# A Multiobjective Approach for the Heuristic Optimization of Compactness and Homogeneity in the Optimal Zoning

B. Bernábe-Loranca*[1], C. A. Coello-Coello[2], M. Osorio-Lama[3]

[1] Facultad de Ciencias de la Computación
[3] Facultad de Ingeniería Química
Benemérita Universidad Autónoma de Puebla, México
*beatriz.bernabe@gmail.com
[2] Departamento de Computación
Centro de Investigación y Estudios Avanzados del Instituto Politécnico Nacional
México, D. F., México

## ABSTRACT

This paper presents a multiobjective methodology for optimal zoning design (OZ), based on the grouping of geographic data with characteristics of territorial aggregation. The two objectives considered are the minimization of the geometric compactness on the geographical location of the data and the homogeneity of any of the descriptive variables. Since this problem is NP hard [1], our proposal provides an approximate solution taking into account properties of partitioning algorithms and design restrictions for territorial space. Approximate solutions are generated through the set of optimum values (Maxima) and the corresponding minimals (dual Minima) [2] of the bi-objective function using Variable Neighborhood Search (VNS) [3] and the Pareto order defined over this set of values. The results obtained by our proposed approach constitute good solutions and are generated in a reasonably low computational time.

Keywords: Optimal Zoning, compactness-homogeneity, maxima, multiobjective optimization.

## RESUMEN

Se presenta una propuesta de optimización multiobjetivo para la zonificación óptima (ZO) basada en la agrupación de datos geográficos bajo características de agregación territorial. Los dos objetivos a minimizar son la compacidad geométrica en la ubicación geográfica de los datos y la homogeneidad de alguna de sus variables descriptivas. Dado que este problema es NP Duro [1], nuestra propuesta proporciona una solución aproximada tomando en cuenta las propiedades de los algoritmos de particionamiento y de las restricciones espaciales para diseño territorial. Se generan soluciones aproximadas a través del conjunto de valores máximos (Maxima) y el dual (Minima) [2] de la función bi-objetivo con la heurística de Búsqueda por Entorno Variable (BEV) [3] y el orden Pareto definido sobre este conjunto de valores. Los resultados obtenidos por nuestra propuesta constituyen buenas soluciones, y se produjeron en tiempos de cómputo razonablemente cortos.

## 1. Introduction

In the problem of Optimal Zoning (OZ), the goal is to obtain a spatial data partitioning named BGAs (Basic Geostatistical Areas). Its composition consists of two components: geographical coordinates in the plane R2 and a vector of 171 census descriptive characteristics [4]. The first component allows us to obtain a distance matrix for the process of calculating the geometric compactness, which is one of the objective functions to minimize. The second objective function is the minimization of the homogeneity of any of the census variables stored in the vector´s

description, and selected on the basis of a particular interest. From this point of view, we look for a partition that consists of a set of classes with components that are very close geographically, and balanced according to one of its census variables.

To solve the issue of the bi-objective geographical partitioning, a heuristic methodology named Variable Neighborhood Search (VNS) [3] was used in a multiobjective optimization framework that allowed us to find a set of pairs of nondominated

and non-comparable solutions named minimals [2]. This paper is organized as follows. The Introduction is in Section 1. Section 2 presents the basic definitions and concepts used throughout this study. The problem statement of the Optimal Zoning (OZ) problem is described in Section 3. Section 4 shows the method used to find the set of nondominated solutions (minimals). The conclusions are shown in Section 5.

## 2. Preliminary and theoretical aspects

Multiobjective problems can be more clearly understood if the relationships among their characteristics, constraints and main objectives to be improved are identified. For such problems, it is possible to have a model that consists of mathematical functions and restrictions expressed as constraints.

We have taken some definitions from [5] which are necessary to formally introduce the multi-objective partitioning problem on which this paper is focused. The definitions are the following

**Notions of Optimality**
The minimization in multicriteria optimization problems in general is:
"min" $(f_1(\boldsymbol{x}), f_2(\boldsymbol{x}),...,f_p(\boldsymbol{x}))$ subject to $\boldsymbol{x} \in X$

The fundamental importance of efficiency (Pareto optimality) is based on the observation that any x which is not efficient cannot represent a most preferred alternative for a decision maker, because there exists at least one other feasible solution $x' \in X$ such that $f_k(x') \le f_k(x)$ for all $k = 1,..,p$, where a strict inequality holds at least once, i.e., $x'$ should clearly be preferred to $x$.

Let $S$ be any set. A binary relation on $S$ is a subset $R$ of S×S:

*Definition 1.* A binary relation $R$ on a set $S$ is:
● an equivalence relation if it is reflexive, symmetric, and transitive,
● a preorder (quasi-order) if it is reflexive and transitive.

In the case of $R$ being a preorder set the pair $(S, R)$ is called a *preordered set*. For convenience, we will write $s^1 \le s^2$ as shorthand for $(s^1, s^2) \in R$ and $s^1 \nleq s^2$ as a shorthand for $(s^1, s^2) \notin R$. Also, we will

interchangeably refer to the relation $R$ or to the relation $\le$. This notation can be read as "preferred to".

Given any preorder $\le$, two other relations are closely associated with $\le$ :
$s^1 \prec s^2$ iff $s^1 \le s^2$ and $s^2 \nleq s^1$
$s^1 \sim s^2$ iff $s^1 \le s^2$ and $s^2 \le s^1$

By the choice of an order $\le$ on $\mathbf{R^p}$, now it's possible define "min"

"min" $f(\boldsymbol{x}) = $ "min" $(f_1(\boldsymbol{x}), f_2(\boldsymbol{x}),.., f_p(\boldsymbol{x}))$ $\boldsymbol{x} \in X$
With the multiple objective functions, we can evaluate objective value vectors $(f_1(\boldsymbol{x}), f_2(\boldsymbol{x}),..,f_p(\boldsymbol{x}))$. However, we can see that these vectors $y = f(\boldsymbol{x})$, $\boldsymbol{x} \in X$ are not always directly compared in objective space, i.e., $\mathbf{R^p}$.

In general, the objective function vectors are mapped from $\mathbf{R^p}$ to an ordered space, e.g. $(\mathbf{R^p}, \le)$, where comparisons are made using the order relation $\le$ (model map). Now, we can summarize the elements of a Multicriteria Optimization Problem (MOP):

The feasible set $X$, the objective function vector $f = (f_1,..,f_p) : X \to \mathbf{R^p}$, the objective space $\mathbf{R^p}$, the ordered set $(\mathbf{R^p}, \le)$, the model map $\theta$.

Feasible set, objective function vector $f$, and objective space are data of the problem to be solved. The model map provides the link between objective space and ordered set, in which, the meaning of the minimization is defined. Thus, with the three main aspects (i.e., data, model map, and ordered set) the classification $(X, f, \mathbf{R^p})/\theta/(\mathbf{R^p}, \le)$ completely describes a multicriterion optimization problem.

*Definition 2.* A feasible solution $x^* \in X$ is called an optimal solution of a Multiobjective Problem (MOP) $(X, f, \mathbf{R^p})/\theta/(\mathbf{R^p}, \le)$ if there is no $x^* \in X$, $x \ne x^*$ such that $\theta(f(x)) \le \theta(f(x^*))$.

For an optimal solution $x^*$, $\theta(f(x^*))$ is called an optimal value of the MOP. The set of optimal solutions is denoted by $Opt((X, f, \mathbf{R^p})/\theta/(\mathbf{R^p}, \le))$.

*Definition 3.* A feasible solution $x^\wedge \in X$ is called efficient or Pareto optimal, if there is no other $x \in X$ such that $f(x) \le f(x^\wedge)$. If $x^\wedge$ is efficient, $f(x^\wedge)$ is called

nondominated point. If $x^1$, $x^2$, $\in X$ and $f(x^1) \leq f(x^2)$ we say $x^1$ dominates $x^2$ and $f(x^1)$ dominates $f(x^2)$.

The set of all nondominated $x^\wedge \in X$ is denoted as $X_E$ and called the efficient set. The set of nondominated points $y^\wedge = f(x^\wedge) \in Y$ where $x^\wedge \in X_E$ is denoted $Y_N$ and called the nondominated set [5]. Finally, we present concepts that we have adapted to the problem discussed in this paper. The previous definitions allow us to define the problem of our interest as follows:

*Definition 4.* A MOP is defined in the case of minimization (and similarly for the case of maximization) as:
Minimize f(x) since $f: F \subseteq R^n \rightarrow R^q$, $q \geq 2$ and evaluated in

$$A = \{a \in F: g_i(a) \leq 0, i = 1, ..., n\} \neq \emptyset.$$

The constraints set A, called the feasible region, and consists of the regular functions
$g_i; R^n \rightarrow R$.

*2.1 Pareto Dominance*

*Definition 5.* Pareto dominance (DP). A vector $u = (u_1,..,u_k)$ dominates a vector $v = (v_1,..,v_k)$ (denoted by $u \preccurlyeq v$) if and only if $u$ is partially less than $v$ .

That is:
$u_i \leq v_i$ for any $i = 1,..,k$ and denoted by
$u < v$ if $u \preccurlyeq v$ and $u \neq v$ .

Another common concept used as Pareto dominance is expressed as:

*Definition 6.* Given the multiobjective problem Minimize $f(x)$ where $f : F \subseteq R^n \rightarrow R^q$, $q \geq 2$ with $A \subseteq F$ the feasible region. A vector $x^* \in A$ is *not dominated* or a *Pareto optimal solution* if a vector $x \in A$ such that $f(x) < f(x^*)$ does not exist [6].

Two vectors $x, x' \in A$ are not comparable iff

$$f(x) \npreccurlyeq f(x') \text{ and } f(x') \npreccurlyeq f(x).$$

The answer to the problem of finding the best solutions (nondominated solutions) to a multiobjective problem is known as the solution set of the problem, and the group of values of the objective functions, with a domain restricted to the vectors of the solution set (i.e., nondominated vectors) is known as the Pareto frontier.

In summary, the set of Pareto optimal solutions is the space solution of the problem and the Pareto Frontier is its image with respect to the function $f : F \subseteq R^n \rightarrow R^q$, q ≥ 2 being optimized [6] .

*Definition 7.* Let $(A, \preccurlyeq)$ be a partially ordered set, where an element $x^* \in X \subseteq A$ is called a minimal element of X, iff an element $x \in X$ such that $x < x^*$ does not exist. The set of all minimal elements is denoted as $M(A, \preccurlyeq)$.

In this paper, we propose an adaptation of definitions 5, 6 and 7 to obtain:

*Definition 8.* **Let** $F \subseteq R^n$ and C = {**P**: **P** is a partition of *F*}

Minimize *f(p)* where $f : C \subseteq 2^F \rightarrow R^q$, q ≥ 2 with $A \subseteq$ **C** the feasible region. A partition $p^* \in A$ is *not dominated* or a *Pareto optimal solution* if a partition $p \in A$ such that $f(p) < f(p^*)$ does not exist.

Two partitions $p, p' \in A$ are not comparable iff

$$f(p) \npreccurlyeq f(p') \text{ and } f(p') \npreccurlyeq f(p)$$

## 3. The Optimal Zoning Problem

The Zoning Problem is informally defined as a grouping process of geographical areas with the assumption of the existence of a relationship between the characteristics of the data that constitutes the metropolitan area, the population variables and the properties of the problem. Here, the main issue is to know how a variable census - population is distributed or concentrated in certain territorial spaces.

The applications of the Zoning Problem are diverse, and primarily focused on solving population issues. For example, in order to serve a segment of the population that has no basic services, as drainage, the problem can be addressed by grouping units of territorial areas (in this case BGAs), in larger groups, so that the areas that belong to each group, are very close and compact, and can be considered in the grouping of the variables related to sewer services [7].

This procedure involves creating and analyzing small groups because it is not possible to study the full extent as a single territorial unit. At this point, and according to the territorial design properties, Optimal Zoning can be seen as a case arising in territorial design. Since the problem of optimal zoning seeks answers to several population issues (distribution, concentration, density or centralization of the population in a metropolitan area), the method for addressing optimal zoning uses compact territorial units, in this case BGAs with two objectives to consider: a) to satisfy the geometric compactness property required in territorial design, and b) to balance a specific population variable in each group (homogeneity). This means that the partitioning method for BGAs minimizes two objective functions simultaneously and checks that the solutions generated are minimals [7].

### 3.1 Solution Strategy for Optimal Zoning

Generally, the clustering can be addressed as a combinatorial optimization problem where the clusters are a partition of a set of objects. If only one objective function is optimized, the clustering algorithm simply incorporates a heuristic method to find a global solution [8]. But given the strong relationship between spatial design problems and geographical partitioning, we know that these problems, in order to be solved, require a clustering methodology applied on the geographical space where the conditions for the grouping are territorial design spatial properties such as continuity, connectedness, compactness and homogeneity. Similarly, the combination of geographical data in territorial design has produced good results if one optimizes only one objective function f and leaves other properties as constraints.

The multiobjective partitioning in territorial design is an alternative solution to such problems as it offers a set of solutions where the decision maker can have more than one solution and choose the most appropriate for a particular application. There are few works on multiobjective partitioning due to its various sources of complexity: modelling and characterization of the problem, the development of implementations articulated with a heuristic method and the generation of nondominated solutions [9, 10, 11, 12, 13].

Since optimal zoning is a multiobjective problem, the vector of census population variables and the geographical coordinates are data that should be considered in the clustering process in order to create the BGAs. In the aggregation process, which is NP-hard in nature, the kinds of clusters obtained are subject to compliance with the minimization of two cost functions: a measure of distance in geographic space and a balance or uniformity of population variables.

This implies that the optimization of these objective functions will be addressed with a heuristic approach, while the solution to the competition between these two functions will be solved by multi-objective methods for obtaining the set of nondominated solutions (minimals).

The general strategy used to solve the problem of optimal zoning is:
a) To develop a geographical partitioning combinatorial model that considers the spatial layout properties (BGAs restrictions of compactness and homogeneity).

b) To solve first the geographical partitioning problem, considering only one objective function: geometric compactness. Once a compact partition is obtained, its homogeneity is simultaneously calculated.

c) To develop an algorithm that optimizes both compactness and homogeneity with the use of VNS to ensure the generation of quality solutions, according to a) and b) [7].

d) While the solutions are generated with the heuristic algorithm, it is necessary to use a multi-objective method for the construction of an efficient set of solutions defining a Pareto frontier [2, 10]. The method used to find nondominated solutions is based on the theory of order and obtains the minimals [2, 7].

In particular, we emphasize the use of [2] to solve the multiobjective problem because of the importance of the concepts of Maxima and their dual Minima used to find nondominated solutions for optimal zoning problems. On the other hand, other proposals to address multiobjective problems have been studied [10, 11, 12, 13]. The aim has been to identify aspects of these methods that

could be adaptable to the multiobjective problem discussed in this article. However, after an analysis of several works, the core support used for the multiobjective optimal zoning method proposed in this work was the algorithm proposed in [2].

### 3. 2 Discretization and representation

The geographical clustering is done in a physical space. In the finite geographical units called BGAs (Basic Geostatistical Areas) in our research, each element is represented by its spatial location and an array of descriptive variables (census quantified variables). The problem is discrete, combinatorial and mixed-integer, and the aggregation is performed under the partitioning properties. The geographically clustering process produces compact groups of BGAs. To do this, we use an objective function that minimizes the sum of the distances between the elements of the BGAs of each group and its centroid. The homogeneity is optimized seeking a grouping balance in a census variable of interest.

The clustering strategy randomly chooses a certain number of BGAs as centroids, which are used to represent and identify their respective groups. Those BGAs which are not centroids and have the shortest distance to a particular BGA-centroid, are members of a group. This is defined as geometric compactness [8]. Once the groups that minimize the sum of the distance of the members of each group with its centroid is formed, the homogeneity of the groups created is estimated, since both objectives are defined in the same partition, i.e., the functions to optimize have the same domain for all the objectives of the multiobjective problem [10]. Thus, on the same partition, the compactness and homogeneity are optimized.

The characteristics of this multiobjective problem are:

1. Each BGA must belong to only one group (constraint for compactness)
2. In one group, the value of each parameter is the value of the census variable (constraint for homogeneity).
3. The groups are disjoint (constraint for compactness)
4. There are no empty groups (constraint for compactness)

5. The population variables may or may not be bounded (constraint for homogeneity)
6. All the variables or only a subset of them can be in the cluster (constraint for homogeneity).
7. The BGAs assigned to each group must conform a compact group (Compactness Objective).
8. The groups should be balanced with respect to a measurable characteristic (Homogeneity Objective).

The following equations, will use this notation:

Let $GU = \{x_1, x_2, .., x_n\}$ be the initial set of $n$ geographical units, where $x_i$ is the $i^{th}$ geographical unit, $k$ is the number of zones or groups, $Z_i$ is the set of geographical units that belong to area $i$, $n$ is the number of geographical units that belong to area $i$, $c_t$ is the centroid and $d(i, j)$ is the Euclidean distance from node $i$ to node $j$ (from one BGA to another).

### 3.2.1 First Objective: Minimization of Distances (Compactness Objective)

It minimizes the intra-class distance between BGAs, defined as

$$\underset{k=1,...,n}{Min}\left\{\left\{\sum_{t=1}^{k}\sum_{i\in c_t}d(i,c_t)\right\}\right\} \qquad (1)$$

(where $n$ is the number of iterations, $t$ the number of centroids and $i$ the number of BGAs in the proximity of that centroid). Once the number $k$ of centroids $c_t, t = 1, .., k$ is decided, the other BGAs are assigned in a random selection, to the nearest centroids $c_t$. For each BGA $i$ is fulfilled:

$$\underset{t=1,...,k}{Min}\left\{d(i,c_t)\right\} \qquad (2)$$

The sum of the distances of the BGAs assigned to each centroid $k$ is calculated, selecting the arrangement that minimizes the sum in (2).

*3.2.2 Second objective: Homogeneity for census variables*

When the BGAs in a metropolitan area are clustered, seeking for a balance in any census variable (*VA*), it can be said that the BGAs are partitioned under a criteria of homogeneity. The equilibrium or homogeneity for a specific census variable is the second objective considered. To balance the groups, an ideal average for the variable of interest should be obtained; it ideally happens when all the members in the group have the same value. As it is not common practice, the actual average is calculated for each group and subtracted from the ideal average. Minimizing the sum of this difference is the homogeneity objective function to be optimized.

Let *VA* be the set of measurable attributes, from which a subset will be selected according to the problem.

$VA_{kj}$ is the value of the $k^{th}$ attribute contained in the $j^{th}$ geographical unit (*GU*).

$\alpha_k$, $\beta_k$ are the parameters and tolerances for $VA_k$ in any geographical unit *GU*.

$\alpha_k \leq VA \leq \beta_k$ are bounds on the variables.

Then $VA_{ki} = \sum_{j=1}^{n} VA_{kj} X_{ij}$ (3)

.is the value for the $k^{th}$ territorial group (*TG*), where $X_{ij}$ is a binary variable that equals one if the territorial group $i$ belongs to the basic geographic unit $j$ and zero, otherwise. This value is defined as the target for the $k^{th}$ attribute in any geographical unit *GU*.

$$\overline{VA_k} = i / m \sum_{j=1}^{n} VA_{kj} \qquad (4)$$

Is the ideal target value for the $k^{th}$ attribute in every territorial group.

From (3) and (4), the objective function for homogeneity can be defined as:

$$Hom = (\overline{VA_k} - VA_{ki}) \qquad (5)$$

Rewriting (1) and defining *BGU* as a Basic Geographic Unit, we have:

$$D_{ij} = \sum_{j=1}^{n} d(c_i, BGU) X_{ij} \qquad (6)$$

Finally, the multiobjective model for the optimal zoning problem (compactness and homogeneity) can be written as:

Minimize $f(x) = (f1, f2)$ (7)

Where
$f1$ : is the cost of minimizing the distance between BGAs according to equation (6).

$f2$ : is the cost of minimizing the homogeneity for a census variable of the BGAs according to equation (5).

The functions 1 and 2 in (7) are subject to the following constraints:

$Z_i \neq \varnothing$ for $i = 1, \ldots, k$ (the groups are not empty)

$Z_i \cap Z_j = \varnothing$ for $i \neq j$ (Every BGA can only belong to one group)

$\bigcup_{i=1}^{k} Z_i = UG$ (The union of all groups are all the BGAs)

$\sum_{i=1}^{m} X_{ij} = 1$ is the allocation of BGAs (BGA mean *GU*),

with $X_{ij} = 1$ if $GU \in TG_i$ o
$X_{ij} = 0$ if $GU \notin TG_i$

## 4. Non-dominated solutions to Optimal Zoning

In order to find the Pareto optimal solutions to the problem of our interest, Pareto ordering (implicit in the definition of Pareto dominance) was initially applied to the VNS generated solutions of the problem at hand. These solutions are a partially

ordered sets of pairs $(c_l, h_l)$ according to the definition of Pareto dominance [6].

However, it is necessary to verify that the set of final solutions $(c_l, h_l)$ obtained in a test, satisfies the definition of Pareto dominance and is indeed a set of nondominated solutions. For that purpose, the optimal zoning solutions have been processed with an application named NODOM or NDOMINATED (nondominated), built according to the algorithm presented in [2].

The general description NODOM is:
Let $U_1, U_2, \ldots, U_d$ be totally sets ant let $V$ be a set of $n$ d-dimensional vectors in the cartesian product $U_1 \times U_2 \times \ldots \times U_d$. For any vectors $v$ in $V$ let $x_i(v)$ denote the $i^{th}$ component of $v$. A partial ordering $\leq$ is definided on $V$ in a natural way, that is, for $v, u \in V, v \leq u$ if only if
$x_i(v) \leq x_i(u)$ for all $i = 1, \ldots, d$ where $\leq$ is the total ordering on $U$. For $v \in V$, $v$ is defined to be a maximal element of $V$ if there does not exist $u \in V$ such that $u \geq v$ and $u \neq v$.

The algorithm was implemented in the C programming language and identifies nondominated solutions of a dataset. The source code and information about its use are available in [11].

NODOM works like this: accepts as an input file a set of vectors (txt format document). NODOM needs as well the number of objectives as input. The output is a file that contains the set of nondominated solutions.

Once it is understood that the maxima are "nondominated points" but also non-comparable (for the problem solved in this paper), it is necessary to propose a method to discover all pairs of equally non-comparable solutions.

Given this situation, we have to revisit the ordering imposed by Pareto dominance, since this relation must include noncomparability.
A Pareto ordering implies:

Given a solution $(a, b)$ the following solution $(a', b')$ is accepted if:

$$(a' > a \land b' = b) \lor (b' > b \land a' = a) \lor (a' > a \land b' > b) \lor (a = a' \land b = b') \qquad (1a)$$

When a comparison of the pair of solutions is made using expression 1a, a point at the Pareto boundary is reached.

The negation of expression 1a allows the production of approximations to the Pareto boundary by several chains. However, it is also necessary to iteratively examine that the non-comparable solutions fulfill the Pareto dominance. Lastly, the solution set obtained is a set of minimals (Pareto boundary).

A pair of solutions is non-comparable in a given partial order if it does not fulfill the trichotomy property, in this case it means that two pairs $(a, b)$ and $(a', b')$ are Pareto non-comparable if

$$\neg ((a, b) < (a', b')) \land \neg ((a', b') < (a, b)), \text{ that is}$$
$$(a > a' \lor b > b') \land (a' > a \lor b' > b) \qquad (2b)$$

Under this strict partial order, that we have called "Pareto Non-Comparable" (2b), suitably combined in an iterative manner with Pareto dominance, we obtain all pairs of minimal solutions; a fact that is confirmed when NODOM has been applied to all the solutions generated by optimal zoning.

Some other authors also considered the properties of the non-comparable relation as an extension of the classical Pareto ordering and used these properties to find the Pareto frontier for their own specific problems [12, 13].

Finally, note that the average of the minimal solutions is non-comparable [7].

If the solutions are not comparable (non-comparable), this does not imply that they are minimal. However, if we have solutions that are not comparable and also nondominated then these solutions are minimal.

This set of solutions is defined as minimal for optimal zoning and are precisely the solutions forming the Pareto frontier for the optimal zoning

problem. This problem has been a well-defined Pareto front of nondominated solutions consisting of minimal-optimal zoning solutions (see Figures 1, 2, 3 ,4, 5 and 6).

This point can be illustrated in the following example:

Example 1: The BGAs of the metropolitan area of Toluca Valley are going to be grouped in five compact and homogeneous partitions that only include elements whose variables have values in the ranges indicated below. It is important to note that these variables are bounded in a value that is above the average:

Male Population under 6 years (X001).
Male population between 6 and 11 years (X003).
Male population between 15 and 17 (X007).
The homogeneity will be obtained on the variable X003.

We considered two different neighborhood structures and 15 iterations for the local search in the VNS.
In this example of five groups, we can see the optimal zoning minima pairs obtained as solutions with the program NODOM [2].

Table 2 shows the subset of the generated solutions accepted as non-comparables according to the proposed definition in 8 of section 2. The bold numbers in Table 2 are the solutions that match the solutions obtained with NODOM and shown in Table 1, i.e., the solutions in the table are obtained, by applying NODOM to all the solutions generated with the VNS heuristic, according to the characteristics described in example 1.

In all tables, the first column corresponds to the values of homogeneity (HOM) and the second is the compactness (COMP).

Test 1 (example 1):

| HOM | COMP |
|---|---|
| 55262 | 3256.4 |
| 37111 | 4419.6 |
| 73647 | 2162.4 |
| 94983 | 1217.2 |

Table 1. NODOM (test 1, example 1).
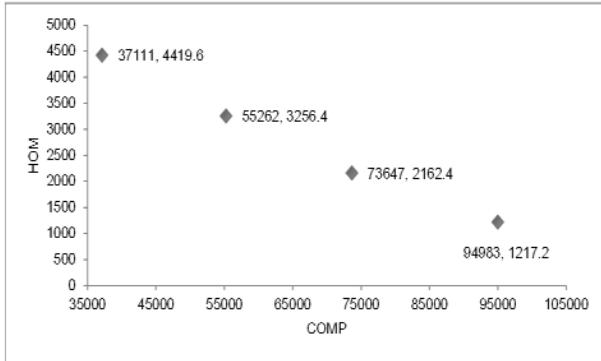


Figure 1. Pareto frontier for NODOM
(test 1, example1).

| COM | HOM |
|---|---|
| 75083 | 3184.4 |
| 42396 | 4646.8 |
| **37111** | **4419.6** |
| 45867 | 4419.6 |
| 44397 | 4419.6 |
| 65229 | 4556.4 |
| **55262** | **3256.4** |
| 57265 | 4251.6 |
| **73647** | **2162.4** |
| **94983** | **1217.2** |

Table 2. Non-comparable
(test 1, example 1.)
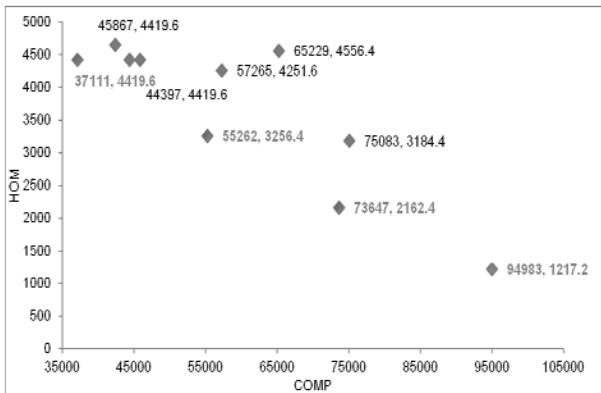


Figure 2. Pareto frontier Non-comparable
(test 1, example1).

Test 2 (example 1):

| HOM | COMP |
|------|--------|
| 66123 | 2010 |
| 30578 | 3090.7 |
| 14839 | 3250.7 |
| 37876 | 2218.7 |

Table 3. Pareto frontier for NODOM (test 2, example 1).
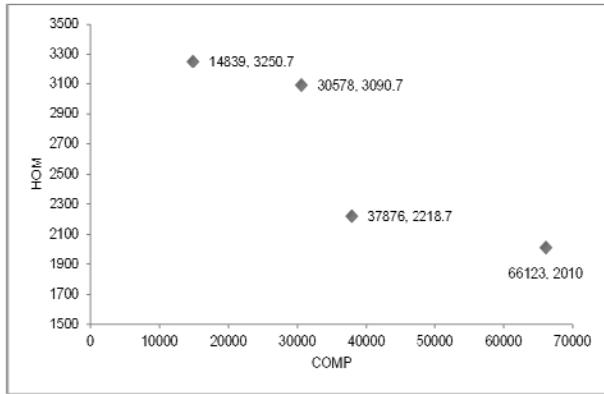


Figure 3. Pareto Frontier for NODOM
(test 2, example 1).

| HOM | COMP |
|--------|-----------|
| 53450 | 4926 |
| **66123** | **2010** |
| 50792 | 3736 |
| 65280 | 5064 |
| **30578** | **3090.667** |
| 47952 | 4792 |
| **14839** | **3250.667** |
| 28715 | 5364.6667 |
| **37876** | **2218.667** |
| 50007 | 4332.6667 |

Table 4. Non-comparable
(test 2, example 1).
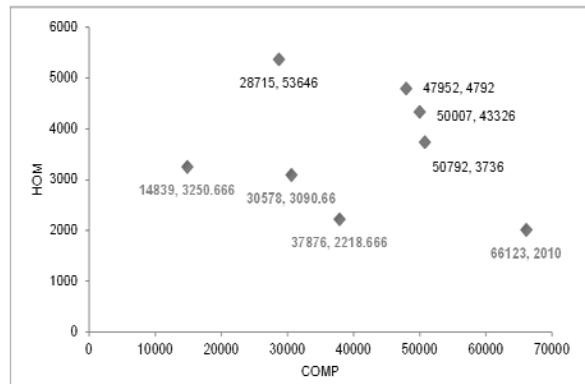


Figure 4. Pareto frontier for Non-comparable
(test 2, example 1.

Comparing Figures 3 and 4, we can see the intersection between the minima obtained with NODOM and those obtained with the Non-comparable definition proposed in this paper. However, a future contribution of this work will be to "release" some additional solutions far from the minima, which should be otherwise filtered.

Example 2: We have taken the characteristics example 1 but considering 200 different neighborhood structures and 150 iterations for the local search in the VNS and 30 groups. The homogeneity changed on to the variable male population:

| HOM | COMP |
|---------|---------|
| **257091** | **17655.3** |
| 274789 | 28974.0 |
| **288349** | **2021.5** |
| 267543 | 23452.0 |
| 282018 | 22171.3 |
| **252626** | **23854.7** |

Table 5. Pareto solutions NODOM and
Non-comparable (test 1, example 2).
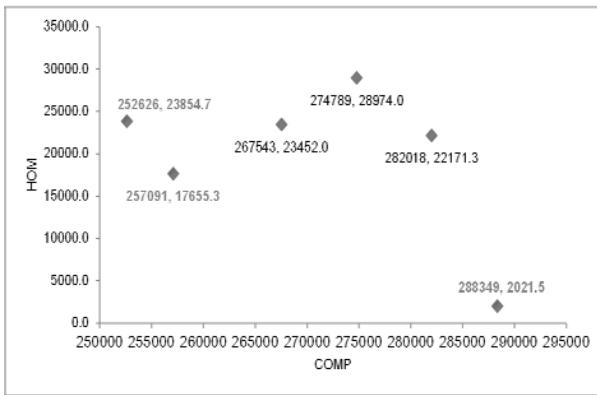
Figure 5. Pareto frontier NODOM and Non-comparable (test 1, example 2).



Figure 6. Pareto Frontier NODOM and Non-comparable (test 2, example 2).

Test 2, example 2: in this test, we consider the characteristics of the example 2 with 15 groups, 200 different neighborhood structures and 150 iterations for the local search in the VNS:

| HOM | COMP |
|---|---|
| 1764511 | 156618 |
| 1805407 | 175342 |
| **1708714** | **147318** |
| **1602015** | **243672** |
| **2026415** | **93894** |
| **1626607** | **208688** |
| 1890192 | 140800 |
| **1967572** | **122920** |
| **1917498** | **140052** |
| **1668354** | **197974** |
| **1852111** | **140558** |
| **1925789** | **137184** |
| **1920692** | **139924** |
| **1973937** | **110058** |

Table 6. Pareto solutions NODOM and Non-comparable (test 2, example 2).

## 5. **Conclusions**

One of the main contributions of this paper was the design of a partitioning model for bi-objective combinatorial optimization, where the objectives were the minimization of the compactness and the homogeneity of a particular census variable, solved with VNS. The set of solutions generated was a selected subset of the non dominated solutions that form the Pareto frontier of the problem.

The results were satisfactory and showed that the heuristic adopted to find the values of compactness and homogeneity and the method used to find minimal bi-points have been appropriate, as evidenced by the Pareto frontier obtained.

Our method can also be applied to other problems with special data, but adjusting the implementation of the particular problem.

We have chosen VNS as our search engine because of the good performance that it has shown in single-objective partitioning problems [8]. However, the comparison of VNS with other search methods is also desirable in the near future.

As part of our future work, we also intend to design a factorial experiment for the multiobjective problem and to improve the Pareto Non-comparable's algorithm in order to obtain a more accurate approximation of the Pareto frontier.

### Acknowledgements

## References

[1] M. Altman,"The Computational Complexity of Automated Redistricting: Is Automation the Answer?", *Rutgers Computer and Technology Law Journal*, vol. 23, no. 1, pp. 81-142, 1997.

[2] H. Kung, F. Luccio and F. Preparata, "On Finding the Maxima of a Set of Vectors", *Journal of the ACM* (*JACM*), vol. 22, no.4, pp. 469-476, 1975.

[3] P. Hansen and N. Mladenovic, "Variable neighbourhood search, In Fred Glover and Gary A. Kochenberger editors", *Handbook of Metaheuristics, Kluwer*, 2003, chapter 6.

[4] INEGI. (2000). Sistema para la consulta de información censal 2000, (SCINCE), XII Censo General de Población y Vivienda 2000. Retrieved from http://www.inegi.org.mx/prod_serv/contenidos/espa nol/catalogo/Default.asp?accion=2&upc=702825496371.

[5] M. Ergohtt, "Multicriteria Optimization", Springer Berlin Heidelber New York, second edition, 2005, pp. 12-42.

[6] L. A. Lara, "Un estudio de las Estrategias Evolutivas para problemas Multi-objetivo", Tesis de Maestria en Ciencias en la especialidad de Ingeniería Eléctrica Opción Computación, Cinvestav IPN, Codirectores Dr. Carlos A. Coello Coello y Dr. Alin Carsteanu, 2003, pp. 12-29.

[7] B. Bernábe, "Desarrollo de un Modelo para la Determinación de Zonificación Óptima", PhD thesis, División de Estudios de Posgrado de la Facultad de Ingeniería DIMEI, UNAM, Investigación de Operaciones, 2006, pp. 73-91.

[8] B. Bernábe, J. Espinosa, J. Ramírez, and M. A Osorio, "Statistical comparative analysis of Simulated Annealing and Variable Neighborhood Search for the Geographical Clustering Problem", *Computación y Sistemas*, vol. 42, no. 3, pp. 295-308, 2009.

[9] J. Handl, J. Knowles, "Multiobjective cluster and conglomerado validation", *In multiobjective machine learning, edited by Yaochu Jin, Springer Series on Computational Intelligence*, vol. 16, pp. 21-47, 2006.

[10] I. D. Ríos, "Sobre soluciones optimas en problemas de optimización multiobjetivo", *Trabajos de Investigación Operativa. Editor Springer Berlin/ Heidelberg*, ISSN 0213-8204, vol. 2, no. 1, pp. 49-67, 2008.

[11] NODOM http://www.cs.cinvestav.mx/~emoobook/nodom/nonodo m.html

[12] J. M. Fernandez Giangreco, "Equipo de Algoritmos Evolutivos Multiobjetivo Paralelos", (Team Algorithm of pMOEAs), Proyecto de Tesis Doctoral, Informática, Facultad de Ciencias y Tecnología, Universidad Católica Nuestra Señora de la Asunción, pp. 16-21,1993.

[13] Q. Lima, J. Baran, "Optimización de enjambre de partículas aplicada al problema del cajero viajante bi-objetivo", *Revista Iberoamericana de Inteligencia Artificial, Asociación Española para la Inteligencia Artificial Valencia España*, ISSN 1137-3601, vol. 10, no. 032, pp. 67-76, 2006