*} in a way that the areas belonging to each region form a geographical contiguity.

To define these regions it will be necessary to select *n-m* arcs from the global set of arcs that define the contiguity graph. These *n-m* arcs can be understood as a necessary but not sufficient condition to form *m* regions in a way that areas belonging to each region are totally interconnected but disconnected from the areas belonging to other regions. This selection should take into account the following conditions: each region must have a number of arcs equal to the number of areas belonging to the region less one, each region should be formed by a minimum of two areas and, last, in each region, every couple of areas should be connected by a one and only one combination of arcs[16]. This kind of regional configurations implies that the minimum number of areas in each region will be two (one arc connecting two areas), this is $m = [n/2]$. This condition is less restrictive as the number of areas forming the territory increases[17]. Figure 3.3 shows a possible solution to design 2 regions from 7 areas.

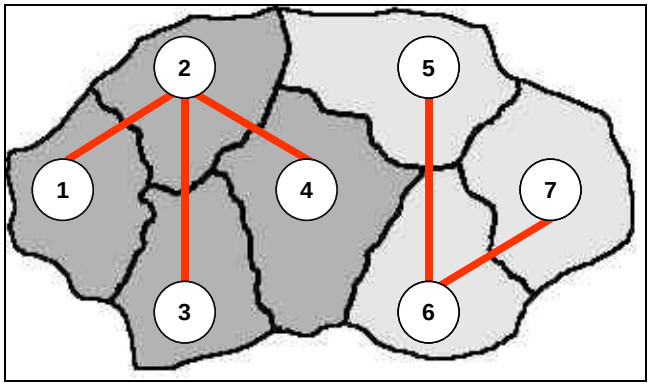
---

[15] For more information, see Gower and Legendre (1986).

[16] For more information about the properties of this (and other) configurations, see Ahuja *et al.* (1993).

[17] If we have one area that is considered as an outlier it should be treated as a region, the solution will be to exclude from the analysis and forming m-1 groups with the other n-1 areas.

Figure 3.3. Feasible result for the design of two regions.



Source: Own elaboration.

The location of arcs in each region does not have influence on the final result. For example, the region formed by the areas connected by arcs 1-2, 2-3 and 2-4 can be also configured with arcs 1-3, 2-4 and 3-4. This equivalence is related with the fact that the arcs function is only to ensure geographical contiguity, because of they do not have any value assigned. This strategy can be very useful to identify regional configurations with a high variety of shapes (longed or compact regions), as it does not rely on centroids, which tend to produce compact areas.

3.1.4. Considered criteria for the configuration of regions: the objective function.

The objective of grouping *n* areas in *m* regions is that the areas belonging to each region form a homogeneous geographical contiguity. So, a partition criterion considering which one of the possible configurations of *n* areas in *m* regions is the most adequate should be defined.

With this aim, it is necessary to define a measure of adequacy of a regional configuration. One possibility consists of calculating the degree of heterogeneity of the areas assigned to a region or, other alternative could be to calculate the degree of isolation of the areas of one region related to the rest. The heterogeneity measure selected in this paper consists of the sum of the elements of the upper triangular matrix of dissimilarity relationships between the areas in the considered region. Following Gordon (1999), the heterogeneity measure for region *r*, C_r can be calculated as follows:

$$H(C_r) \equiv \sum_{\{i,\, j \in C_r \mid i<j\}} d_{ij} \qquad (3\cdot4)$$

Taking this into account, the problem of obtaining *r* homogeneous classes (regions) can be formulated as the minimisation of the sum of the heterogeneity measures of each class (region) *r*:

$$P(H,\Sigma) \equiv \sum_{r=1}^{c} H(C_r) \qquad (3\cdot5)$$

or, following the MIN-MAX strategy, we can also try to minimise the value of the most heterogeneous region as this imply that the rest of the regions would be equal or less heterogeneous:

$$P(H,Max) \equiv \max_{\{r=1,\ldots,c\}} H(C_r) \qquad (3\cdot6)$$

One disadvantage associated to the second strategy is that once the value of the most heterogeneous region is minimised, the configuration of the rest of the regions will not be revised, avoiding the possibility of making changes that could improve their heterogeneity. For this reason, the selected strategy has been the minimisation of the sum of the heterogeneity measures of each region $(P(H,\Sigma))$.

It is worth mentioning that both objectives, minimising internal heterogeneity $H(C_r)$ and maximising the isolation among regions $I(C_r)$, are not independent. In fact, we can formulate an equivalent objective in terms of isolation criteria:

$$P(H,\Sigma) \equiv P(I,\Sigma) \equiv \sum_{r=1}^{c} I(C_r) \text{ with } I(C_r) \equiv \sum_{i \in C_r} \sum_{j \notin C_r} d_{ij} \qquad (3\cdot7)$$

3.2. Mathematical model.

Parameters:

i, I index and set of areas, $i = \{1,...,n\}$;

k, K index and set of regions, $k = \{1,...,m\}$;

c_{ij} $\begin{cases} 1, \text{if } i \text{ and } j \text{ are continuous (share a border), with } i < j, \\ 0, \text{otherwise;} \end{cases}$

M $Max\left(\sum_{j=1}^{n} c_{1j},...,\sum_{j=1}^{n} c_{nj} \right)$;

N_i $\{j | c_{i,j} = 1\}$;

$D_{i,j}$ Dissimilarity relationships between areas *i* and *j*, with *i*<*j*;

Decision Variables:

X_{ijk} $\begin{cases} 1, \text{if areas } i \text{ and } j | j \in N_i \text{ belong to the same region } k, \text{ with } i < j, \\ 0, \text{otherwise;} \end{cases}$

Y_{ik} $\begin{cases} 1, \text{if area } i \text{ belongs to region } k, \\ 0, \text{otherwise;} \end{cases}$

T_{ij} $\begin{cases} 1, \text{the disimilarity relationship between } i \text{ and } j \text{ is considered if both areas} \\ \quad \text{belong to the same region } k, \ i < j, \\ 0, \text{otherwise;} \end{cases}$

Objective function: $Min \sum_{i=1}^{n} \sum_{j=1}^{n} D_{ij} \cdot T_{ij}$

Subject to:

$$T_{ij} \geq Y_{ik} + Y_{jk} - 1, \qquad \forall i, \forall j = 1,...,n \ ; \ \forall k = 1,...,m \qquad (3\cdot8)$$

$$\sum_{i=1}^{n} Y_{ik} \geq 2, \qquad \forall k = 1,...,m \qquad (3\cdot9)$$

$$\sum_{k=1}^{m} Y_{ik} = 1, \qquad \forall i = 1,...,n \qquad (3\cdot10)$$

$$\sum_{j \in N_i} X_{ijk} \leq Y_{ik} \cdot M, \qquad \forall i = 1,...,n \ ; \ \forall k = 1,...,m \qquad (3\cdot11)$$

$$\sum_{j \in N_i} X_{jik} \leq Y_{ik} \cdot M, \qquad \forall i = 1,...,n \ ; \ \forall k = 1,...,m \qquad (3\cdot12)$$

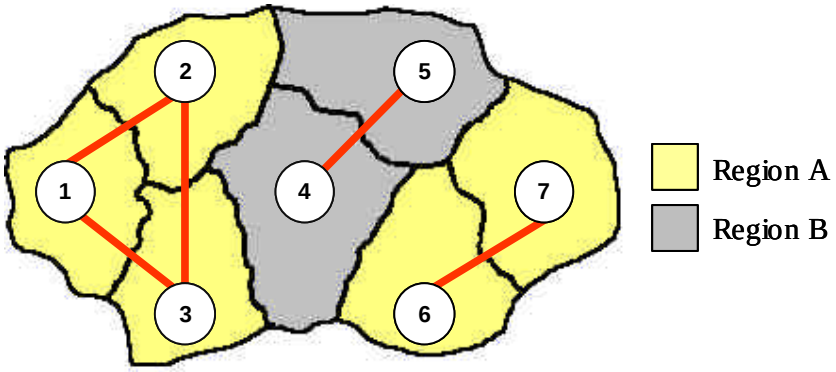
$$\sum_{i=1}^{n}\sum_{j\in N_i} X_{ijk} = \sum_{i=1}^{n} Y_{ik} - 1, \qquad \forall k = 1,...,m \tag{3·13}$$

$$\sum_{i,j\in C} X_{ijk} \leq |C| - 1, \ \forall \text{ non - empty subset of } C \subseteq \{3,...,(n\text{-}2m+1)\}; \quad \forall k = 1,...,m \tag{3·14}$$

$$X_{ijk} \in \{1,0\}\ ;\ Y_{ik} \in \{1,0\};\ T_{ij} \geq 0, \ \ \forall i, \forall j = 1,...,n\ ;\ \forall k = 1,...,m \tag{3·15}$$

As it was previously mentioned, the objective function looks for the minimisation of the total heterogeneity, measured as the sum of the elements of the upper triangular matrix (D_{ij}) of dissimilarity relationships between areas belonging to the same region (the elements defined by the binary matrix T_{ij}). Restriction (3·8) controls the assignation of the values of matrix T_{ij} where, by the nature of the objective function, the relationship between areas *i* and *j* will only be taken into account if they belong to the same region. Restriction (3·9) imposes that the minimum number of areas defining a region is two. As it was previously mentioned, the restriction is less strong as the number of areas increases. Restriction (3·10) imposes that each area must be assigned to one and only one region. Restrictions (3·11) and (3·12) imposes that only when the area *i* is assigned to region *k*, it will be possible to establish arcs to the neighbourhoods of the area ($j \in N_i$). To avoid an excessive reduction of feasible regional configurations, the number of arcs from an area can be greater than one. Restriction (3·13) imposes that the number of arcs to ensure geographical contiguity of the areas assigned to one region must be equal to the number of areas in the region less one. However, this restriction does not totally ensure that the final solution will be formed by contiguous regions. There are cases such as the one shown in Figure 3.4, where region *A*, formed by areas 1, 2, 3, 6 and 7, satisfies restriction (3·13) –there are four connecting arcs for five areas– but the combination of arcs 1-2, 1-3, 2-3 generates a cycle that breaks the geographical contiguity of the region. For this reason, it will be necessary to control, a part of the number of arcs, if there are cycles and this is the origin of restriction (3·14).

Figure 3.4. Non-feasible regional configuration.



Source: Own elaboration.

The problem of cycles has been treated in the literature as the analysis of *subtour* in transport models such as the *Vehicle Routing Problem* (VRP)[18]. The VRP consists of defining vehicles routes with a given origin and end in the same node (called *depot*) and trying to minimize costs. The design of a tour for a certain vehicle cannot contain subtours and to control this condition, the VRP incorporates the following constraint:

$$\sum_{j,i\in S} X_{ijk} \leq |S| - 1, \ \forall \text{ non-empty subset of } S \subseteq \{2,...,n\};\ k=1,...,m. \qquad (3\cdot16)$$

The main disadvantage of this approach is that the number of restrictions increases exponentially with n and m. For this reason, and although the proposal is theoretically adequate, at the practical level it has been necessary to implement other restrictions to solve this problem in a more efficient way. These alternatives can be appropriated for the specific problem of the VRP (although they do not ensure the elimination of subtours in problems of a certain dimension), but not for the regionalisation problem. For example, it is required to establish *a priori* a *depot* node that will be the origin and end of all the tours, and it is also necessary to establish a sequential order among nodes.

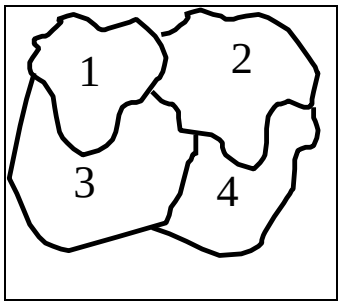
However, the theoretical restriction of the VRP can be adapted in an efficient way in this geographical context as we know the number of elements of the set S. For

[18] This problem was first proposed by Dantzing and Ramser (1959). A survey about the models derived from this approach can be found in Laport and Osman (1995).

example, in the territorial configuration of Figure 3.5 we can clearly identify the different combination of arcs c_{ij} that can generate cycles. The combination of arcs 1-2, 1-3, 2-3 (or 2-3, 2-4, 3-4) will produce a cycle where 3 areas would be involved, 1, 2 and 3 (or 2, 3, 4), while the combination of arcs 1-2, 1-3, 3-4, 2-4 will generate a cycle among the four areas.

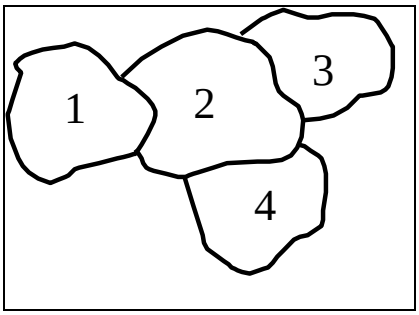
Moreover, in a territorial configuration as the one shown in Figure 3.6, there is no combination of arcs c_{ij} that could generate a cycle. For this reason, at the territorial level, not every subset *S* can have cycles as the number of potential arcs c_{ij} is limited to those combinations *i,j* where the value of the contact matrix *wij* =1. This is the set of potential arcs c_{ij} that are included in $\boldsymbol{N_i}$.

Figure 3.5. Configuration of areas with potential cycles



Source: Own elaboration.

Figure 3.6. Configuration of areas without potential cycles



Source: Own elaboration.

But, is there any special pattern that could help to detect potential cycles in a specific territorial configuration? Yes, we only have to identify those combinations of arcs where the number or arcs is equal to the number of areas connected through them. For example, in the case shown in Figure 3.5, the three arcs 1-2, 1-3, 2-3 (or 2-3, 2-4, 3-4) connect three areas, 1,2,3 (or 2,3,4), and as a result, 3 arcs and 3 areas imply the existence of a cycle. The same happens with the combination of arcs 1-2, 1-3, 3-4, 2-4 that connect four areas (1,2,3,4). Again, 4 arcs and 4 areas imply the existence of a cycle of 4 elements.

But, for a territorial configuration of *n* areas that will be grouped in *m* regions, which is the maximum number of areas that can be involved in a cycle? As the model, in restriction (3·9), requires that the minimum number of areas in a region is 2, in the case where (*m*-1) regions are formed by two areas, there will be no possibility of cycles, as each region will only have one possible arc (restriction 3·13). For this reason, when

creating m-1 regions with 2 areas, we will have a region formed by n-2(m-1) areas with (n-2(m-1))-1 arc, which is the maximum number of arcs that can create a cycle. Simplifying this expression, we have that:

$$n\text{-}2m\text{+}1 \qquad (3\cdot17)$$

So, the minimum number of areas where the possibility of finding a cycle should be evaluated is three, as it is impossible that for a lower number of areas we find this problem.

As a result, restriction (3·14) is related with the modification of the set S as proposed in the VRP. Using this modification, we achieve an important reduction in the number of restrictions to satisfy, avoiding that the number of restrictions increases exponentially with n and m. This fact allows to use commercial software in the context of regionalisation problems with a high number of areas and regions.

Last, restriction (3·15) only implies that X_{ijk} and Y_{ik} should be binary variables. Although the variable T_{ij} has been defined as positive, and not as binary, it will always take values 0 or 1 because of the combination of restriction (3·8) with the objective of minimisation of the model[19].

3.3. Application of the model.

In this subsection, different examples are shown with the aim of illustrating the model capacity to design regional configurations with different characteristics. Thus, it has been implemented a first set of four examples each one with a different dissimilarity matrixes ($D_{i,j}$), where values $d_{i,j}$ have been established in such a way that it is possible to know *a priori* the optimum regional configurations. The procedure to obtain the dissimilarity matrix in each example has been the following:

[19] The possibility of defining a variable taking values 0 or 1 as positive and not as a binary variable has an advantage when using the branch and bound algorithm, as the number of sub-problems is drastically reduced. For more information about this algorithm, see Hiriart *et al.* (1983).

1. n areas have been grouped in m contiguous regions, assigning each area $i = \{1,...,n\}$ to a region $k = \{1,...,m\}$. This aggregation permits to built the set R_k $\{i|i \hat{I} k\}$.
2. A value has been assigned to each of the areas $i = \{1,...,n\}$ depending on the region they have been assigned. This value is given by the sum of a constant with a random term, generated from a uniform distribution among 0 and 1. The value of the constant is different for each region, as there should be a big enough difference (D) in order to obtain significant different average values for each region. The applied expression has been the following:

$$A_{i\in R_k} = C + (D * k) + e \quad \forall i = 1,...,n; \forall k = 1,...,m; e \sim U[0,1] \quad (3\cdot18)$$

3. Next, the relationships between areas has been calculated using a distance function. The weighted Euclidean distance has been applied in order to calculate distances among the elements of the A_i vector after centering it.

$$d_{ij} = \sqrt{\left(\frac{A_i^c}{S} - \frac{A_j^c}{S}\right)^2}, \quad \forall i, j = 1,...,n \mid i < j \qquad (3\cdot19)$$

where S is the standard deviation of the A_i, vector and A_i^c is a centered vector calculated as follows from A_i:

$$A_i^c = A_i - \left(\sum_{i=1}^{n} A_i \Big/ n\right), \ \forall i = 1,...,n \qquad (3\cdot20)$$

The matrixes obtained with this procedure are shown in Table 3.1 and the obtained regional configurations after applying the optimisation model with the different relationship matrices are shown in the maps in Table 3.2. The solutions coincide with the optimal regional configurations predefined above and, so, it seems that the model can design regions with a high variety of shapes.

Table 3.1. Relationships matrixes for examples 1 to 4.

Example 1

area	2	3	4	5	6	7	8	9	10	11
1	1.04	1.21	1.18	1.11	0.17	0.14	2.26	2.31	0.09	2.31
2		0.17	0.14	0.07	1.22	1.18	1.22	1.27	1.14	1.27
3			0.03	0.10	1.38	1.35	1.05	1.10	1.31	1.10
4				0.07	1.35	1.32	1.08	1.13	1.27	1.13
5					1.29	1.25	1.15	1.20	1.21	1.20
6						0.03	2.43	2.48	0.08	2.49
7							2.40	2.45	0.05	2.45
8								0.05	2.36	0.05
9									2.41	0.00
10										2.41

Example 2

area	2	3	4	5	6	7	8	9	10	11
1	0.06	0.02	0.03	2.42	2.49	1.23	0.03	1.19	0.04	0.02
2		0.07	0.03	2.37	2.44	1.18	0.09	1.13	0.01	0.04
3			0.04	2.44	2.51	1.25	0.01	1.20	0.06	0.03
4				2.40	2.47	1.21	0.06	1.16	0.02	0.01
5					0.07	1.19	2.45	1.23	2.38	2.40
6						1.26	2.52	1.31	2.45	2.48
7							1.27	0.05	1.19	1.22
8								1.22	0.07	0.05
9									1.14	1.17
10										0.02

Example 3

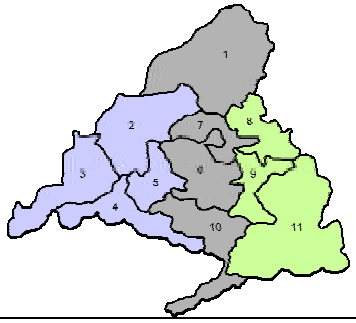
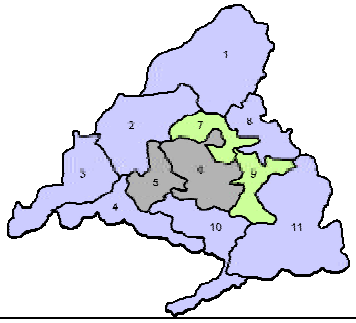
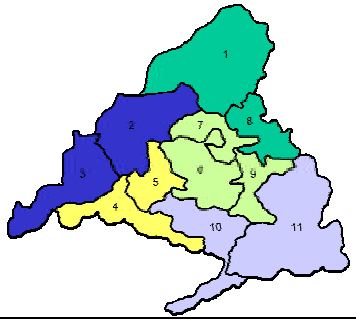
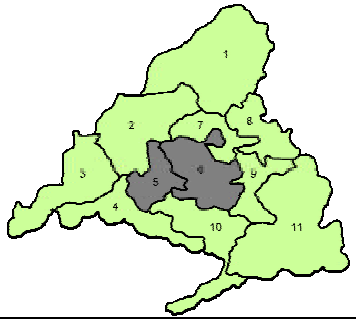
area	2	3	4	5	6	7	8	9	10	11
1	0.64	0.80	1.36	1.27	2.03	1.98	0.08	1.98	2.78	2.79
2		0.15	0.72	0.62	1.39	1.34	0.73	1.34	2.13	2.14
3			0.57	0.47	1.23	1.19	0.88	1.18	1.98	1.99
4				0.10	0.67	0.62	1.45	0.62	1.41	1.42
5					0.76	0.72	1.35	0.71	1.51	1.52
6						0.05	2.11	0.05	0.75	0.76
7							2.07	0.00	0.79	0.80
8								2.06	2.86	2.87
9									0.79	0.80
10										0.01

Example 4

area	2	3	4	5	6	7	8	9	10	11
1	0.23	0.27	0.16	2.45	2.56	0.22	0.04	0.06	0.17	0.04
2		0.05	0.06	2.23	2.34	0.00	0.27	0.28	0.40	0.27
3			0.11	2.18	2.29	0.05	0.31	0.33	0.45	0.31
4				2.29	2.40	0.06	0.21	0.22	0.34	0.21
5					0.11	2.23	2.49	2.51	2.63	2.50
6						2.34	2.61	2.62	2.74	2.61
7							0.26	0.28	0.40	0.26
8								0.02	0.13	0.00
9									0.11	0.02
10										0.13

Source: Own elaboration.

Table 3.2. Solutions for the relationships matrixes from Table 3.1.

Example 1 	***Example 2*** 
n=11 and m=3	***n=11 and m=3***
Example 3 	***Example 4*** 
n=11 and m=5	***n=11 and m=2***

n: number of areas, m: number of regions.
Source: Own elaboration.

3.4. Computational results.

One of the most interesting features of optimisation models when applied in real problems is the required computational time to achieve the optimal solution.

With the aim of testing the computational capacity of the model, it was applied to different random territorial configurations. The procedure to obtain these random configurations was the following:

a. For a given number *n* of areas, a triangular matrix was randomly generated following a [0,1] uniform distribution.
b. A threshold point, between 0 and 1, was fixed in a way that random numbers above this point were replaced by 1, and 0 otherwise. The obtained binary matrix can be interpreted as a contact matrix, which should be evaluated in terms of contiguity. The threshold value was assigned taking into account that the resulting territorial configuration (or connecting arcs) was realistic in term of the neighbourhoods of each area. The selected matrices have an average density of 28.3% and a median of neighbourhoods of 3 per area, ranging from 1 to 8.
c. Every randomly generated matrix was evaluated in terms of geographical contiguity and the feasible ones have been selected[20].
d. Last, the relationships between the *n* considered areas were randomly generated from a [0,1] uniform distribution. Using this method, it is assuming a scenario where relationships between areas are not geographically dependent.

Table 3.3 shows the average running times[21] for different combinations of areas and regions (5 examples for each combination).

[20] Although the decision of evaluating *a posteriori* the contiguity of the matrix would imply a higher computation time for the generation of the different examples, this methodology assures that the territorial configurations in each example are totally random.

[21] The calculations in this paper have been performed using Extended LINGO/PC 6.0 in a PC computer with a Pentium 4 processor at 2.40C GHz and 256 Mb of RAM memory.

Table 3.3. Average running time, in seconds, for different combinations (areas-regions).

		Regions		
		2	**4**	**6**
Areas	**5**	<1*	-	-
	8	<1*	3.00	-
	11	<1*	19.00	-
	14	5.80	117.40	2,571.00
	17	2.20	2,458.20	42,283.80

Note: Five examples for each combination of areas and regions.
* Execution times lower than a second.
Source: Own elaboration.

Although the number of restrictions was clearly reduced with the modification of restriction (3·14), that controls the elimination of cycles, the running time stills very high. In fact, for those cases with more than 17 areas the running time increases substantially. For this reason, other alternative that would permit to increase the computational capacity of the model will be considered in the next section.

4. A SOLUTION FOR THE "COMPUTATIONAL PROBLEM": THE RASS ALGORITHM

In this section, a new algorithm called *RASS* (*Regionalisation Algorithm with Selective Search)*, is proposed. The most relevant characteristic of this algorithm is related with the fact that the way it operates is inspired in the own characteristics of regionalisation processes, where available information about the relationships between areas can play a crucial role in directing the searching process in a more selective and efficient way (less random).

The *RASS* incorporates inside its algorithm the optimisation model presented in section 3 in order to achieve local improvements in the objective function. These improvements can generate significant changes in regional configurations, changes that would be very difficult to obtain using other iterative methods.

4.1. Steps for the application of *RASS.*

Step 1: Take as a starting point, a feasible solution of *m* regions that group *n* areas.

Step 2: Select from these *m* regions the more heterogeneous geographical contiguity formed by *r* regions with $2 \le r \le (m-1)$.

$$H(C_m) \equiv \sum_{\{i,\, j \in C_m | i<j\}} d_{ij} \rightarrow Max\left(\sum_{m \in M_i} H(C_m) \right) \qquad (4\cdot1)$$

where M_i is the set formed by the different alternatives of selection of *r* contiguous regions of the available *m* regions.

Step 3: Application of the direct optimisation model to the areas of the *r* selected regions to create r^* regions.

Step 4: Select a region to include (*e*): From the (*m*-*r*) regions that were not considered, identify those areas bordering on territory formed by the *r** regions and select the one with higher similarities with any of the regions in *r*.

$$I(C_{d,f}) \equiv prom\left(\sum_{i \in C_f} \sum_{j \in C_d | j>i} d_{ij} \right) \rightarrow Min\left(I(C_{d,f}) \right) \qquad (4\cdot2)$$

where *d* is the set of the r^* regions which are inside, and *f* is a subset of regions bordering on *d*. Each of the (*m*-*r*) regions that were not selected in the step 2 will only be selected once in every cycle (steps 2 to 8).

Step 5: Select the region that will be removed (*s*): The region with higher differences with the region to be included (*e*) in step 4 will be removed from *d*. The region to be removed cannot destroy the internal contiguity of *d*.

$$I(C_{d,e}) \equiv prom\left(\sum_{i \in C_e} \sum_{j \in C_d | j>i} d_{ij} \right) \rightarrow Max\left(I(C_{d,e}) \right) \qquad (4\cdot3)$$

Step 6: Include in the set of *r* regions the region (*e*) and remove (*s*): *d*=(*d*+*e*-*s*). The direct optimisation model will be applied to the new configuration of *r* regions to create r^* regions.

Step 7: Repeat steps 4 to 6 until the (*m*-*r*) regions that where not selected in step 2 have been included at any time in *d*, or until there are no more candidates to be selected in the bordering on *d*.

Step 8: Calculate the value of the objective function.

Step 9: If the value of the objective function improves, step 2 would be repeated. If the value of the objective function does not improve, step 2 would be repeated but selecting the next more heterogeneous group. Steps 2 to 8 would be repeated until no significant improvement in the objective function is found in a given number of cycles (*C*) or until the list of alternative *r* contiguous regions is exhausted.

Some characteristics to highlight from the *RASS* algorithm are the following:

a. The application of direct optimisation to a group of regions, in steps 3 to 5, permits to achieve improvements in the objective function that can be accompanied by important changes in regional configurations because of the reassignation of an important number of areas.
b. The criteria used in step 2 for the selection of *r* regions and the criteria for including/removing regions in steps 4 and 5 try to keep in the optimisation model, step 3, those regions with a higher potential to improve the objective function after reconfiguration. The objective is to ensure that the included region is the one that presents the higher probability of containing areas belonging to other regions. This potential reassignation is identified assuming that two regions with exchanged areas, decreases the dissimilarities among these regions. Last, when the region to be included (*e*) is selected, the next step establishes that the region to be removed (*s*) (in order to keep an appropriated number of areas for the optimisation model) is the more different one from the region to include. This region has lower possibilities of exchanging areas with the region to be included (*e*).
c. The conditions in steps 7 and 9 try to avoid repetitive searching patterns. Moreover, the criteria for including/removing regions and the use of the

optimisation model clearly improve the capacity of *RASS* of escaping from local optimum.

d. The fact of applying the optimisation model only to a part of the considered territory does not imply that each local improvement could worsen the global solution. In fact, after each cycle, the value of the objective function will be always lower or equal to the value of the objective function at the beginning of the cycle.

4.2. Computational results and comparison with the direct optimisation.

This section tries to evaluate the performance of the *RASS* algorithm respect the direct optimisation model proposed in section 3. The solved examples are the ones that were randomly generated in section 3.6[22]. In order to apply the algorithm to these examples, it was necessary to define an initial feasible partition that could be used as a starting point for *RASS*. The initial partition was randomly generated following these steps:

a. Generate a vector with *n* values (as many as areas) using a uniform distribution between 0 and 1.

b. The interval [0,1] is divided in equal sized intervals, as many as the number of regions to design. For example: for 2 regions we used the intervals [0, 0.5) and [0.5, 1] and for 4 regions, the intervals were [0, 0.25), [0.25, 0.5), [0.5, 0.75) and [0.75, 1). Each of these intervals represents a region, in such a way that the elements of the random vectors can be transformed in a vector that assignates areas to regions (potential initial partition).

c. If the initial partition is feasible in terms of geographical contiguity, this partition is used as starting point for *RASS*.

Some descriptive of the results for the 30 considered problems (5 for each combination of regions and areas) are shown in Table 4.1. *RASS* achieved the optimal

[22] In this analysis we have excluded the examples where 2 regions should be formed, as in this case the application of the *RASS* would be equivalent to the direct application of the optimisation model: there is no difference between the values of parameters *m* and *r* of *RASS* and, as a result, the application of step 3 will take directly to the optimal solution.

solution in the 100% of the considered examples in a considerably lower time than the direct solution method.

Table 4.1. Comparison of *RASS* with the direct solution method.

Regions	Areas	Optimum/5	Seconds (*RASS*)	Seconds (Direct)	(FOI - FO1c) / (FOI - SO*)
4	8	5/5	3.40	3.00	76.45%
	11	5/5	5.80	19.00	86.70%
	14	5/5	29.00	117.40	74.31%
	17	5/5	247.20	2,458.20	69.46%
6	14	5/5	25.20	25,710.00	85.93%
	17	5/5	250.00	42,283.80	66.71%

FOI= Initial objective function, FO1c= Objective function after the first cycle, SO= Optimal solution.*
Source: Own elaboration.

In the last column, it can be seen that after the first cycle of the *RASS*, the value of the objective function is reduced in an 80% of the total reduction required to achieve the global optimum.

Using the available information about running times of both regionalisation methods, the direct method and the *RASS*, it is possible to calculate the time savings by applying the algorithm. Figure 4.1 shows the relationship between the savings and an indicator of complexity that has been defined as the product between the number of considered areas and the number of considered regions. The results in this figure show that in less complex models the direct method is a better option, while in complex models the *RASS* provides better results. According to these results, this change happens for models with a complexity over 57.83 (58 if we keep the discrete nature of the variable[23]).

In order to obtain a better measure of the time savings achieved with *RASS*, we have estimated a quadratic model between time savings and the measure of complexity[24,25]. The results of estimating this model are shown in Table 4.2. There is a

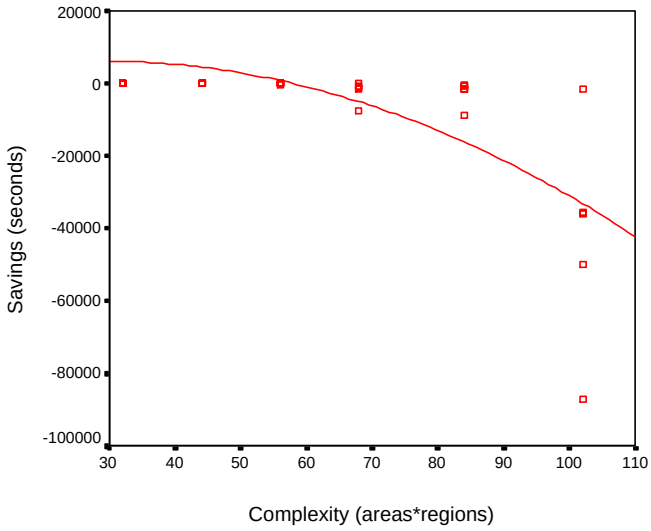
[23] It should be highlighted that this value can be obtained with different combinations of areas and regions.

[24] We have considered together the effects of the number of areas and regions because when introduced separately in the regression, there is a problem of collinearity due to the high correlation among them.

[25] We have excluded the intercept from this regression in order to impose that the execution time is equal to zero when the complexity is equal to zero.

significant relationship between the two variables at 1% significance level. In front of a marginal increase in the complexity of the problem, the use of *RASS* implies a time saving of 426.08-14.73 (*areas*regions*), a result that confirms the previously mentioned intuition.

Figure 4.1. Relationship between the complexity of the problem and the time savings obtained after applying *RASS*.



Source: Own elaboration.

Table 4.2. Quadratic regression among the time savings obtained with *RASS* and the complexity indicator.

n=30	Coefficient
(areas×regions)	426.078*
$(areas \times regions)^2$	-7.367*
R^2	*0.566*
F	*18.269**

** Significant at 1%*
Source: Own elaboration.

4.3. Capacity of the RASS to achieve global optimums in more complex problems.

As in more complex problems, it is impossible to compare the results obtained by the *RASS* and direct optimisation because the execution method for the second would be very high, in this section the obtained solution for a regionalisation process where 38 areas are grouped in 10 regions (complexity of 38*10 = 380) is presented. For this comparison, it was applied the same procedure than in the examples of section 3.4: A

relationship matrix D_{ij} is defined in a way that it is possible to know *a priori* the optimal solution of the regionalisation process. This optimal solution can be compared with the solution obtained by the *RASS*.

4.3.1. Data

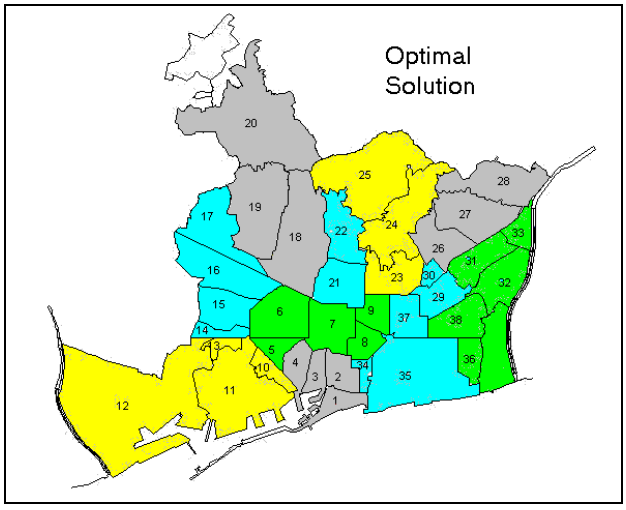
4.3.1.1. Characteristics of the territory to regionalise.

The selected areas for this example are the 38 areas (*Zones Estadístiques Grans*) that form the city of Barcelona. The first step consists of considering the contiguity relationships among these 38 areas or, in other words, in obtaining the contact matrix.

4.3.1.2. Relationships among areas.

The relationships among areas (see Table 4.3) were created in a way that the optimal solution grouped the 38 areas in 10 regions, each of them with different shapes and sizes (among 2 and 6 areas by region). This optimal solution is shown in Figure 4.2, and this is the solution that the *RASS* algorithm should be able to identify.

Figure 4.2. Preestablished optimal regional configuration.



Source: Own elaboration.

Table 4.3. Relationships matrix between the 38 areas.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37
2	0.006																																				
3	0.000	0.006																																			
4	0.038	0.032	0.038																																		
5	0.328	0.322	0.328	0.290																																	
6	0.330	0.324	0.330	0.291	0.002																																
7	0.319	0.313	0.319	0.281	0.009	0.011																															
8	0.359	0.353	0.359	0.321	0.031	0.029	0.040																														
9	0.340	0.335	0.340	0.302	0.013	0.011	0.022	0.018																													
10	0.676	0.670	0.676	0.638	0.348	0.346	0.357	0.317	0.335																												
11	0.661	0.655	0.661	0.622	0.333	0.331	0.342	0.302	0.320	0.015																											
12	0.671	0.666	0.671	0.633	0.344	0.342	0.353	0.312	0.331	0.004	0.011																										
13	0.682	0.677	0.682	0.644	0.354	0.353	0.363	0.323	0.342	0.006	0.022	0.011																									
14	0.951	0.945	0.951	0.913	0.623	0.621	0.632	0.592	0.610	0.275	0.290	0.279	0.269																								
15	0.977	0.971	0.977	0.939	0.649	0.647	0.658	0.618	0.636	0.301	0.316	0.305	0.295	0.026																							
16	0.972	0.966	0.972	0.934	0.644	0.642	0.653	0.613	0.631	0.296	0.311	0.300	0.290	0.021	0.005																						
17	0.964	0.958	0.964	0.925	0.636	0.634	0.645	0.605	0.623	0.288	0.303	0.292	0.281	0.013	0.013	0.008																					
18	1.270	1.264	1.270	1.231	0.942	0.940	0.951	0.911	0.929	0.594	0.609	0.598	0.587	0.319	0.293	0.298	0.306																				
19	1.316	1.310	1.316	1.278	0.988	0.986	0.997	0.957	0.976	0.640	0.656	0.645	0.634	0.365	0.339	0.344	0.353	0.046																			
20	1.287	1.282	1.287	1.249	0.960	0.958	0.969	0.929	0.947	0.612	0.627	0.616	0.605	0.337	0.311	0.316	0.324	0.018	0.029																		
21	1.603	1.597	1.603	1.565	1.275	1.273	1.284	1.244	1.263	0.927	0.943	0.932	0.921	0.652	0.626	0.631	0.640	0.333	0.287	0.316																	
22	1.625	1.620	1.625	1.587	1.298	1.296	1.307	1.266	1.285	0.950	0.965	0.954	0.943	0.675	0.649	0.654	0.662	0.356	0.309	0.338	0.022																
23	1.928	1.923	1.928	1.890	1.601	1.599	1.609	1.569	1.588	1.252	1.268	1.257	1.246	0.978	0.951	0.957	0.965	0.659	0.612	0.641	0.325	0.303															
24	1.898	1.893	1.898	1.860	1.571	1.569	1.580	1.539	1.558	1.223	1.238	1.227	1.216	0.948	0.922	0.927	0.935	0.629	0.582	0.611	0.295	0.273	0.030														
25	1.928	1.923	1.928	1.890	1.601	1.599	1.609	1.569	1.588	1.252	1.268	1.257	1.246	0.978	0.951	0.957	0.965	0.659	0.612	0.641	0.325	0.303	0.000	0.030													
26	2.243	2.237	2.243	2.205	1.915	1.913	1.924	1.884	1.902	1.567	1.582	1.571	1.560	1.292	1.266	1.271	1.279	0.973	0.927	0.955	0.640	0.617	0.314	0.344	0.314												
27	2.235	2.229	2.235	2.197	1.907	1.905	1.916	1.876	1.895	1.559	1.574	1.564	1.553	1.284	1.258	1.263	1.272	0.965	0.919	0.948	0.632	0.610	0.307	0.337	0.307	0.008											
28	2.257	2.252	2.257	2.219	1.930	1.928	1.939	1.898	1.917	1.581	1.597	1.586	1.575	1.307	1.280	1.286	1.294	0.988	0.941	0.970	0.654	0.632	0.329	0.359	0.329	0.015	0.022										
29	2.565	2.560	2.565	2.527	2.238	2.236	2.247	2.206	2.225	1.890	1.905	1.894	1.883	1.615	1.588	1.594	1.602	1.296	1.249	1.278	0.962	0.940	0.637	0.667	0.637	0.323	0.330	0.308									
30	2.558	2.552	2.558	2.520	2.230	2.228	2.239	2.199	2.217	1.882	1.897	1.886	1.876	1.607	1.581	1.586	1.594	1.288	1.242	1.270	0.955	0.932	0.629	0.659	0.629	0.315	0.323	0.300	0.008								
31	2.882	2.876	2.882	2.844	2.554	2.552	2.563	2.523	2.541	2.206	2.221	2.210	2.199	1.931	1.905	1.910	1.918	1.612	1.566	1.594	1.279	1.256	0.953	0.983	0.953	0.639	0.647	0.624	0.316	0.324							
32	2.866	2.861	2.866	2.828	2.539	2.537	2.548	2.507	2.526	2.191	2.206	2.195	2.184	1.916	1.890	1.895	1.903	1.597	1.550	1.579	1.263	1.241	0.938	0.968	0.938	0.624	0.631	0.609	0.301	0.309	0.015						
33	2.873	2.868	2.873	2.835	2.546	2.544	2.555	2.514	2.533	2.198	2.213	2.202	2.191	1.923	1.897	1.902	1.910	1.604	1.557	1.586	1.270	1.248	0.945	0.975	0.945	0.631	0.638	0.616	0.308	0.316	0.008	0.007					
34	2.574	2.568	2.574	2.535	2.246	2.244	2.255	2.215	2.233	1.898	1.913	1.902	1.891	1.623	1.597	1.602	1.610	1.304	1.258	1.286	0.971	0.948	0.645	0.675	0.645	0.331	0.339	0.316	0.008	0.016	0.308	0.293	0.300				
35	2.581	2.575	2.581	2.542	2.253	2.251	2.262	2.222	2.240	1.905	1.920	1.909	1.898	1.630	1.604	1.609	1.617	1.311	1.265	1.293	0.978	0.955	0.652	0.682	0.652	0.338	0.346	0.323	0.015	0.023	0.301	0.286	0.293	0.007			
36	2.901	2.895	2.901	2.863	2.573	2.571	2.582	2.542	2.560	2.225	2.240	2.229	2.219	1.950	1.924	1.929	1.937	1.631	1.585	1.613	1.298	1.275	0.972	1.002	0.972	0.658	0.666	0.643	0.335	0.343	0.019	0.034	0.027	0.327	0.320		
37	2.586	2.580	2.586	2.548	2.258	2.256	2.267	2.227	2.245	1.910	1.925	1.914	1.904	1.635	1.609	1.614	1.622	1.316	1.270	1.298	0.983	0.960	0.657	0.687	0.657	0.343	0.351	0.328	0.020	0.028	0.296	0.281	0.288	0.012	0.005	0.315	
38	2.853	2.847	2.853	2.815	2.525	2.523	2.534	2.494	2.513	2.177	2.192	2.182	2.171	1.902	1.876	1.881	1.890	1.583	1.537	1.566	1.250	1.228	0.925	0.955	0.925	0.610	0.618	0.596	0.288	0.295	0.029	0.013	0.020	0.279	0.272	0.048	0.267

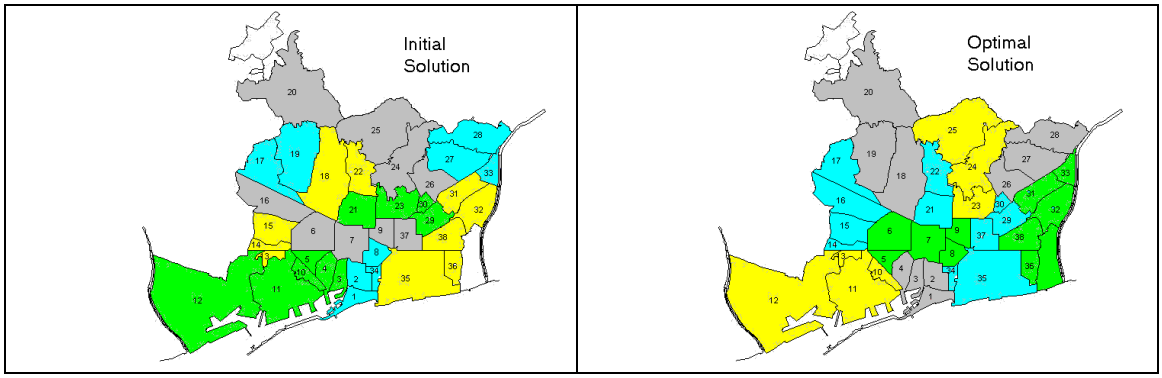
Source: Own elaboration.

4.3.2. Evaluation of results.

The initial partition is shown in Table 4.4. This is the partition that is considered by the *RASS* in the step 1. It is worth mentioning that this configuration is very different to the optimal one. After 5 cycles, the *RASS* algorithm properly reaches the optimal solution.

The different regional configurations considered by the *RASS* in the different steps and iterations are shown in the Annex.

Table 4.4. Initial partition and solution obtained by the *RASS*



Source: Own elaboration.

In order to evaluate the evolution of the results from the initial partition up to the final results, Table 4.5 presents the value of the objective function at the end of each cycle in the application of the algorithm. The value of the objective function for the initial partition is 34.36 and in the first cycle a reduction of 24.15 is achieved. This value is reduced in the following cycles until achieving its minimum value in 1.08.

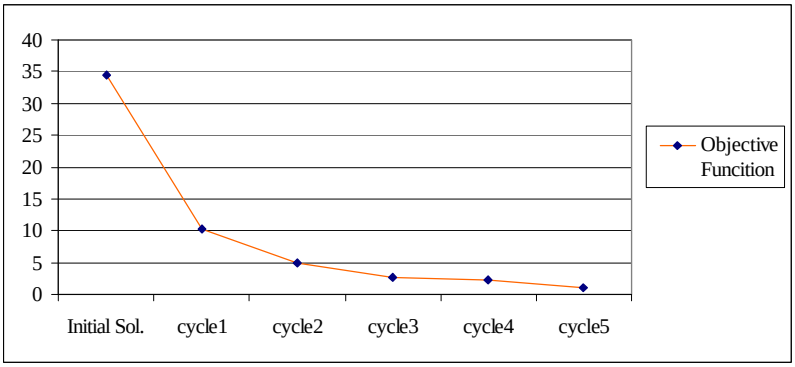
As it can be appreciated in Figure 4.3, the behaviour of the objective function is similar to the expected one: in the first cycles is where higher improvements are achieved. Also, it is confirmed that in every cycle the value of the objective function is improved, or at worst equal, in relation to the previous cycle.

Table 4.5. Values of the objective function in the initial partition and at the end of each cycle.

Regions	Initial	cycle 1	cycle 2	cycle 3	cycle 4	cycle 5
1	10.35	5.21	2.21	1.04	1.04	0.23
2	8.07	2.21	1.04	0.93	0.30	0.18
3	5.61	1.70	0.93	0.23	0.23	0.16
4	3.52	0.60	0.23	0.16	0.18	0.13
5	2.89	0.13	0.13	0.13	0.16	0.10
6	1.34	0.10	0.11	0.09	0.13	0.09
7	1.28	0.09	0.09	0.07	0.09	0.07
8	0.59	0.07	0.07	0.04	0.07	0.06
9	0.36	0.06	0.06	0.02	0.04	0.04
10	0.35	0.04	0.04	0.02	0.03	0.02
Objective function	**34.36**	**10.21**	**4.91**	**2.73**	**2.27**	**1.08**

Source: Own elaboration.

Figure 4.3. Evolution of the objective function during the application of *RASS*.



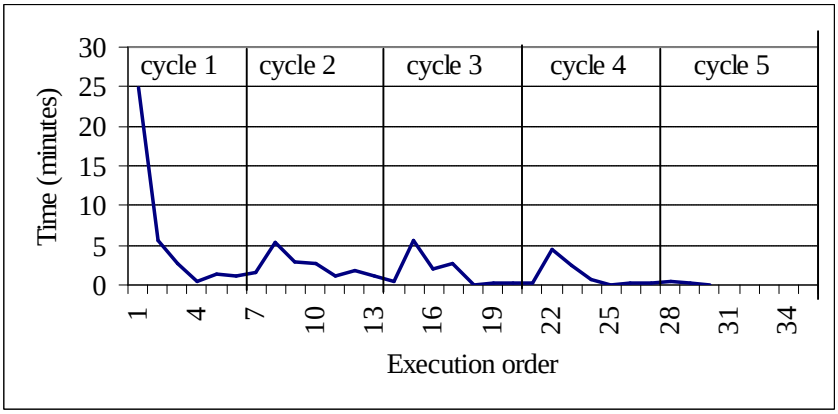
Source: Own elaboration.

The number of regions in the optimisation model was set to 4 (r = 4). With this value, the average number of areas where each optimisation model was running was 15. This number was enough to permit that the running times where appropriated with an average running time of 2.43 minutes by model. These running times are shown in Figure 4.4.

As it can be seen, the running times of the different optimisation models were higher at the beginning of each cycle and, in particular, for the first time it is executed (although it is also when a higher reduction in the objective function is achieved). This is related with the fact that in the first model of each cycle is executed considering the 4 (r) most heterogeneous regions, which can imply that the reassignation of the areas in these r regions can be very high. For this example, the first model has reassignated the

37% of these areas (or a 18.4% if we take into account the 38 areas) and has achieved a reduction in the objective function of 13.18 points, a 54.6% of the reduction obtained in the first cycle (or a 39.6% of the total reduction).

Figure 4.4. Running times of optimisation models.



Source: Own elaboration.

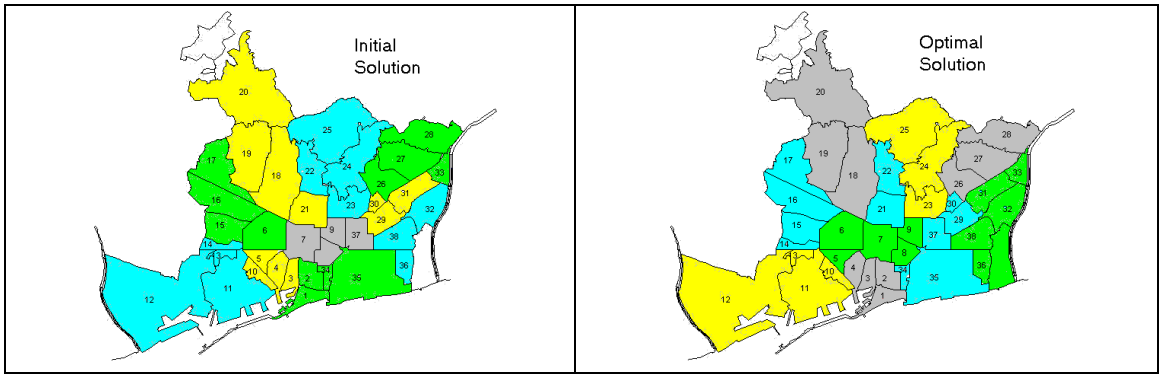
4.3.3. Sensitivity of the results to the initial partition.

How can the initial partition affect to the final result? In this sub-section, a different initial partition is used to solve the same problem as above. Thus, the initial partition in the step 1 of *RASS* will be closer to the optimum regional configuration. With this partition, a lower number of cycles and similar results as in the previous sub-section should bee expected.

In this case, the optimal configuration was found after 2 cycles (see Table 4.6), 3 cycles less than in the previous example. The results shown in the Table 4.7 and in the Figure 4.5, permit to conclude that, as before, the higher reductions in the objective function are achieved in the initial cycles of the *RASS*.

Regarding the impact of the first optimisation model on the objective function, now there is a reduction of 19.33 points (from 26.94 to 7.61), a 79,25% of the total obtained reduction in the first cycle. The 50% of the areas in the 4 (r) considered regions are now reassigned (a 21.1% in the 38 areas are considered).

Table 4.6. Initial partition (close to optimum) and obtained solution



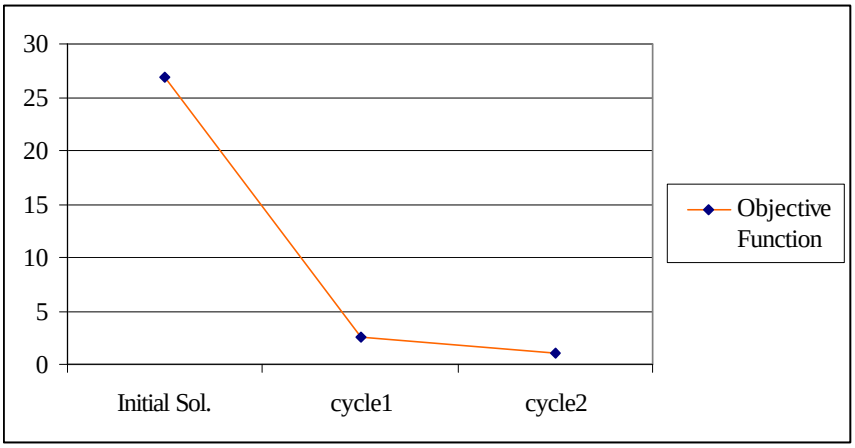
Source: Own elaboration.

Table 4.7. Values of the objective function in the initial partition (closes to the optimal solution) and at the end of each cycle.

Regions	Initial	cycle 1	cycle 2
1	10.31	1.71	0.23
2	6.83	0.18	0.18
3	2.33	0.15	0.16
4	1.95	0.13	0.13
5	1.93	0.10	0.10
6	1.04	0.09	0.09
7	0.93	0.07	0.07
8	0.88	0.06	0.06
9	0.65	0.04	0.04
10	0.09	0.02	0.02
Objective function	**26.94**	**2.55**	**1.08**

Source: Own elaboration.

Figure 4.5. Evolution of the objective function during the application of *RASS* with the initial partition closes to the optimal solution.



Source: Own elaboration.

4.4. Final remarks.

The obtained results permit to conclude that the *RASS*, due to the incorporation of a direct optimisation routine as part of the algorithm, has a big capacity to achieve global optimums in the context of regionalisation problems. However, it is worth mentioning that the relationship between the number of regions (*m*) and the number of areas (*n*) should be defined as a way that the number of regions considered by the direct optimisation model (*r*) must be 2 or higher and these regions should contain a number of areas in line with the computational capacity of the model. It has been calculated that the most appropriate relationship *m/n* must be above the 14%. For example, if it is considered a territory formed by 8,000 areas, the number of regions that can be obtained will be higher or equal than 1,120 regions (an average size of 7 areas per region). This relationship ensures that *r* can take values higher or equal than 2 without increasing substantially the running time.

If the relationship between regions and the number of areas is very low, one possible strategy could consist in designing nested regionalisation problems, which would imply the sequential application of the *RASS*. For example, the city of Barcelona is divided in 1,919 statistical sections (*Seccions Estadístiques*, SE), which are grouped in 248 small research areas (*Zones of Recerca Petites*, ZRP). These areas are also grouped in 110 basic statistical units (*Unitats Estadístiques Bàsiques*, UEB) that form the 38 big statistical areas (*Zones Estadístiques Grans*, ZEG). Last, the big statistical areas are grouped to obtained the 10 districts of the city[26]. Each territorial level is formed grouping the previous one, and this also guarantees that the different grouping levels are self-contained.

5. CONCLUSIONS

In this paper new methodologies to design regions from lower level territorial units (areas) were proposed considering not only their characteristics but also the relationships among them.

[26] For more information, see: http://www.bcn.es/estadisitica/catala/terri/index.htm.

These methodologies permit to avoid the use of *ad-hoc* regionalisation to obtain territorial units that are representative of the considered phenomenon. This aspect is especially relevant if it is taken into account that statistical and econometrical results are sensitive to different levels of aggregation and scale.

After it was made a survey of more relevant regionalisation methods, in section 2 a linear optimisation model has been proposed to find the optimal aggregation of different areas in a given number of regions from the consideration of a geographical contact matrix and a relationships matrix. The minimisation of the "internal" heterogeneity of each region permits to find homogeneous regions according to the considered criteria.

The possibility of treating the regionalisation problem as a linear model permits to ensure that, by its mathematical properties, the feasible region is convex and, as a result, it is possible to find the optimal solution. Another advantages of this kind of formulation are that it is easy to implement in a great variety of commercial software without paying a high price for it, and the flexibility when some changes or additional constraints are needed.

The obtained empirical evidence permits to affirm that the proposed methodology has a great capacity to identify different complex territorial configurations. The model takes into account the contiguity constraint but without conditioning the shapes that those regions can adopt.

It is also important to highlight that the model permits to easily introduce additional restrictions in the regionalisation process. As an example, it has been shown the possibility of introducing two additional restrictions: the minimum population requirement and the mandatory isolation.

In a second stage, and according to the second specific objective formulated in this paper, an algorithm called *RASS* (*Regionalisation Algorithm with Selective Search*) has been formulated in section 3 as a way of improving the computational capacity of the direct optimisation model. This algorithm tries to take profit of the advantages of applying direct optimisation to a given territorial portion that varies in each iteration, thanks to a selective search strategy. These characteristics permit to the *RASS* to escape from local optimum.

The obtained results with the *RASS* have shown its utility, as in a 100% of the considered simulations the global optimum was found and in a running time considerably lower than the one obtained applying the direct optimisation model.

Table 6.1 shows the main characteristics of regionalisation models proposed in this paper and the previous models. As it can bee seen, both linear optimisation model and *RASS* algorithm overcome some inconvenient in existing methodologies.

A common characteristic in all presented regionalisation methods is that the **number of regions** to be formed is a exogenous variable. These **regions are obtained automatically** by applying the proposed models. This is an advantage with regard to regionalisation models based on clustering techniques, where it is necessary to make several proves before obtaining the decides number of regions.

To take into account the **relationships between areas** to be grouped, the proposed models allow to incorporate them through a squared matrix that contain a relationship measure between each pair o areas. Cutting models only take into account relationships between contiguous areas ($w : E \rightarrow N$), and iterative realocation algoritms as AZP do not uses these relationships in their searching processes.

Non metric relationships between areas can be used in proposed models. In contrast, models based on centroids selection have to use metric relationships in order to assure that the obtained regions are contiguous after assignation process.

The **contiguity relationships** between areas to be grouped are an important input in proposed models, such information is not taken into account in clustering models when two stages regionalisation strategy is applied. Centroid based regionalisation models do not use the contiguity relationships because in the assignation process contiguity is obtained by using metric relationships between areas.

Shapes flexibility is an important characteristic in some regionalisation processes where it is necessary that regional shapes only depend on data characteristics and are not imposed by the considered methodology. Centroid based regionalisation process tend to produce compact areas.

To find the **global optimum** solution of a regionalisation problem can only be guarantied by applying linear optimisation models as cutting models, centroid models and the linear regionalisation model propose in this paper. In iterative models the optimal solution could not be founded, but these kind of models are suitable to solve **large regionalisation problems**.

Finally, only in iterative regionalisation models it is necessary an **initial feasible solution** in order to start the searching process.

Table 3.1. Comparison between revised regionalisation models and the linear optimisation model propose in this paper.

Regionalisation methodology / Characteristics	Clustering		Mathematical programming				Iterative algorithms	
	Conventional	Adapted	Non linear regionalisation	Cutting models	Based on centroids	**Linear optimisation model**	Iterative realocation algorithms	***RASS***
The number of groups is given	✓	✓	✓	✓	✓	✓	✓	✓
Automatic regionalisation	×	×	✓	✓	✓	✓	✓	✓
Relationships along areas (n×n)	✓	✓	✓	×*	✓	✓	✓**	✓
Non metric relationships	✓	✓	✓	✓	×	✓	✓**	✓
Contiguity relationships	×	✓	✓	✓	×	✓	✓	✓
Shapes flexibility	✓	✓	✓	✓	×	✓	✓	✓
Optimal solution	×	×	×	✓	✓	✓	×	×
Acceptable for large problems	×	×	×	×	×	×	✓	✓
Initial feasible solution	×	×	×	×	×	×	✓	✓

Shared column: models proposed in this paper.

* It is only taken into account the relationships between each area an its neighbouring areas (i.e. first order relationships).

** They can be incorporated into de objective function but they are not taken into account in any step of the algorithm.

6. REFERENCES

Aarts, E. and Lenstra, J. K. (1997): *Local search in combinatorial optimization*. Chichester, New York [etc.]: John Wiley and Sons.

Ahuja, R. K., Magnanti, T. L. and Orlin, J. B. (1993): *Network flows: theory, algorithms, and applications.* Englewood Cliffs, N.J: Prentice Hall, cop.

Albert, J. M., Mateu, J. and Orts, V. (2003): *Concentración versus dispersion: Un análisis especial de la localización de la actividad económica en la U.E.*, mimeo.

Amrhein, C. G. and Flowerdew, R. (1992): 'The effect of data aggregation on a Poisson regression model of Canadian migration', *Environment and Planning A*, 24, 1381-91.

Anderberg, M. R. (1973): *Cluster analysis for applications.* New York: Academic Press.

Aurenhammer, F. (1991): 'Voronoi diagrams - a survey of a fundamental geometric data structure', *ACM Computing Surveys*, 23, 345-405.

Benabdallah, S. and Wright, J. R. (1992): 'Multiple subregion allocation models', *Journal of Urban Planning and Development*, 118 (1): 24-40.

Browdy, M. (1990): 'Simulatea Annealing - an improved computer model for political redistricting', *Yale Law and Policy Review*, 8, 163-79.

Calciu, M. (1996): 'Une méthode of classification sous contrainte of contiguïté en géo-marketing' Cahiers of recherche of l'IAE of Lille (96/5). Université des Sciences et Technologies of Lille.

Crescenzi, P. and Kann, V. (2004): 'A compendium of NP optimization problems'. http://www.nada.kth.se/~viggo/wwwcompendium/

Dantzing, G. B. and Ramser, J. H. (1959): 'The truch dispatching problem', *Management Science*, 6, 80-91.

EUROSTAT (2004): 'Nomenclature of territorial units for statistics – NUTS. Statistical Regions of Europe'. http://europa.eu.int/comm/eurostat/ramon/nuts/home_regions_en.html. (01/03/04).

Ferligoj, A. and Batagelj, V. (1982): 'Clustering with relational constraint', *Psychometrika*, 47, 413-26.

Ferligoj, A. and Batagelj, V. (1983): 'Some types of clustering with relational constraints', *Psychometrika*, 48, 541-52.

Fisher, M. M. (1980): 'Regional taxonomy', *Regional Science and Urban Economics*, 10, 503-37.

Fotheringham, A. S. and Wong, D. W. S. (1991): 'The modifiable areal unit problem in multivariate statistical analysis', *Environment and Planning A*, 23, 1025-44.

Glover, F. (1977): 'Heuristic for integer programming using surrogate constraints', *Decision Sciences*, 8, 156-66.

Glover, F. (1989): 'Tabu search, part I', *ORSA Journal on Computing*, 1, 190-206.

Glover, F. (1990): 'Tabu search, part II', *ORSA Journal on Computing*, 2, 4-32.

Gordon, A. D. (1996): 'A survey of constrained classification', *Computational Statistics & Data Analysis*, 21, 17-29.

Gordon, A. D. (1999): *Classification*. 2nd edition, Boca Raton [etc.]: Chapman & Hall/CRC.

Gower, J. C. and Legendre, P. (1986): 'Metric and euclidean properties of dissimilarity coefficients', *Journal of Classification*, 3, 5-48.

Graham, R. R. and Hell, P. (1985): 'On the history of the minimum spanning tree problem', *Annals of the history of computing*, 7, 43-57.

Haining, R. P., Wise, S. M. and Ma, J. (1996): 'The design of a software system for interactive spatial statistical analysis linked to a GIS', *Computational Statistics*, 11, 449-66.

Hiriart, J. B., Oettli, W. and Stoer, J. (1983): *Optimization: Theory and Algorithms*. New York [etc.]: Marcel Dekker, cop.

Horn, M. E. T. (1995): 'Solution techniques for large regional partitioning problems', *Geographical Analysis*, 27, 230-48.

Jobson, J. D. (1991): *Applied multivariate data analysis: Categorical and multivariate methods*. New York [etc.]: Springer.

Kirkpatrick, S., Gelatt, C. D. and Vecchi, M. P. (1983): 'Optimization by simulated annealing', *Science*, 220, 671-80.

Kohonen, T. (1984): *Self-Organisation and Associative Memory*. Berlin [etc.]: Springer.

Laport, G. and Osman, I. H. (1995): 'Routing problems: A bibliography', *Annals of Operations Research*, 61, 227-62.

López-Bazo, E., Vaya, E., Mora, A. and Suriñach, J. (1999): 'Regional Economic Dynamics and Convergence in the European Union', *Annals of Regional Science*, 33, 343-70.

Macmillan, B. and Pierce, T. (1994): 'Optimization modelling in GIS framework: the problem of political redistricting', in Fotheringham, S. and Rogerson, P. (eds.), *Spatial analysis and GIS*, London [etc.]: Taylor & Francis, pp 221-46.

Martin, D., Nolan, A. and Tranmer, M. (2001): 'The application of zone-design methodology in the 2001 UK Census', *Environment and Planning A*, 33, 1949-62.

Matula, D. W. and Sokal, R. R. (1980): 'Properties of Gabriel graphs relevant to geographic variation research and the clustering of points in the plane', *Geographical Analysis*, 12, 205-22.

Murtagh, F. (1985): 'A survey of algorithms for contiguity-constrained clustering and related problems', *The Computer Journal*, 28 (1): 82-8.

Neves, M. C., Freitas, C. C. and Câmara, G. (2001): 'Mineraçâo of datos em grandes bancos of datos geographical.' Brasil: Instituto nacional of pesquisas espaciais. Ministério da ciência e tecnologia.

Ohsumi, N. (1984): 'Practical techniques for areal clustering', in Diday, E., Jambu, M., Lebart, L., Pagès, J. and Tomassone, R. (eds.), *Data analysis and informatics*, Vol III, North-Holland, Amsterdam, pp 247-58.

Openshaw, S. (1977): 'Algorithm3: a procedure to generate pseudo random aggregation of N zones into M zones where M is less than N', *Environment and Planning A*, 9, 1423-28.

Openshaw, S. (1984): The modifiable areal unit problem, *Concepts and Techniques in Modern Geography*, 38 (GeoBooks, Norwich).

Openshaw, S. (1992): 'Some suggestions concerning the development of artificial intelligence tools for spatial modelling and analysis in GIS', *The Annals of Regional Science*, 26, 35-51.

Openshaw, S. and Rao, L. (1995): 'Algorithms for reengineering 1991 census geography', *Environment and Planning A*, 27, 425-46.

Openshaw, S. and Taylor, P. J. (1981): 'The modifiable areal unit problem', in Wrigley, N. and Bennett, R. J. (eds.), *Quantitative Geography*, London, pp 60-70.

Openshaw, S. and Wymer, C. (1995): 'Classifying and regionalizing census data', in Openshaw, S. (eds.), *Census Users Handbook*, Cambridge, UK: Geo Information International, pp 239-70.

Perruchet, C. (1983): 'Constrained agglomerative hierarchical classification', *Pattern Recognition*, 16, 213-17.

Semple, R. K. and Green, M. B. (1984): 'Classification in human geography', in Gaile, G. L. and Wilmott, C. J. (eds.), *Spatial statistics and models*, Reidel, Dordrecht, pp 55-79.

Toussaint, G. T. (1980): 'The relative neighbourhood graph of a finite planar set', *Pattern Recognition*, 12, 261-68.

Webster, R. and Burrough, P. A. (1972): 'Computer-based soil mapping of small areas from sample data II. Classification smoothing', *Journal of Soil Science*, 23, 222-34.

Wise, S. M., Haining, R. P. and Ma, J. (1997): 'Regionalisation tools for exploratory spatial analysis of health data', in Fisher, M. M. and Gentis, A. (eds.), *Recent Developments in Spatial Analysis: Spatial statistics, behavioural modelling, and computational intelligence*, Berlin [etc.]: Springer, pp 83-100.