Computers, Environment and Urban Systems 39 (2013) 12–26

Contents lists available at SciVerse ScienceDirect

Computers, Environment and Urban Systems

journal homepage: www.elsevier.com/locate/compenvurbsys

A population-based iterated greedy algorithm for the delimitation and zoning of rural settlements

Juan Porta [a,*], Jorge Parapar [a], Ramón Doallo [a], Vasco Barbosa [b], Inés Santé [b], Rafael Crecente [b], Carlos Díaz [c]

[a] Computer Architecture Group, Department of Electronics and Systems, University of A Coruña, 15071 A Coruña, Spain
[b] Land Laboratory, Department of Agroforestry Engineering, University of Santiago de Compostela, 27002 Lugo, Spain
[c] Galician Agency of Rural Development, Galician Regional Government, 15897 Santiago de Compostela, Spain

ARTICLE INFO

Article history:
Received 4 July 2012
Received in revised form 28 January 2013
Accepted 29 January 2013

Keywords:
Rural settlements
GIS
Population-based iterated greedy algorithm
Java
Land-use planning

ABSTRACT

In this paper we present a Population-Based Iterated Greedy (PBIG) algorithm for delimiting and zoning rural settlements. Each cadastral plots is allocated to a category (traditional–historical, common or none) considering restrictions such as the characteristics of the existing edifications and the building density. Since the problem has multiple solutions, heuristic search algorithms, as PBIG, are a good strategy to solve it. Besides the resolution of the problem according to the requirements of the laws, our work explores also new methods of delimitation. The comparison between both types of solutions can help to improve the current methodology. The algorithm, implemented using the Java programming language and integrated into an open-source GIS software, has been tested in rural settlements with different morphological characteristics, providing adjustable solutions to the specific needs of each rural settlement.

1. Introduction

Rural settlements are located in spaces between cities and are small groupings of buildings where predominate residential land use and activities related to agriculture. Although the latter characteristic is becoming less important (Muilu & Rusanen, 2004), it has conditioned the evolution and layout of much of the current villages (Grossman & Katz, 1992). This is a population settlement model representative of the European rural areas, which presents in the North West of the Iberian Peninsula (Ferreira, Condessa, Castro e Almeida, & Pinto, 2010), and specifically in the region of Galicia, a wide variety of cases that involve complex land planning and management. However, this type of settlements and the problematic aspects that stem from it are not exclusive of Europe but common throughout the world (e.g. Feng, Wang, Wang, Li, & Zhang, 2007; Smailes & Molyneux, 1965; Mukerji, 1976; Lerise, 2000; Grossman & Katz, 1992; Stoian & Henkemans, 2000).

The oldest studies on rural settlements focused on identifying the type of spatial distribution of these settlements. For example, Dickinson (1949) distinguishes two extreme types of rural settlements, isolated farm-stead and nucleated village, and numerous gradations between both types (compact irregular village, linear village, rundling village, irregular modern growth, suburban growth, etc.), while Smailes and Molyneux (1965) classify them in dispersed settlements, pastoral agglomerations and village agglomerations. Later, these studies addressed the analysis of the functional and geometrical characteristics of rural settlements. Mukerji (1976) analyzes the morphology of rural settlements in a region of India according to the type (based on functional relationships), form (the geometrical shape of the aggregate of buildings and streets) and pattern (the geometrical arrangement of a large number of settlements suggestive of correlations with natural and cultural features). Meanwhile, Grossman and Katz (1992) identify the rural settlement patterns in Israel by building densities, field systems, physical size, and the presence or absence of detached nuclei. Recent studies seek to distinguish internal functional areas inside the rural settlement. Thus, for example, Stoian and Henkemans (2000) propose a separation between the residential area and the agricultural area in order to achieve clearer delimitations and more compact settlements. Feng et al. (2007) distinguish two types of rural settlement expansion: concentrative expansion and incompact expansion according to the value of a shape index and other characteristics. More recently Banski and Wesolowska (2010) differentiate three types of rural villages based on their residential, tourist-recreational or agricultural functional type. However, there are no studies on scientific methods or techniques for planning the delimitation of the rural settlement and zoning it in different land categories, beyond the specifications and procedures established in the corresponding laws (e.g. Lerise,

* Corresponding author.
E-mail addresses: juan.porta@udc.es (J. Porta), jparaparl@udc.es (J. Parapar), doallo@udc.es (R. Doallo), arqvascobarbosa@gmail.com (V. Barbosa), ines.sante@usc.es (I. Santé), rafael.crecente@usc.es (R. Crecente), carlos.dias.redondo@xunta.es (C. Díaz).

0198-9715/$ - see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.compenvurbsys.2013.01.006

2000) and policies (e.g., Turnock, 1991) or the method proposed by Ferreira et al. (2010) for the delimitation of consolidated urban areas in low density regions.

For this reason, the objective of this study is the development of an algorithm for the delimitation of rural settlements and the zoning of different land categories inside them. The algorithm has been designed so that the rural settlement zoning can be carried out according to the criteria established by the urban planning law of Galicia, either to more general technical criteria based on the distance between buildings, the total number of buildings, the building density rate, the total occupied land, the land suitability for development and the compactness of the delimited area. All of them are applicable to any rural settlement located anywhere. In order to clarify terms, plot is defined as a parcel of land legally defined that is owned by one or several natural or legal persons, rural settlement is an area form by plots identifiable and differentiated by official census, and zone is used in the text as synonymous of area or region.

The implemented algorithm provides valid and satisfactory solutions, that means, delimitations which comply all the restrictions and with a quality useful to the experts' needs. The characteristics of the plots (slope, orientation, land use, etc.) and the relationships between the elements of the settlement (plots, buildings, roads, etc.) are key for assigning one or another category to each plot. Each of these variables has been quantified through an analytic hierarchy process (AHP) (Bhatta & Doppler, 2010) as a multiple-criteria decision-making with the participation of twelve experts in planning processes. The MPC 2.0 software (Rodrfguez & Alboreca, 2011) was used to quantified the weights of each variable.

The rest of the paper is structured as follows: Section 2 explains the legal restrictions and the experts' recommendations that delimitations have to comply. Next, Section 3 gives an overview of the algorithm and details the pre-processing stage, where resource intensive spatial operations are executed. Section 4 explains in details the algorithm and all its phases. Finally, Section 5 shows a case study in several rural settlements and in Section 6 some conclusions are drawn.

2. Criteria for rural settlement planning

The delimitation processes are defined by several rules imposed by laws. Nevertheless, there are some criteria defined by the experience of the experts in land planning, which should be also taken into account in order to achieve acceptable solutions. In accordance with those criteria a new methodology for delimitation and zoning of rural settlements is proposed. Next sections describe the most outstanding aspects of the current laws and the proposed new methodology.

From now on, the term *building* will be used to define any construction, meanwhile the term *residential building* only will refer to constructions intended for living. In addition, a building can be traditional or modern, depending on its construction materials, height, and especially, age.

2.1. Law criteria

The current law that affects to the delimitation of rural settlements in Galicia is the *2/2010 Law of Urgent Actions of Modification of the Law 9/2002 of Urban Planning and Protection of the Rural Environment of Galicia* (Law, 2010). This law defines three different categories of rural settlements: the Traditional-Historical Rural Settlement (THRS), the Common Rural Settlement (CRS), and the Complex Rural Settlement. The last one just defines a rural settlement with THRS and CRS.

The main differences between the traditional-historical category and the common category regarding the future development are that the restrictions over the new buildings in the traditional-historical category are clearly established in the law (building materials, distances from roads, maximum height, etc.) whereas the restrictions over the common category are left to each municipality and may vary from one to another.

According to the law, a zone is considered as *consolidated* when it exceeds a certain *Building Density Rate* (BDR). Being consolidated is a necessary condition to be a rural settlement. The minimum BDR established by law is 50% for THRS and 33% for CRS. Other legal restriction is that plots further than 50 m from traditional buildings can not be part of the THRS.

One of the methods proposed by the Galician Urban Legal Protection Agency (APLU) for the calculation of the BDR of a category is based on the ratio between the number of buildable plots and the current number of buildings (Galician Urban Legal Protection Agency, 2013).

This method has been adapted according to the following equation:

$$BDR(\%) = \frac{NB}{MNBP} \tag{1}$$

where *NB* is the *Number of Buildings* and *MNBP* is the *Maximum Number of Buildable Plots*, that is calculated by:

$$MNBP = 0.8 * \frac{TAC}{MPAB} \tag{2}$$

being *TAC* the *Total Area allocated to the Category* and *MPAB* the *Minimum Plot Area for Building*, that is the minimum area for buildable plots. The factor 0.8 in Eq. (2) means that only the 80% of the total area is taken into account (the remaining 20% is an estimation of the surface of settlements usually occupied by roads, utilities networks, etc.).

2.2. Proposed alternative criteria

Besides the restrictions imposed by law, experts in land planning processes have proposed some criteria to formulate a new alternative methodology for the delimitation and zoning of rural settlements. Moreover, whereas the law refers to traditional buildings in general, in our proposed methodology, it is possible to take into account all traditional buildings or only residential traditional buildings. Following sections describe this methodology.

2.2.1. Characteristic mean distance

As aforementioned, current law indicates that plots further than 50 m from traditional buildings can not be part of a THRS. As an experimental alternative, the *Characteristic Mean Distance* (CMD) is defined as a variable distance calculated according to the morphology of the settlement and directly related to the distance between its buildings.

For calculating the CMD of a settlement, the distances between the centroids of every two buildings are computed and the CMD is the average of the *X* percent of the shortest distances, being *X* a value set by the expert in the input parameters of a preprocessing stage. Two kind of CMD are considered, traditional-historical CMD (TH-CMD) and common CMD (C-CMD) and different types of the buildings can be taken into account for the calculation: traditional residential buildings (TRB), traditional buildings (TB), residential buildings (RB) or all the buildings (B).

2.2.2. Alternative method for the calculation of the BDR

An alternative method for the calculation of the BDR is defined by the experts as follows: let *NPC* be the number of plots with buildings taken into account for calculations, and for the rest of

Table 1
Mandatory input data for the pre-processing stage.

Input data	Data type	Description
Plots	Shapefile	Set of possible plots to be included in the delimitation
Buildings	Shapefile	Buildings with the required attributes (traditional, residential)
MPAB of THRS	Double	Minimum area for buildable plots in THRS
MPAB of CRS	Double	Minimum area for buildable plots in CRS
Distances considered for the calculation of the CMD	Percentage	Percentage of distances taken into account for the calculation of the CMD
Buildings type for the calculation of the TH-CMD	String	Buildings type taken into account for the calculation of the TH-CMD (*traditional residential buildings* or *traditional buildings*)
Buildings type for the calculation of the C-CMD	String	Buildings type taken into account for the calculation of the C-CMD (*residential buildings* or *buildings marked in the shapefile)*
Minimum area of overlap	Double	Minimum area of overlap between a plot and a building to consider the building belongs to the plot

Table 2
Output of the pre-processing stage.

Data	Data type	Description
Pre-processed data	File	Results of the pre-procesing stored in a Java serialized object
Candidate plots	Shapefile	Plots not excluded from the delimitation at the pre-processing stage

plots that touch a road or are nearer than 10 m from one, let *NPN* be the number of them whose area is greater than the MPAB, and *SPM* the sum of the areas of all those plots whose area is smaller than the MPAB. Then, the building density rate is defined by the following equation:

$$BDR(\%) = \frac{NPC}{NPN + \frac{SMP}{MPAB}} \tag{3}$$

The law 2/2010 states that plots whose area is smaller than the MPAB and are located among built plots can be built. The plots whose area is smaller than the MPAB and are not in the previous case can be merged with adjacent plots to be able to be built. For these reasons, those plots and their area must be taken into account in the calculation of the BDR.

2.2.3. Maximum area of a rural settlement

The maximum area of a settlement should be limited taking into account the desirable or estimated future growth of the settlement. Considering the MPAB established by the municipal land use plan, the number of current buildings and the maximum number of new buildings that should be allowed (that is, the maximum number of future buildings), the maximum area for the delimitation of the settlement can be calculated multiplying the MPAB by the total number of buildings (current and future), with the aim of minimizing the land development.

In order to set the maximum number of new buildings that should be allowed, the following recommendations of experts are used:

- If the current number of modern residential buildings (MRB) is equal or exceeds the 50% of the RB of the settlement, the goal is to allow the construction of as many new buildings as existing MRB. For example, if there are 8 TRB and 20 MRB, the settlement must have, at maximum, the area for 20 new MRB.
- If the number of MRB is between 25% and 50% of the RB, the goal is to allow, at maximum, the construction of:
  – In the CRS, as many new buildings as existing MRB in the settlement.
  – In the THRS, 50% of the existing MRB.

For example, if there are 8 TRB and 6 MRB in the whole settlement, then the delimitation must have area for 6 new MRB in the CRS and for 3 new MRB in the THRS.

- If the number of MRB is lower than the 25% of the RB, the goal is to allow the construction of, at maximum:
  – In the CRS, as many new buildings as the 50% of the existing MRB in the settlement.
  – In the THRS, when there is CRS, 25% of the existing MRB.
  – In the THRS, when there is not CRS, 50% of the existing MRB.

For example, if there are 8 TRB and 2 MRB, the settlement must have area for, at maximum:

– If both THRS and CRS exist, 1 new MRB in the CRS (no more buildings are allowed in the THRS).
– If CRS does not exist, 1 new MRB in the THRS.

It matters that the criterion establishes differences between traditional and modern buildings, since the traditional constructions are a priority objective of the political agricultural development of the European Union (Fuentes, 2010).

2.2.4. Minimum number of buildings

The law does not set the minimum number of buildings that a zone must have in order to be considered a settlement. Usually, a minimum of two residential buildings is applied in the delimitation processes, but planning experts consider that some additional restrictions must be imposed. We will refer as polygons to the sets of neighboring plots allocated to the same category (i.e., the connected components of the two categories of the settlement). In order to classify a settlement, totally or partially, as THRS, at least one polygon with three or more traditional residential buildings should be allocated to the THRS category. Also, any polygon allocated to the THRS category must contain at least two traditional residential buildings, and any polygon allocated to the CRS category must contain at least two residential buildings.

3. Algorithm overview and pre-processing stage

To formalize our problem, we assume that the maximum extent of the rural settlement is previously limited by the experts by selecting from the cadastral plot map the set of plots candidates to be included in the delimitation of the settlement. This set of plots should be large enough to include any reasonable delimitation of the settlement, but not so large that increase the computation times too much. So, given a set of cadastral plots, the problem consists on allocating each plot to one of the two categories (THRS, CRS) or excluding it from the delimitation, maximizing a fitness function and verifying the given restrictions. This way, the number of possible solutions of the problem is C^N where C is the number of categories (three in our case) and N the number of plots.

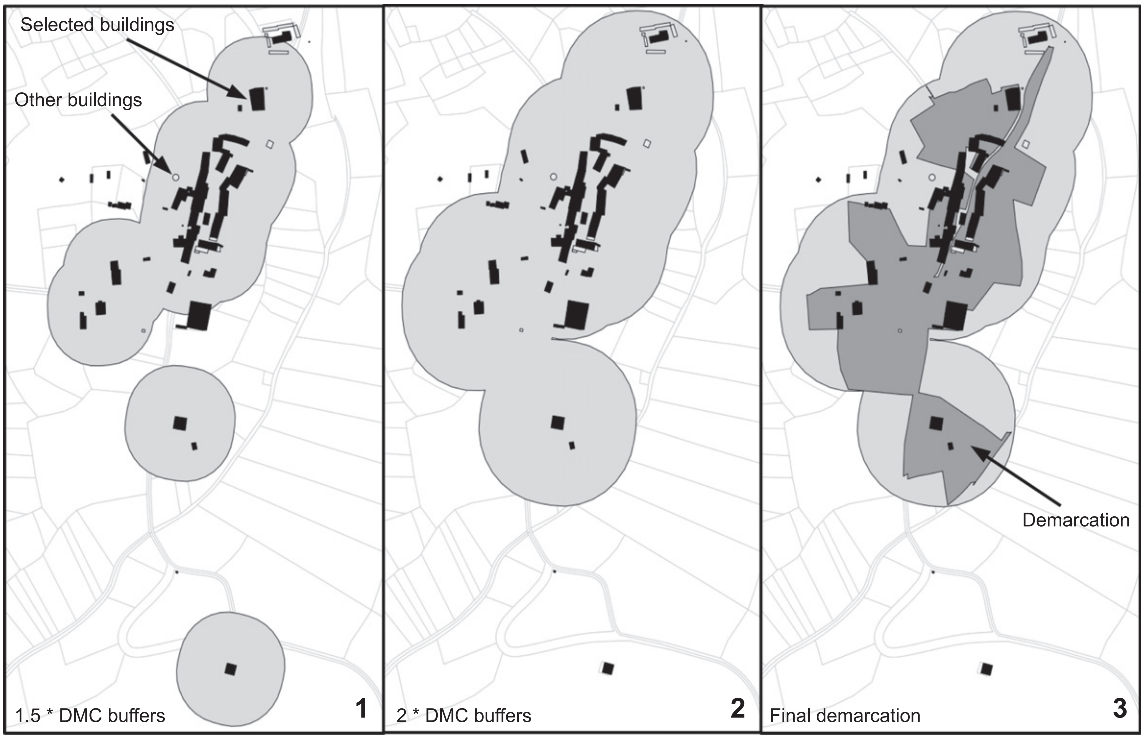

Fig. 1. Buffers around selected buildings and final demarcation.

As delimitation and zoning problems can have multiple solutions, heuristic search algorithms (Edelkamp & Schroedl, 2011) seems to fit perfectly in this kind of processes. Genetic algorithms are a type of heuristic search algorithms that are commonly applied to spatial planning for solving land use allocation problems (e.g. Balling, Taber, Brown, & Day, 1999; Cao, Huang, Wang, & Lin, 2012; Stewart, Janssen, & Herwijnen, 2004; Xin & Zhi-xia, 2008; Ferreira-Neto, Carneiro dos Santos-Junior, Fra-Paleo, Miranda-Barrós, & César de Oliveira-Moreira, 2011; Porta et al., 2013). They are based on the principles of natural evolution and use operators as selection, crossover and mutation, and the survival of fitness evolutionary analogy. Studying the characteristics of our specific problem, some of these operators as the crossover, and thus selection, make no sense: the crossover operator would provide invalid individuals with too high probability because it does not take into account the neighborhood relationships between plots, and the selection operator is not needed because the mutation is applied to all the individuals of the population.

With a population of individuals and only the mutation operator to apply, a Population-Based Iterated Greedy (PBIG) algorithm can be used instead. An iterated greedy (IG) algorithm is a heuristic search algorithm making local optimal choices at each iteration (Cormen, Leiserson, Rivest, & Stein, 2001; Neapolitan & Naimipour, 2010). The standard IG algorithms have been applied to a wide variety of problems (Ruiz & Stntzle, 2007; Pan, Wang, & Zhao, 2008; Ribas, Companys, & Tort-Martorell, 2011; Tuffery, Guyon, & Derreumaux, 2005; Benedettini, Blum, & Roli, 2010; Toyama,

Table 3
Input data for the algorithm.

Data	Data type	Description
Pre-processed data	File	Data obtained in the pre-processing stage (stored in a Java serialized object)
Population size	Integer	Number of individuals in the population
Number of iterations	Integer	Maximum number of iterations the algorithm will run
Type of distance	String	Maximum distance used. Options: *CMD* or *50 m of 2/2010 Land Law*
Minimum BDR for THRS	Percentage	Minimum building density rate for THRS. Default value: 50%
Minimum BDR for CRS	Percentage	Minimum building density rate for CRS. Default value: 33%
BDR calculation method	String	Method to use for the calculation of the BDR. Options: APLU method or alternative method
Weight for the suitability	Double	Weight in the fitness function for the suitability variable. Range: [0,1]
Weight for the THRS building ratio	Double	Weight in the fitness function for the ratio between the number of buildings included in the THRS and the number of buildings in the M-THRS. Range: [0,1]
Weight for the buildings ratio	Double	Weight in the fitness function for the ratio between the number of buildings included in the delimitation and the number of buildings in the M-RS. Range: [0,1]
Weight for the THRS area ratio	Double	Weight in the fitness function for the ratio between the area of the THRS and the area of the M-THRS. Range: [0,1]
Weight for the area ratio	Double	Weight in the fitness function for the ratio between the area of the delimitation and the area of the M-RS. Range: [0,1]
Weight for the compactness	Double	Weight in the fitness function for the compactness variable. Range: [0,1]
THRS suitability weights	File	Text file with the weights for calculating the suitability for THRS
CRS suitability weights	File	Text file with the weights for calculating the suitability for CRS

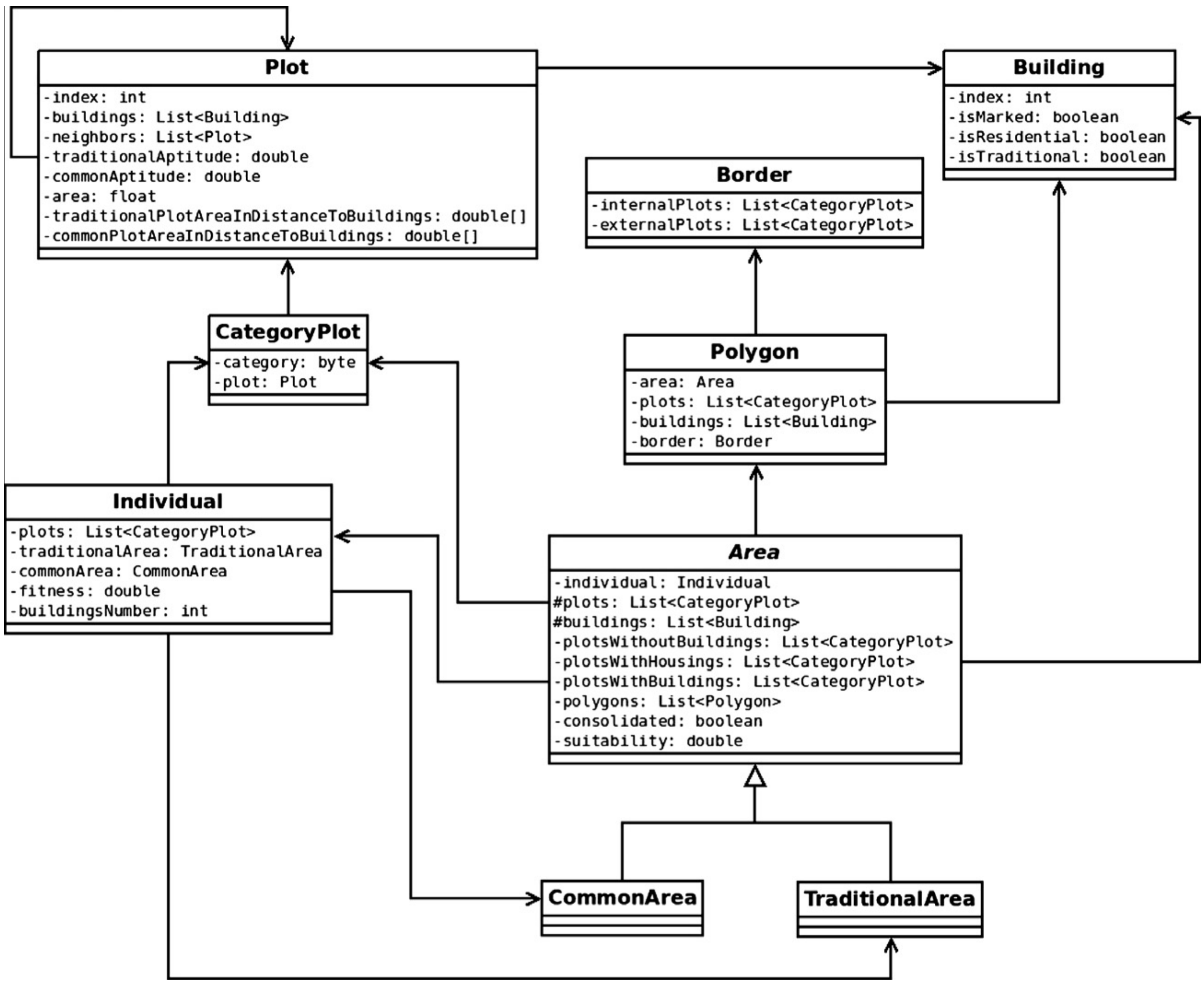

Fig. 2. Class diagram of the objects used in the algorithm.

Shoji, & Miyamichi, 2008; Lozano, Molina, & Garcý́a-Martý́nez, 2011) and only operate with one solution, but the PBIG algorithms extend that behavior using a population of solutions with the aim of improve them in a parallel way (Rodriguez, Blum, Lozano, & Garcfa-Martfnez, 2012), a technique of more recent use (Bouamama, Blum, & Boukerram, 2012; Rodriguez et al., 2012; Ballestfn, Schwindt, & Zimmermann, 2007). In the IG algorithms, the so-called *destruction-and-construction* operators is typically used to improve the solutions. This operator destroys part of a solution and then the construction phase rebuilds a complete solution (Ballestfn et al., 2007). In our case, the algorithm uses the *remove-and-add* operator (R&A op.) instead, with the aim of allocating and deallocating plots to and from the different categories to create new solutions. To keep valid the new individuals, this operator is only applied to the plots on the borders of the THRS and CRS polygons: some of the inner plots are removed from the delimitation and some of the outer plots are added and allocated to the category of the polygon they touch (see Fig. 4). From now on, the term *individual* is also used to refer to a solution belonging to the population.

Table 4
Output data of the algorithm.

Data	Data type	Description
Suitability map	Shapefile	Map of the suitability of plots
Solution	File	Final solution stored in a Java serialized object
Solution map	Shapefile	Plots of the final solution, with their category

3.1. Pre-processing stage

There are some spatial operations computationally expensive which are executed at a previous stage of the algorithm. This operations are independent of some of the input parameters so the algorithm can be executed several times with different input parameters without the need of executing those expensive operations each time.

The pre-processing stage also reduces the set of candidate plots that can be included in the rural settlement by calculating their distance to the buildings and discarding those that exceed the maximum distance allowed. The rejected plots are not passed to the algorithm and so the computation times are reduced.

3.1.1. Input and output data of the pre-processing stage

Table 1 shows the input data required for the pre-processing stage. Additional data can be also introduced in order to calculate the suitability of the plots for the THRS and CRS categories: aspects, slopes, roads, sewage and water supply networks, lightning elements, and parks and recreation areas. Table 2 shows the output of the pre-processing.

```
if (M-RS is valid) { save M-RS }                                  Validation
else if (M-RS not comply polygons restriction)
{
  for (i = 0 to size_of(pol(M − RS))) {
    if (pol(M-RS)[i] is not valid)
      remove pol(M-RS)[i]
  }
  if (M-RS is valid) { save M-RS }
}
for (i = 1 to M) { create(P_i) }                                  Initialization
iterations = 1
while (iterations ≤ max_iterations) {
  for (i = 1 to M) {
    P̃_i = P_i; tries = 1
    while (tries ≤ max_tries) {
      A = if (flip()) then THRS else CRS                          R&A op.
      if (A_ext-P̃_i is null) { rem(A-P̃_i) }
      else if (A_int-P̃_i is null) { add(A-P̃_i) }
      else if (flip()) { add(A-P̃_i) }
      else { rem(A-P̃_i) and add(A-P̃_i) }
      if (P̃_i complies restrictions &                             Replacement
            f(P̃_i) ≥ f(P_i)) {
        P'_i = P̃_i
        break;
      } // if
      tries++
    } // while
    if (tries == max_tries) P'_i = P_i
  } // for
  P = P'; iterations++
} // while
return P̂_γ where f(P̂_γ) =
    = max{f(P_0), ..., f(P_M)}
```

Fig. 3. PBIG algorithm pseudocode.

Table 5
Meaning of the pseudocode variables.

Name	Meaning
P	Population
M	Population size (number of individuals)
P_i	ith individual of the population
$A - P_i$	The part of P_i allocated to category A, where A is *THRS* or *CRS*
$A_{int} - P_i$	Internal border of the category A of P_i, where A is *THRS* or *CRS*
$A_{ext} - P_i$	External border of the category A of P_i, where A is *THRS* or *CRS*
$pol(X)$	Function that returns the polygons of a set of plots X
$f(P_i)$	Fitness function
$add(A - P_i)$	Function that adds a random plot from the external border of $A - P_i$ to the internal border of $A - P_i$
$rem(A - P_i)$	Function that removes a random plot from the internal border of $A - P_i$
$flip()$	Return *true* or *false* with a probability of 50% each one

3.1.2. Phases of the pre-processing stage

The next paragraphs describe the more important phases of the pre-processing stage. From now on, *selected buildings* will mean the buildings with the type selected in the input parameters *Buildings type for the calculation of the TH-CMD* and *Buildings type for the calculation of the C-CMD*.

Relationship between plots and buildings. Many of the operations of the algorithm need information about the number of buildings on each plot. A building can be in several plots at a time, but often the geometries of the plots and the buildings of the shapefiles are not accurate enough and some buildings erroneously overlap more than one plot providing wrong results. To avoid this situation, the area of the intersection of every overlapped plot and building is calculated and it is considered than a building is on a plot only when the area of the intersection is larger than the threshold specified in the corresponding input parameter. If several plots comply that, the building is actually in several plots. When a building do not reach the overlapping threshold with any of the plots, the building is assigned to the plot with a bigger overlapping area.

Identification of roads. The plot layer usually includes the road stretches as plots. They are identified using the land uses of the plots. But, as the geometrical relationship between both layers could be not accurate enough either, the algorithm considers that a plot is a road stretch if the overlapping of the plot with the land use polygons classified as roads exceeds 60% of the area of the plot. As land use layer is an optional input parameter, if the user does not specify it, this operation is not performed.

Calculation of the neighborhood relationships. The neighborhood relationship between plots is a fundamental information for the execution of the algorithm. The way of calculating and storing the neighborhood relationship is the same as in Porta et al. (2013) and Suárez et al. (2011), in order to achieve an optimal performance. Two plots are considered to be neighbors when the length of their boundary is larger than 0. This definition rejects elements that only touch each other in a finite number of points.

Existence of THRS and calculation of its maximum extent. As it is mentioned in Section 2.2.3, the rural settlement should not exceed a certain extent. That extent is calculated by creating buffers around the selected buildings. When the TH-CMD is used to delimit the THRS, a buffer with radius $1.5 * $ *TH-CMD* is created around each selected building. Then, all these buffers are merged together creating a new geometry composed of one or more polygons. If any of them have more than two traditional residential buildings, the algorithm concludes that the rural settlement could have a THRS part. Otherwise, it could not.

At this point, if the existence of THRS is not discarded, the next step is calculating its maximum extent. The algorithm removes the polygons which contain none or only one of the selected buildings and new buffers with radius $0.5 *$ TH-CMD are created around the rest of the polygons. The new buffers that intersect themselves are merged forming new polygons; the other new buffers are discarded and the polygons which originated them are kept.

Fig. 1 shows all those steps applied to the rural settlement of Corredoiros belonging to the Galician municipality of Guitiriz. In the first image we can observe the $1.5 *$ CMD buffers around the selected buildings. As there are some polygons with two or more selected buildings inside, the rural settlement could have a THRS part. The polygon with only one selected building inside (at the bottom of image 1) is rejected. In the second image, existing buffers are extended adding $0.5 *$ *CMD* units; then the two polygons intersect themselves, so this buffer size is kept for the final demarcation, as the third image shows.

Some plots will be totally included in some of the new polygons and others will be only partially included. These last ones must have one or more residential buildings inside the limits of some of the polygons or they will be excluded. The plots outside of the new polygons but that contain residential buildings closer than 10 m to some of the polygons are partially included by creating a buffer of 15 m around the residential building and merging it to the polygon. The resulting geometry is the maximum extent that the THRS should have, and it will be so-called *M-THRS*.

If a distance of 50 m is used instead of the CMD, the steps above would be the same but creating buffers with radius 50 m.

Existence of CRS and calculation of its maximum extent. To check the possible existence of the CRS, two scenarios are considered: the rural settlement can have THRS or it can not. When the rural settlement can not have THRS, a buffer with radius $2 *$ *C-CMD* around each of the selected buildings is created. If any of the resultant polygons has two or more residential buildings, the algorithm concludes that the settlement could have a CRS part. In that case, its maximum extent is calculated in the same way that explained above for the

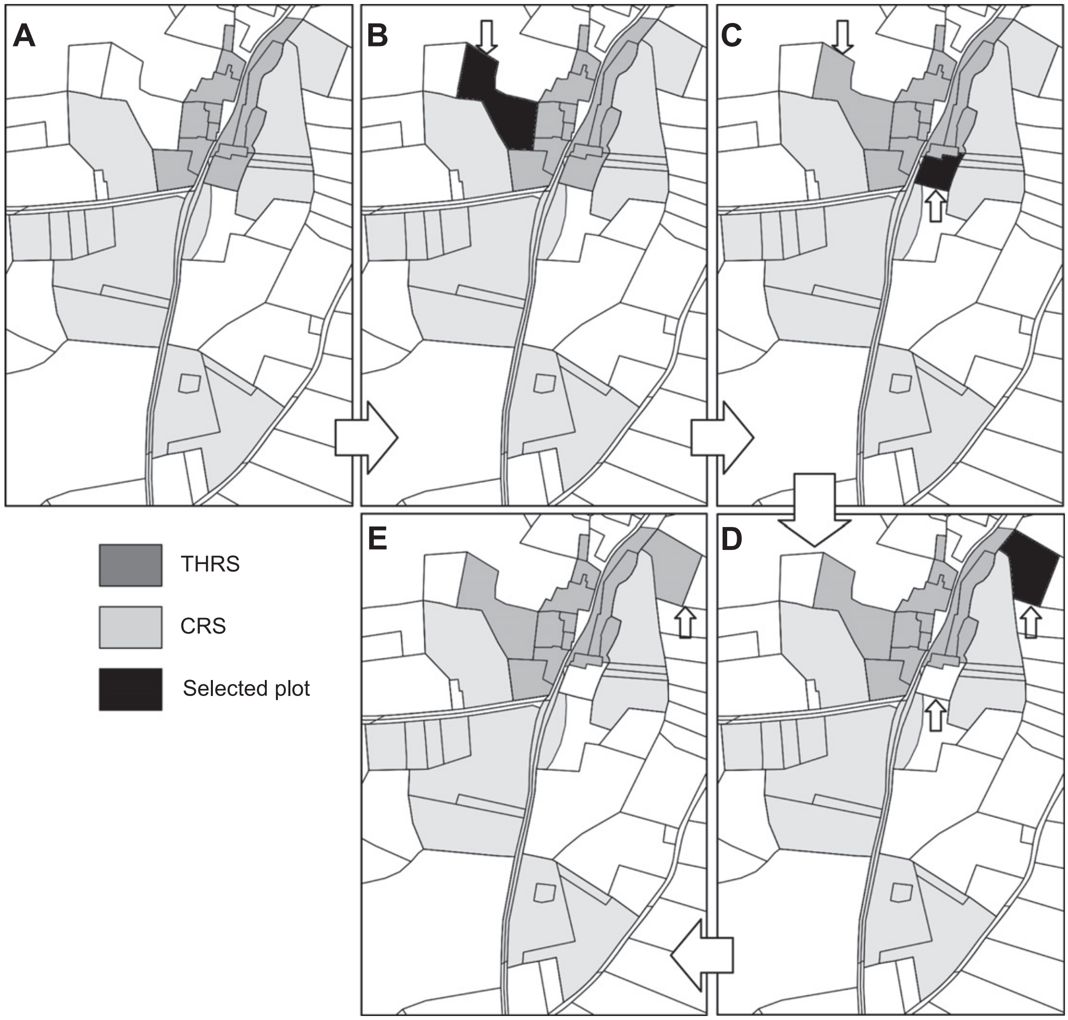

Fig. 4. R&A operator applied to an individual.

Table 6
Configuration of the test executions.

Solution	Distances	Type of buildings		Calculation of BDR
		THRS	CRS	
Legal	50 m	TB	RB	APLU method
A1	CMD (25%)	TRB	RB	Alternative method
A2	CMD (50%)	TRB	RB	Alternative method
A3	CMD (75%)	TRB	RB	Alternative method
B1	CMD (25%)	TB	RB	Alternative method
B2	CMD (50%)	TB	RB	Alternative method
B3	CMD (75%)	TB	RB	Alternative method

THRS but using the C-CMD instead of the TH-CMD. In the first scenario, the existence of CRS depends on the buildings allocated to the THRS, so it has to be calculated by the algorithm at every iteration depending on the delimitation of the THRS part. The maximum extent of the CRS will be so-called *M-CRS*. Finally, the union of the plots in M-THRS and the plots in M-CRS that are not in M-THRS is the maximum extent of the rural settlement (*M-RS*).

4. PBIG algorithm for delimitation of rural settlements

Based on the considerations of Section 3, we have developed a Populated-Based Iterated Greedy algorithm for finding satisfactory solutions to the problem of delimitation and zoning of rural settlements. In following subsections we detail the characteristics of the implementation: parameters, restrictions, fitness function and phases.

4.1. Input and output data

Table 3 shows the input data for the algorithm. All the parameters are mandatory. To simplify the programming of the algorithm, we have used an object-oriented model for representing the data of the problem, instead of using basic data types as in Porta et al. (2013). Basic data types require less computational resources but it is not a critical issue in this problem. Fig. 2 shows the class diagram. An *Individual* contains all the candidate plots (*CategoryPlot*) determined at the pre-processing stage. Each CategoryPlot stores information about the plot details (*Plot*) and its category. The set of plots allocated to the same category is an *Area*. An Area is related to its plots, buildings and polygons (connected components of the Area). Each *Polygon* is related to its Area and to the plots and buildings that it contains. They also store information about their *Borders*.

Table 4 shows the output data of the algorithm.

4.2. Restrictions and fitness function

When a new population is generated, only valid individuals are accepted. To be valid, an individual must verify all the restrictions explained in Section 2:

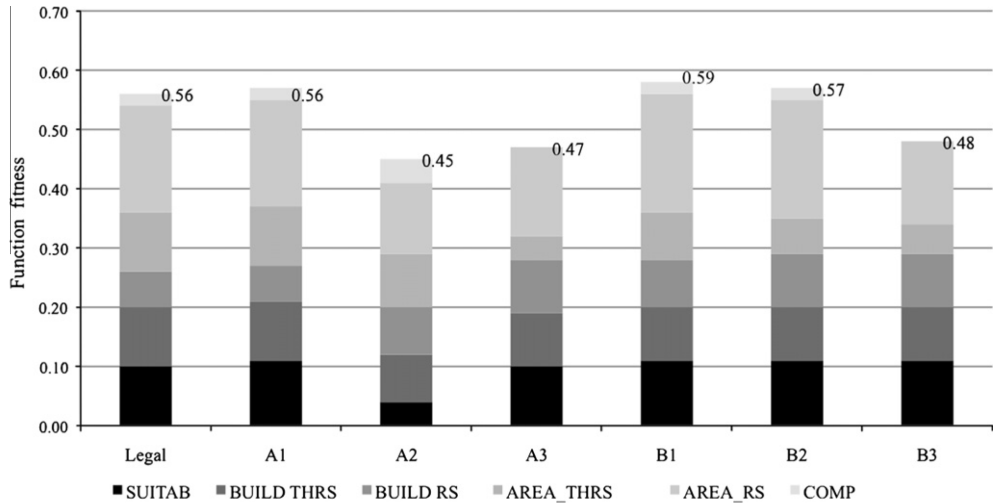

Fig. 5. Fitness values of the results for zoning of the rural settlement of Viladonega.

- *Maximum area*, where the area of the solution must be smaller than or equal to the maximum area of the settlement.
- *Minimum number of buildings*, where every polygon must contain a certain number of residential buildings.
- *BDR*, where the delimitations of the THRS and CRS categories have to be consolidated.
- *Maximum distance to buildings*, where every plot has to be nearer than CMD value or 50 m from some traditional building. All the individuals verify this restrictions because when a plot is allocated to any of the categories, the plot is chosen among the plots that verify this condition. The distance between every plot and building is computed and saved at the pre-processing stage.

The evaluation of individuals is based on the fitness function. We have defined the fitness function as a combination of six weighted variables. The normalized weights are an input parameter. The variables considered are the following ones:

- *Suitability (SUITAB)*. It can be more appropriate that a plot belongs to THRS than to CRS, or vice versa, depending on its own characteristics and the characteristics of the buildings inside it. In our case, the suitability is defined as a combination of several weighted variables. The features checked for the calculation of the plot suitability, previously grouped by territorial and collective units (Regnauld & Revell, 2007), are: buildings; land-uses; proximity to roads, squares, water supplies, among other services; slope; or aspect. Each one of these features is weighted by the experts according to its importance (Barbosa, Crecente, & SantT, 2011). These weights are input parameters. The suitability of an individual is given by the average of the suitability of the plots.
- *Compactness (COMP)*. The compactness is a measure that indicates the shape-regularity of the rural settlement. It is preferred a settlement with a regular shape than with ragged edges. In order to calculate the compactness, the algorithm applies a formula (Eq. (4)) based on the so-called circularity (Montero & Bribiesca, 2009). The more the shape of the polygons seem like a circle, the higher the value of compactness is.

$$COMP = 4\Pi \frac{\sum_{i=0}^{NP_{THRS}} \frac{area_i}{perimeter_i^2}}{NP_{THRS}} + 4\Pi \frac{\sum_{i=0}^{NP_{CRS}} \frac{area_i}{perimeter_i^2}}{NP_{CRS}} \tag{4}$$

being $area_i$ and $perimeter_i$ the area and the perimeter of the ith polygon allocated to the category THRS or to the category CRS respectively, NP_{THRS} the number of the polygons of the THRS and NP_{CRS} the number of the polygons of the CRS.

- *Buildings in the THRS (BUILD_THRS)*. Ratio between the number of buildings in the THRS of the individual and the number of buildings in the M-THRS.
- *Buildings in the whole settlement (BUILD_RS)*. Ratio between the number of buildings in the THRS and CRS of the individual and the number of buildings in the M-RS.
- *THRS area (AREA_THRS)*. Ratio between the area of the THRS of the individual and the area of the M-THRS.
- *RS area (AREA_RS)*. Ratio between the area of the THRS and CRS of the individual and the area of the M-RS.

Thus, the fitness function is defined by:

$$\begin{aligned} Fitness\ Function = {} & w_1 * SUITAB + w_2 * COMP + w_3 \\ & * BUILD\_THRS + w_4 * BUILD\_RS + w_5 \\ & * AREA\_THRS + w_6 * AREA\_RS \end{aligned} \tag{5}$$

being w_i, $i = 1–6$, the weights of each one of the variables, with $0 \leqslant w_i \leqslant 1$ and $\sum_{i=1}^{6} w_i = 1$.

4.3. Phases of the PBIG algorithm

The PBIG algorithm moves through several phases with the goal of obtaining a valid and acceptable solution. Starting from an initial population, the individuals are subject to variations in order to be improved. In the following sections those phases are described. Fig. 3 shows a pseudocode of the algorithm. The meaning of the variables and functions used is described in Table 5.

4.3.1. Validation of the M-RS

If the M-RS calculated in the pre-processing stage satisfies all the restrictions (line 1), M-RS solution is saved in a file because it will be useful to the experts since it includes all the buildings and the plots close enough to those buildings in a compact settlement delimitation (it is a compact solution because it is based on buffers around buildings). If some polygon of the M-RS does not satisfy the minimum number of buildings restriction of some polygon, the algorithm will modify the M-RS removing the non-valid polygons (lines 3 to 5). Then, if the M-RS becomes valid it is saved

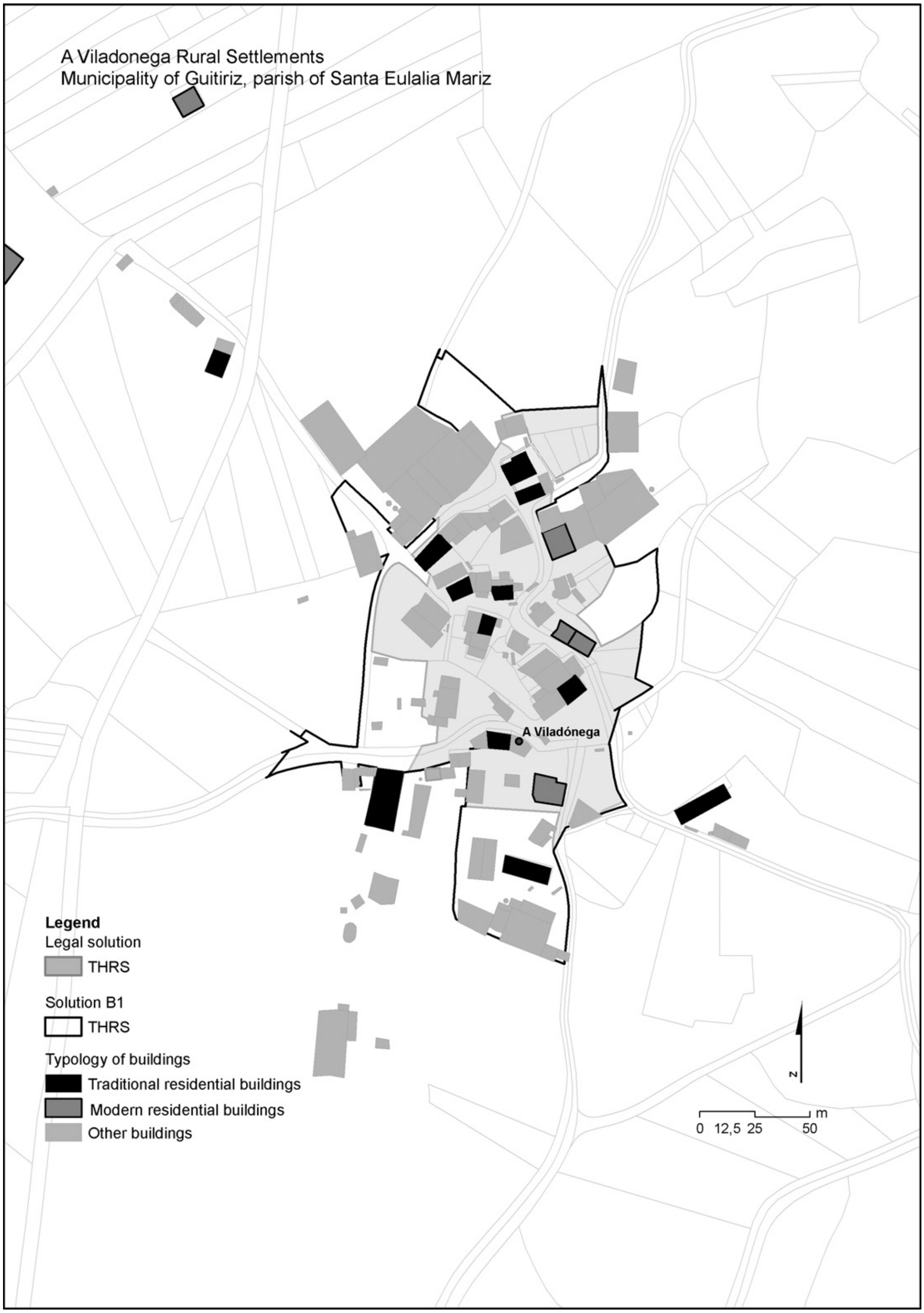

Fig. 6. Zoning of the rural settlement of Viladonega obtained by the solution according to the law and by the solution B1.

without the non-valid polygons (line 6). Otherwise the algorithm continues with the next phase.

4.3.2. Initialization of the population

The algorithm uses the M-RS as the basis for creating the initial population (line 8). In this way, the creation of the initial individuals is easier and faster, and the restrictions complied by the solutions may be exploited directly (Bouamama et al., 2012). This phase proceeds along the following steps:

- If the M-RS has THRS but it is not consolidated, the algorithm randomly removes plots from the borders of the non-consolidated polygons in the following order (until all the THRS polygons are consolidated): plots without buildings; plots with buildings which are not residential buildings; plots with residential buildings. The order in which plots are removed facilitates the consolidation of the polygons. The algorithm will stop removing plots when all the polygons are consolidated.

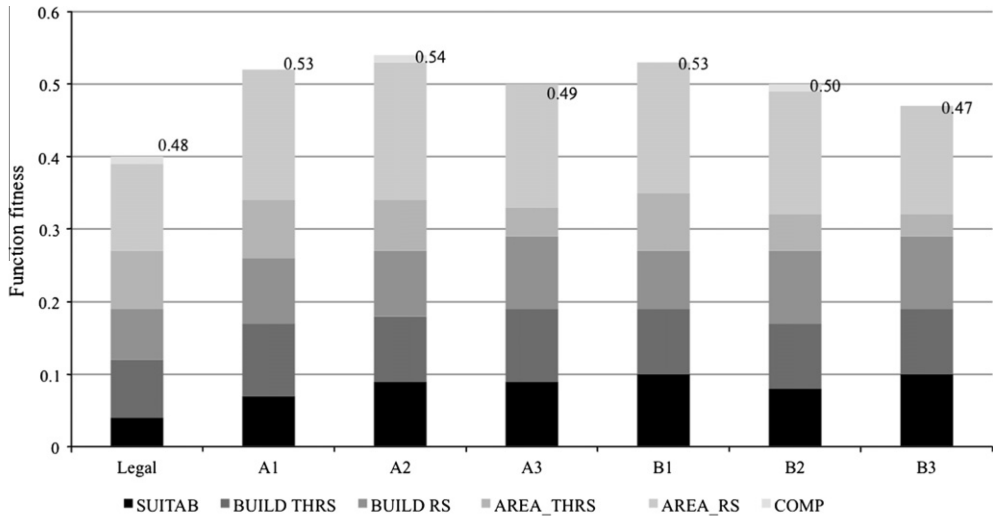

Fig. 7. Fitness values of the results for zoning of the rural settlement of Ferreira.

- If the M-RS, with the previous modifications, has CRS but it is not consolidated, the algorithm proceeds as in the previous case until all the CRS polygons are consolidated.
- If the M-RS, with the previous modifications, exceeds the maximum area, the algorithm randomly removes plots from the borders in the following order: plots without buildings in the CRS; plots without buildings in the THRS; plots without residential buildings in the CRS; plots without residential buildings in the THRS; plots with residential buildings in the CRS; plots with residential buildings in the THRS. It will stop when the area restriction is satisfied.
- If the result of all the previous modifications does not comply with the minimum number of buildings by polygon, the algorithm removes the non-valid polygons.
- All the restrictions are checked again. If any of them fails, the individual is rejected and another one is created. Otherwise, the individual becomes part of the initial population.

Note that the removed plots are selected in a random way but following the defined order. Previous steps are repeated until the necessary number of individuals are created. In the case that, after a given number of attempts, the number of individuals generated is not enough to complete the population, the algorithm stops the execution and notifies this fact to the user.

4.3.3. Remove-and-add operator

At this phase the individuals are evolved applying the R&A operator. This phase is executed as many times as set by the user.

The R&A operator is applied to each one of the individuals in the current population. We call the internal border of a polygon to the set of plots inside the polygon that touch its border, and the external border is the set of plots that touch its border but are outside it. The internal and external border of a category are the union of the internal and external borders of all the polygons allocated to it, respectively.

First, the algorithm randomly chooses one of the categories, THRS or CRS, both with a probability of 50% (line 14). If the individual does not have any plot allocated to the selected category, the individual does not suffer any modification. If the selected category has not any plot in its external border, the algorithm removes some plot from the internal one (line 15). If the selected category has not any plot in its internal border, the algorithm adds some plot from the external border (line 16). Otherwise, the selected category has plots at both external and internal borders so the algorithm randomly chooses, with a probability of 50%, between: adding a plot from the external border to the internal border, or removing a plot from the internal border and then adding another one from the external border (lines 17 to 18). This is done to try to maximize the area of the delimitation and the buildings inside. Note that the added plot must have a building close enough to satisfy the distance restriction explained in Section 2.

Fig. 4 shows two R&A operations over an individual. The original individual is represented in Fig. 4A. The first operation occurs in Fig. 4B, the selected category is the THRS and the R&A operator adds a plot from the external border to the internal border of the THRS. In Fig. 4C, the selected plot is already allocated to the THRS. The second operation consists of removing and adding: the black plot in Fig. 4C is removed from the THRS and the black plot in Fig. 4D is allocated to the THRS. Note that this last plot is *stolen* from the CRS, but the CRS is not allowed to *steal* plots from the THRS, it is only able to incorporate unallocated neighboring plots.

Once the operations are done, the algorithm checks if the new individual satisfies all the restrictions and if its fitness value is greater than the fitness value of the original individual. If both conditions are true, the new individual is included in the new population (replacement, lines 19 to 22). Otherwise, the individual is processed again. If after ten attempts the mutated individual does not satisfy those conditions, the original individual is the one selected to be part of the new population (line 25).

4.4. Final solution

The final solution is the individual with the greater fitness value in the last generation. The solution, represented by the objects defined in Section 4.1, is serialized and saved in a file. A shape file with the plots allocated to each category is also created. As it will be shown in Section 5, in all the analyzed settlements the proposed methodology always achieved better solutions than that provided by the legal criteria. This is due to the possibility of varying the values of the calculation parameters to adapt the zoning to the settlement morphology, obtaining in this way delimitations more specific for each settlement.

5. Case study

Galicia is a region of North West Spain characterized by the dispersion of population in small rural settlements. Almost half of the Spanish settlements are located in Galicia (30,091 settlements). It

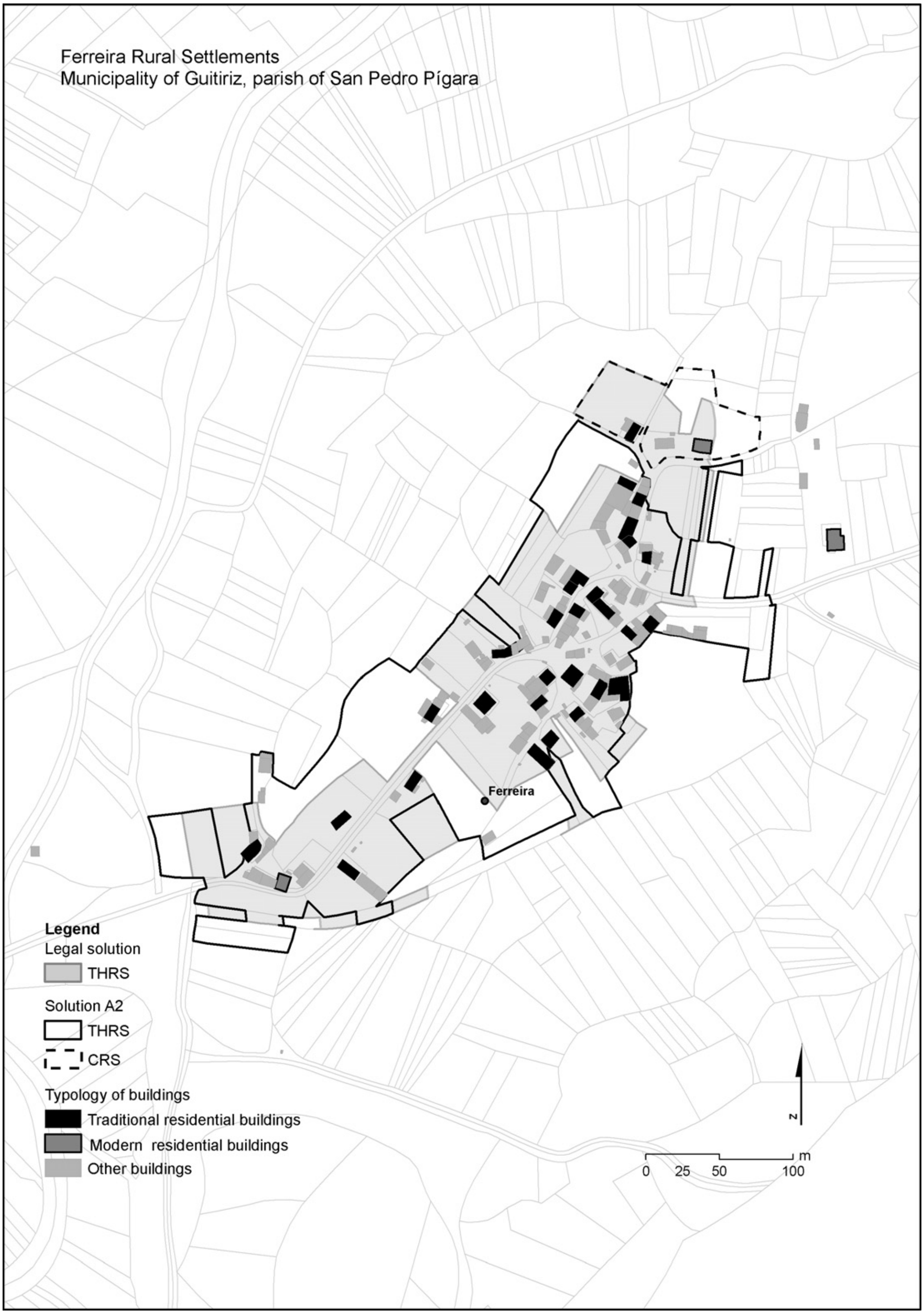

Fig. 8. Zoning of the rural settlement of Ferreira obtained by the solution according to the law and by the solution A2.

is also worth noting the fact that the 89% of Galician population settlements have a population of less than 100 inhabitants and the 16% of the Galicia population lives in settlements with less than 10 inhabitants (Enríquez & Rodríguez, 2007). This explains the importance given by the land planning law to the zoning of rural settlements, as well as the problematic aspects that stem from it.

Three rural settlements of the Galician municipality of Guitiriz (Viladonega, Ferreira and Saa) with very different morphological characteristics and spatial patterns were selected for the evaluation of the algorithm. The algorithm was run for seven sets of input parameters, the first set corresponding to the criteria specified in the planning law and the remaining ones using the alternative method for the BDR calculation, considering only traditional residential buildings (TRB) or all the traditional buildings (TB) and using different percentages of distances for the CMD calculation (Table 6). The solutions were evaluated according to the values achieved for the fitness function defined in Section 4.2. The weights used for its six variables were: 0.3 for the variable SUITAB,

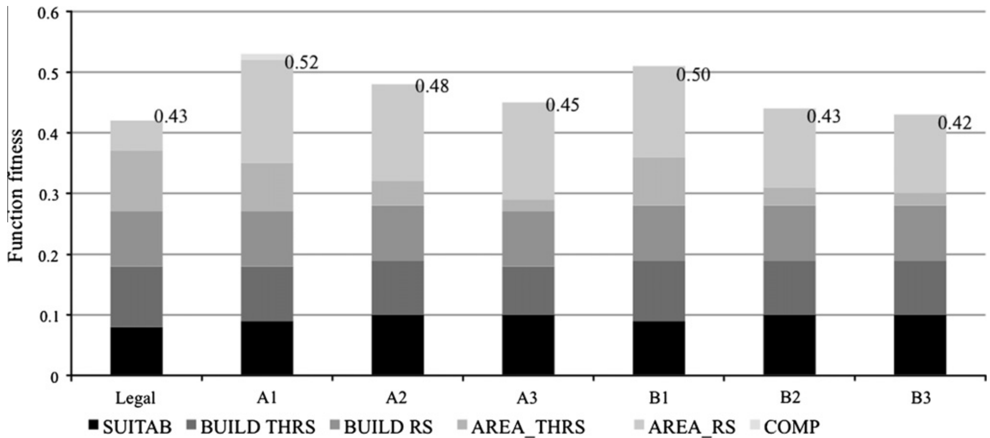

Fig. 9. Fitness values of the results for zoning of the rural settlement of Saa.

0.2 for COMP, 0.2 for AREA_RS, 0.1 for AREA_THRS, 0.1 for BUILD_THRS and 0.1 for BUILD_RS, all of them based on a previous study Barbosa et al. (2011).

The rural settlement of Viladonega is characterized by a mononuclear spatial pattern with a grouped distribution of the built environment. It has 536 plots and it is organized along a secondary road network, close to a main road axis. Results obtained for this settlement (Fig. 5) show that the solution of zoning following the legal criteria achieves a fitness value of 0.56 and the best solution of zoning following the alternative methodology is the B1, with a fitness value of 0.59. Both solutions allocate only the THRS category to the settlement. The main differences between both solutions occur in the variables BUILD_RS and AREA_RS of the fitness function. As can be observed in Fig. 6, the solution according to the law excludes one more traditional residential building than the solution B1 and allocates less area for the future growth of the settlement. The solution B1 shows a better adaptation to the morphological characteristics of the settlement since its delimitation includes a higher number of traditional residential buildings and vacant plots, as evidenced by the values of the variables BUILD_RS and AREA_RS of the fitness function. This is due to the use of only the 25% of the shortest distances between buildings for the CMD calculation. The CMD value for B1 is 31.53 m in front of the 50 m established in the law, which allows a more compact delimitation of the built environment. This can be seen also by comparing the solutions B1, B2 and B3, which show a decrease in the total fitness value when the percentage of distances used in the calculation of the CMD increases and consequently the CMD values increase (31.53 m for B1, 48.69 m for B2 and 64.94 m for B3). From this, it can be inferred that the possibility of modifying the values of the parameters used in the designed methodology allows a more effective capture of the spatial pattern of the settlement and, consequently, a better adaptation of the zoning to the specific characteristics of the settlement.

The rural settlement of Ferreira has 979 plots and a linear spatial pattern, with a distribution of the built environment organized along a road axis. Regarding to the type of buildings with residential function, the modern residential buildings are located at the ends of the settlement, mainly at the north end, while the remaining area is occupied by traditional residential buildings. The solution obtained with the legal criteria has a total fitness value of 0.48, while all the solutions obtained with the alternative methodology provide a higher fitness value (with the exception of the solution B3), being the solution A2 the one that provides the highest fitness value, 0.54 (Fig. 7). Comparing the solution according to the law and the solution A2 regarding the fitness value, the biggest differences are observed in the variables SUITAB, AREA_RS and COMP. The solution according to the law allocates the whole settlement to the THRS category, while the solution A2 divides the settlement into the THRS and CRS categories (Fig. 8). This causes that the solution according to the law presents a fractionation in terms of spatial continuity of the delimited area, while the solution A2 does not present this fractionation but a higher spatial homogeneity. That is why the COMP and AREA_RS variables have a better value. This better spatial pattern is achieved because the solution A2 was calculated using the CMD and the traditional residential buildings for the delimitation of the THRS and all the residential buildings for the delimitation of the CRS whereas the solution according to the law is calculated using a fixed distance of 50 m, in spite of the differences that exist between different areas of the settlement.

The rural settlement of Saa has 987 plots and it presents a binuclear spatial pattern with buildings distributed along a main road axis and its extension into a secondary road axis, forming two areas of spatial organization of the settlement. In both areas the dispersion of traditional residential buildings is lower than in the previously analyzed settlements, while the new residential buildings are located at the north and south ends of the settlement and are more dispersed. The zoning solution according to the legal criteria provides a total fitness value of 0.43, while, as in the previous case, all the solutions obtained with the alternative methodology provide a fitness value equal or higher (with the exception of solution B3), being in this case the solution A1 the one that provides the highest fitness value, 0.52 (Fig. 9). The biggest differences between the solution according to the law and the solution A1 are in the variable AREA_RS of the fitness function. Both solutions allocate a THRS area and a CRS area (Fig. 10), but the solution according to the law presents a division of the THRS category into two polygons, separated by the polygon of CRS, which in turn shows an irregular geometric delimitation that causes the null value of the COMP variable in the fitness function. The solution A1 presents only one polygon for the THRS category, allowing a spatial union with a regular shape between the two areas of traditional residential buildings, so in this intermediate area there are vacant plots available for future development. At the same time, a smaller CRS area is delimited at the south end of the settlement. The zoning of the solution A1 allows to consolidate the vacant space between the two cores of the settlement for future development, obtaining in this way a higher spatial cohesion, as the high value of the variable AREA_RS evidences. In addition, it can be observed in the solutions obtained with the alternative methodology that the total fitness value decreases when the percentage of distances used in the CMD calculation increases.

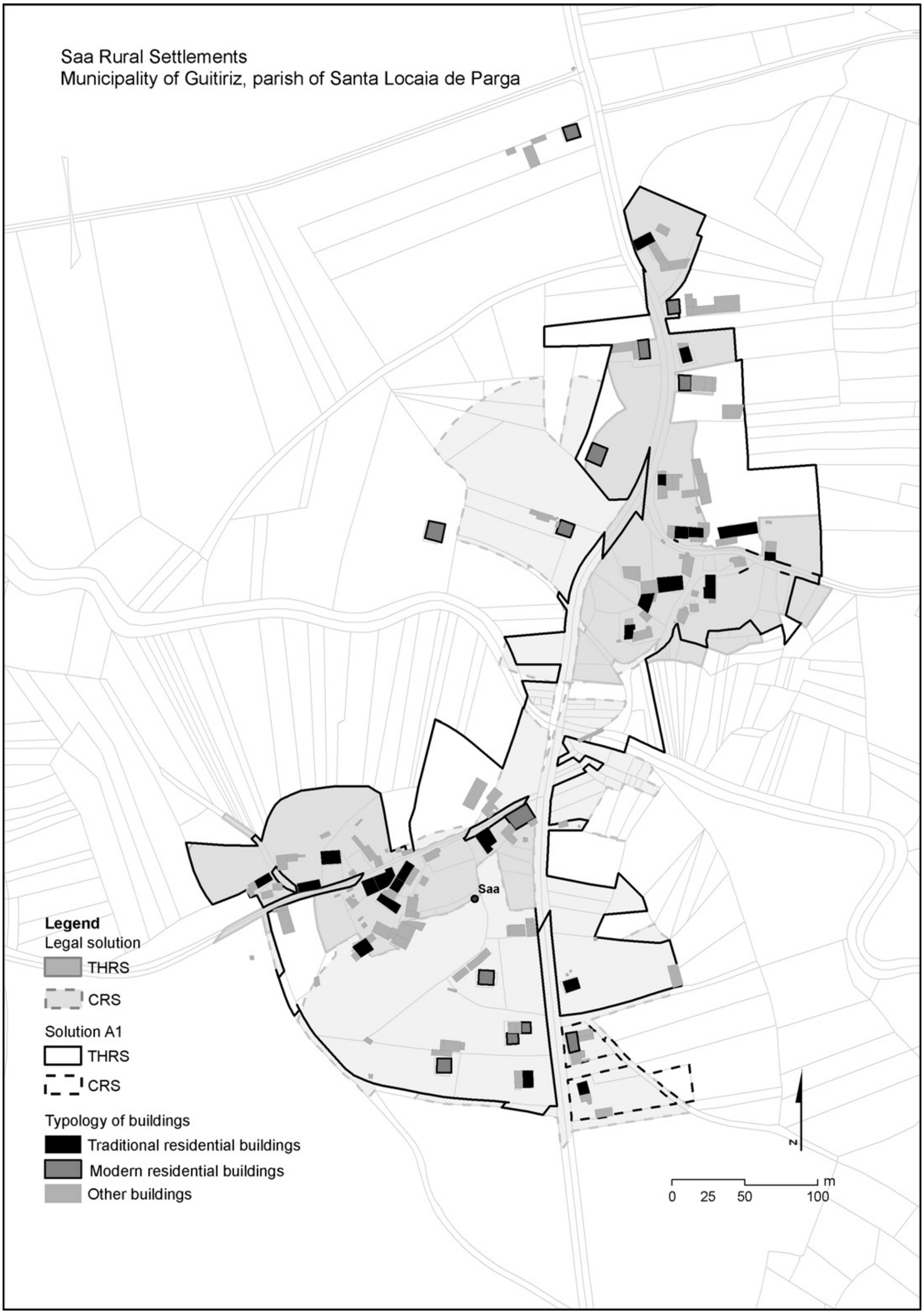

Fig. 10. Zoning of the rural settlement of Saa obtained by the solution according to the law and by the solution A1.

Regarding the behavior of the algorithm during its execution, Fig. 11 shows some examples of the evolution of the fitness values over iterations. The settlements of Ferreira, with the configuration A2, and Viladonega, with the configuration B1, have been used in these tests. Three executions have been run in each case. As Viladonega is the smallest rural settlement analized, with 536 plots, the number of plots involved in the R&A operator is smaller. This fact explains why the fitness reaches its maximum value earlier than in other settlements and all the executions get the same result.

It can be concluded that the alternative methodology can always provide a better solution than the legal criteria in terms of fitness value. This is due to the possibility of varying the values of the parameters for adjusting the zoning to the settlement morphology, obtaining in this way better spatial delimitations. The analysis of

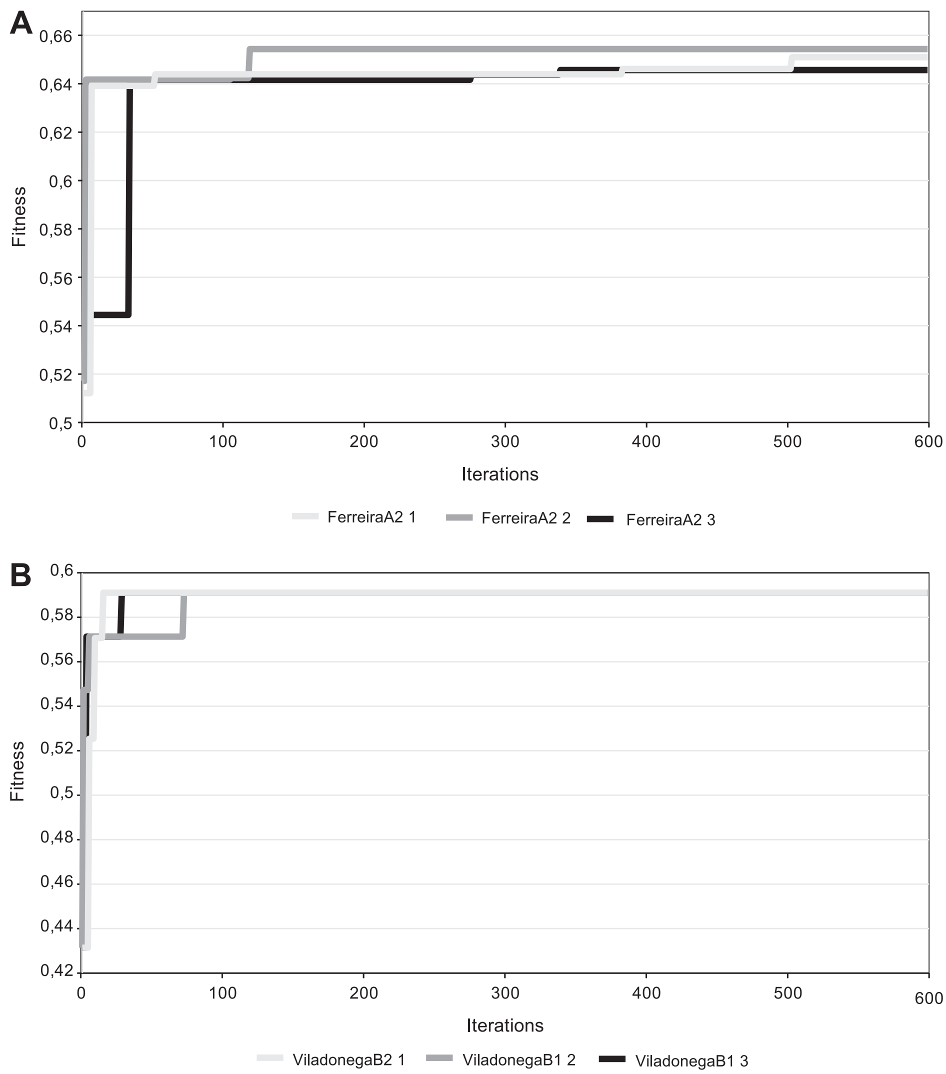

Fig. 11. Fitness evolution over iterations.

three rural settlements with very different morphologies and spatial patterns has revealed that the best solution of delimitation and zoning depends on the morphological characteristics of the settlement. Consequently, in some cases it is obtained by considering only the traditional residential buildings and in other cases considering all the traditional buildings, as well as using different percentages of distances for the CMD calculation.

The results also show zones with different categories between the solutions obtained using the legal criteria and the solutions using the alternative criteria, which demonstrates the ability of the algorithm to differentiate and capture the type of built environment, according to the type and distribution of buildings, as well as to generate delimitations better adapted to the functional characteristics of the settlement.

6. Conclusions and future work

In this paper an PBIG algorithm for the delimitation and zoning of rural settlements is proposed. The algorithm provides a tool of great potential for the generation and comparison of several zoning alternatives, following strictly the criteria established in the planning law either using a proposed alternative methodology.

The application of the algorithm to several rural settlements with very different spatial patterns has demonstrated its ability to provide solutions of delimitation and zoning adjustable to the specific needs and characteristics of the different types of rural settlements. The algorithm capability to easily delineate and zoning rural settlements allows to test multiple parameter values, which provides planners with a deeper knowledge about how the settlement morphology must be taken into account at the time of zoning the settlement. The results obtained can be used for setting a starting point, for guiding the experts, for comparing with other delimitations, etc. In any case, the algorithm does not intend to find the perfect delimitation, an expert supervision is always needed.

Regarding future work, implementing other geometric methods for the calculation of the building density rate that take into account legal restrictions for plot building, such as the plot accessibility to roads or the plot geometry, is an interesting task to do. Moreover, as the algorithm implements functions to validate and evaluate delimitations, if those methods are generalized, a generic tool for validating and evaluating delimitations of rural settlements could be created. It would allow the comparison between solutions provided by the algorithm and solutions modified by experts, in a fast and easy way. It would detect automatically which restrictions a delimitation fails (detecting it visually is usually hard).

Other interesting challenge for the algorithm is the delimitation and zoning of multiple rural settlements at once. This is not a trivial task, because settlements could have competing interests between them.

Finally, the development of a web user interface for the algorithm and to offer it as a cloud service would facilitate the use of the algorithm by the end-users since they could execute it remotely without the need of having installed it in their computers. Moreover, using OGC standards as WMS, WFS and WPS, users do not even have to store the data in their computers, but it could be obtained from remote servers. Regarding to increase the computational capabilities of the algorithm, high performance computing (HPC) can be used as in Porta et al. (2013). The HPC would help to generate more solutions (both valid and invalid) in the same period of time, extending the search space.

Acknowledgements

This work is included in the project named *Geographical Information Systems for Urban Planning and Land Management using Optimization Techniques on Multicore Processors* (Ref. 08SIN011291PR), and supported by the projects *Consolidation of Competitive Research Groups* (Ref. 2010/06), *Geospatial Information Technologies for Land Management Network* (Ref. CN2012/323), and *Strategic Groupings* (Ref. CN2012/211), all of them funded by the Galician Regional Government (Xunta de Galicia), Spain, and the latter partially supported by FEDER funds.

References

Ballestfn, F., Schwindt, C., & Zimmermann, J. (2007). Resource leveling in make-to-order production: modeling and heuristic solution method. *International Journal of Operations, 4*(1), 50–62.

Balling, R. J., Taber, J. T., Brown, M. R., & Day, K. (1999). Multiobjective urban planning using genetic algorithm. *Journal of Urban Planning and Development, 125*(2), 16–99.

Banski, J., & Wesolowska, M. (2010). Transformations in housing construction in rural areas of Poland's Lublin Region – influence on the spatial settlement structure and landscape aesthetics. *Landscape and Urban Planning, 94*(2), 116–126.

Barbosa, V., Crecente, R., & SantT, I. (2011). The rural connection as a re-construction of a territorial structure: sustainable spatial critera to delimit rural settlements following an algorithm. *The International Journal of the Constructed Environment, 2*(2), 111–129.

Benedettini, S., Blum, C., & Roli, A. (2010). A randomized iterated greedy algorithm for the founder sequence reconstruction problem. In C. Blum & R. Battiti (Eds.). *Learning and Intelligent Optimization. Lecture Notes in Computer Science* (Vol. 6073, pp. 37–51). Berlin/Heidelberg: Springer.

Bhatta, G. D., & Doppler, W. (2010). Farming differentiation in the rural-urban interface of the middle mountains, Nepal: application of analytic hierarchy process (AHP) modeling. *Journal of Agricultural Science, 2*, 37–51.

Bouamama, S., Blum, C., & Boukerram, A. (2012). A population-based iterated greedy algorithm for the minimum weight vertex cover problem. *Applied Soft Computing, 12*(6), 1632–1639.

Cao, K., Huang, B., Wang, S., & Lin, H. (2012). Sustainable land use optimization using boundary-based fast genetic algorithm. *Computers, Environment and Urban Systems, 36*, 257–269.

Cormen, T., Leiserson, C., Rivest, R., & Stein, C. (2001). *Introduction To Algorithms*. MIT Press.

Dickinson, R. E. (1949). Rural settlements in the German lands. *Annals of the Association of American Geographers, 39*(4), 239–263.

Edelkamp, S., & Schroedl, S. (2011). *Heuristic Search: Theory and Applications*. Morgan Kaufmann.

Enríquez, R. O., & Rodríguez, S. G. (2007). Urban deconcentration in the territorial system of galicia: a conceptual and empirical approach. *Boletín de la Asociación de Geógrafos Españoles, 44*, 383–386.

Feng, W., Wang, N., Wang, C., Li, G., Zhang, C., 2007. Study on characteristics of rural settlements in the northeast loess Plateau of China. In: IEEE International Geoscience and Remote Sensing Symposium IGARSS. IEEE, Barcelona, Spain.

Ferreira, J. A., Condessa, B., Castro e Almeida, J., & Pinto, P. (2010). Urban settlements delimitation in low-density areas – an application to the Municipality of Tomar (Portugal). *Landscape and Urban Planning, 97*, 156–157.

Ferreira-Neto, J. A., Carneiro dos Santos-Junior, E., Fra-Paleo, U., Miranda-Barrós, D., & César de Oliveira-Moreira, M. (2011). Optimal subdivision of land in agrarian reform projects: an analysis using genetic algorithms. *Ciencia e Investigación Agraria, 38*(2), 169–178.

Fuentes, J. (2010). Methodological bases for documenting and reusing vernacular farm architecture. *Journal of Cultural Heritage, 11*(2), 119–129.

Galician Urban Legal Protection Agency, 2013. URL: <http://www.axenciaurbanistica.es/index.php?id=100>, visited on 23 January.

Grossman, D., & Katz, Y. (1992). Rural settlement patterns in Eretz-Israel. *Geografiska Annaler. Series B, Human Geography, 74*(1), 57–73.

Law 2/2010, 2010. Galician Official Diary 61, 4,639 – 4,666, Law 2/2010, of Urgent Actions of Modification of the Law 9/2002 of Urban Planning and Protection of the Rural Environment of Galicia.

Lerise, F. (2000). Centralised spatial planning practice and land development realities in rural Tanzania. *Habitat International, 24*(2), 185–200.

Lozano, M., Molina, D., & Garcýa-Martýnez, C. (2011). Iterated greedy for the maximum diversity problem. *European Journal of Operational Research, 214*(1), 31–38.

Montero, R. S., & Bribiesca, E. (2009). State of the art of compactness and circularity measures. *International Mathematical Forum, 4*(25–28), 1305–1335.

Muilu, T., & Rusanen, J. (2004). Rural definitions and short-term dynamics in rural areas of Finland in 1989–97. *Environment and Planning A, 36*(8), 1499–1516.

Mukerji, A. (1976). Rural settlements of the Chandigarh Siwalik Hills (India): a morphogenetic analysis. *Geografiska Annaler. Series B, Human Geography, 58*(2), 95–115.

Neapolitan, R., & Naimipour, K. (2010). *Foundations of Algorithms*. Jones & Bartlett Learning.

Pan, Q.-K., Wang, L., & Zhao, B.-H. (2008). An improved iterated greedy algorithm for the no-wait flow shop scheduling problem with makespan criterion. *The International Journal of Advanced Manufacturing Technology, 38*, 778–786.

Porta, J., Parapar, J., Doallo, R., Rivera, F. F., SantT, I., & Crecente, R. (2013). High performance genetic algorithm for land use planning. *Computers, Environment and Urban Systems, 37*(0), 45–58.

Regnauld, N., & Revell, P. (2007). Automatic amalgamation of buildings for producing ordnance survey® 1:50 000 scale maps. *Cartographic Journal, 44*(3), 239–250.

Ribas, I., Companys, R., & Tort-Martorell, X. (2011). An iterated greedy algorithm for the flowshop scheduling problem with blocking. *Omega, 39*(3), 293–301.

Rodrfguez, F. P., & Alboreca, A. R. (2011). MPC 2.0©, Software for the application of AHP method in multiple-criteria decision-making. *Recursos Rurais, 7*, 27–33.

Rodriguez, F., Blum, C., Lozano, M., & Garcfa-Martfnez, C. (2012). Iterated greedy algorithms for the maximal covering location problem. In J.-K. Hao & M. Middendorf (Eds.). *Evolutionary Computation in Combinatorial Optimization. Lecture Notes in Computer Science* (Vol. 7245, pp. 172–181). Berlin/ Heidelberg: Springer.

Ruiz, R., & Sttntzle, T. (2007). A simple and effective iterated greedy algorithm for the permutation flowshop scheduling problem. *European Journal of Operational Research, 177*(3), 2033–2049.

Smailes, P. J., & Molyneux, J. (1965). The evolution of an australian rural settlement pattern: Southern New England, NSW. *Transactions of the Institute of British Geographers, 36*, 31–54.

Stewart, T. J., Janssen, R., & Herwijnen, M. (2004). A genetic algorithm approach to multiobjective land use planning. *Computers & Operations Research, 31*, 2293–2313.

Stoian, D., & Henkemans, A. B. (2000). Between extractivism and peasant agriculture: differentiation of rural settlements in the Bolivian Amazon. *International Tree Crops Journal, 10*(4), 299–320.

Suárez, M., Santé, I., Rivera, F.F., Crecente, R., Boullón, M., Porta, J., Parapar, J., Doallo, R., July 2011. A parallel algorithm based on simulated annealing for land use zoning plans. In: The 2011 International Conference on Parallel and Distributed Processing Techniques and Applications (PDPTA). Las Vegas, Nevada, EEUU, pp. 360–366, ISBN: 1-60132-193-7, 1-60132-194-5 (1-60132-195-3).

Toyama, F., Shoji, K., & Miyamichi, J. (2008). An iterated greedy algorithm for the node placement problem in bidirectional manhattan street networks. In *Proceedings of the 10th Annual Conference on Genetic and Evolutionary Computation. GECCO '08* (pp. 579–584). New York, NY, USA: ACM.

Tuffery, P., Guyon, F., & Derreumaux, P. (2005). Improved greedy algorithm for protein structure reconstruction. *Journal of Computational Chemistry, 26*(5), 506–513.

Turnock, D. (1991). The Planning of Rural Settlement in Romania. *The Geographical Journal, 157*(2), 251–264.

Xin, H., Zhi-xia, Z., 2008. Application of genetic algorithm to spatial distribution in urban planning. In: IEEE International Symposium on Knowledge Acquisition and Modeling Workshop, Wuhan, China. pp. 1026–1029.