A spatial genetic algorithm for automating land partitioning

Demetris Demetriou\textsuperscript{a}, Linda See\textsuperscript{ab} & John Stillwell\textsuperscript{a}
\textsuperscript{a} School of Geography, University of Leeds, Leeds, United Kingdom
\textsuperscript{b} International Institute for Applied Systems Analysis (IIASA), Schlossplatz 1, Laxenburg, Austria


To link to this article: http://dx.doi.org/10.1080/13658816.2013.819977
A spatial genetic algorithm for automating land partitioning

Demetris Demetriou*, Linda Seea,b and John Stillwella

School of Geography, University of Leeds, Leeds, United Kingdom; b International Institute for Applied Systems Analysis (IIASA), Schlossplatz 1, Laxenburg, Austria

(Received 11 March 2013; final version received 19 June 2013)

Land fragmentation is a widespread situation which may often hinder agricultural development. Land consolidation is considered to be the most effective land management planning approach for controlling land fragmentation and hence improving agricultural efficiency. Land partitioning is a basic process of land consolidation that involves the subdivision of land into smaller sub-spaces subject to a number of constraints. This paper explains the development of a module called LandParcelS (Land Parcelling System) that is a part of an integrated planning and decision support system called LACONISS (LAndCONsolidation Integrated Support System) which has been developed to assist land consolidation planning in Cyprus. LandParcelS automates the land partitioning process by designing and optimising land parcels in terms of their shape, size and value. The methodology integrates geographical information systems and a genetic algorithm that has been applied to two land blocks that are part of a larger case study area in Cyprus. Partitioning is treated as either a single or multi-objective problem for various optimisation cases. The results suggest that a step forward has been made in solving this complex spatial problem, although further research is needed to improve the algorithm. This approach may have relevance to other spatial planning tasks that involve single or multi-objective optimisation problems, especially those dealing with space partitioning.

Keywords: land partitioning; GIS; genetic algorithms; single and multi-objective optimisation; Thiessen polygons

1. Introduction

Land fragmentation (e.g. Bentley 1987, Van Dijk 2003) can be a problem in many countries around the world, particularly when it hinders agricultural development. The most favoured land management solution to this problem is land consolidation (e.g. Sonnenberg 2002), which consists of two main components: the reallocation of land and the provision of public infrastructure such as roads and irrigation networks. Land reallocation, which is the most critical, complex and time-consuming part of the land consolidation process (e.g. Thomas 2006, Ayranci 2007), can itself be subdivided into two sub-processes: land redistribution and land partitioning (Demetriou et al. 2012a). The first process involves restructuring the land tenure based on legislation and other related documents that set out the principles governing this restructuring, rules of thumb and the experience of the planners where these factors vary from country to country. The output of this process is a preliminary plan that divides the area to be consolidated into land blocks. Each land block, which is a sub-area

*Corresponding author. Email: demdeme@cytanet.com.cy

© 2013 Taylor & Francis
The land redistribution process is input to the second process that comprises land reallocation, i.e. land partitioning (or re-parcelling). This latter process is the focus of this paper and involves the subdivision of land into smaller ‘sub-spaces’ (i.e. land parcels) to produce a final land reallocation plan. This process is carried out one block at a time and, conventionally, is a trial and error process based on legislation, empirical design criteria, physical and technical constraints and rules of thumb. The aim of this process is to obtain a land partitioning plan with regular-shaped parcels which all have access to a road. Other constraints such as the minimum parcel size as defined by the legislation and existing physical boundaries, e.g. a stream, a river or a series of trees, should be taken into account if possible. Other technical constraints are also considered within this process such as the existence of buildings (e.g. a farmstead) or other kinds of construction (e.g. fencing).

The process of land partitioning is normally carried out by a planner and/or a designer by utilising a computer-aided design (CAD) system or a bespoke module within some other software, e.g. a geographical information system (GIS) or a surveying engineering application. This process is laborious and may take weeks or months depending upon the size and complexity of the project. Therefore, it is very unlikely that a human being could find a single optimum solution or a set of global optimum solutions if all of the problem parameters and the constraints noted above are taken into account. Two land blocks (denoted as I and II) will be used in this paper, which are taken from a case study land consolidation area in Cyprus (Demetriou et al. 2011). Figure 1 shows the subdivision of these blocks as carried out manually by land consolidation experts in the past, where the number within each parcel represents the parcel ID.

Figure 1. The subdivision of land blocks I and II as carried out by land consolidation experts.
There are only two studies in the literature that consider the development of an automated approach to land partitioning – these are studies by Buis and Vingerhoeds (1996) and Tourino et al. (2003) and both contain limitations. In particular, the former uses expert systems (ES) and GIS to produce a semi-automated module but requires the intervention of a planner. Moreover, the problem is not conceptualised as an optimisation search process and hence an optimum solution cannot be obtained using the methodology. Moreover, the GIS and the ES are loosely coupled and hence communication between them is not efficient.

The methodology in the latter study by Tourino et al. (2003) combines a region-growing algorithm (Brooks 1997) and a simulated annealing optimisation routine in a raster-based GIS for solving the land partitioning problem. However, the use of simulated annealing, as well as other classical optimisation methods, does not guarantee the global optimum solution although it has been clearly recognised that this algorithm is robust and efficient (Datta et al. 2006). In particular, classical optimisation methods, where most are point-by-point algorithms, only guarantee an optimum solution when there is a unique optimum solution, i.e. only one objective function, and hence they may adequately solve single-objective problems or multi-objective problems that are transformed into single-objective problems. In contrast, multi-objective problems with conflicting objectives involve a different optimum solution for each objective. As a result, the outcome is a set of solutions that are all optimal by varying degrees of trade-off between the objectives. Thus, the objective of global optimisation is to find the best set of solutions globally (usually for a nonlinear problem), in the presence of multiple local optima.

The results obtained by Tourino et al. (2003) were shown to be strongly influenced by the shape of the block and the size of the original parcels. Therefore, the results were only moderately good but very different from that which experts would have produced. Furthermore, raster GIS cannot provide the required accuracy for land partitioning which is inherently a vector-based problem. The authors suggest that alternative optimisation techniques should be considered and that the objective functions need improvement. Genetic algorithms (GAs) are an example of an alternative, powerful, stochastic optimisation technique that has been used for solving complex spatial optimisation problems in the past (e.g. Matthews et al. 1999, Brooks 2001, Van Dijk et al. 2002, Stewart et al. 2004, Bacao et al. 2005, Porta et al. 2013). In particular, GAs have the unique capability of producing multiple solutions at each iteration (instead of one solution in classical methods) because they use a population of solutions in each run. Hence, they may reach a set of global optima in the case of multi-objective problems. On the other hand, simulated annealing and other classical approaches are not capable of providing a set of trade-off solutions when a problem involves conflicting objectives such as land partitioning. It should be noted, however, that GAs do not always guarantee a global optimum solution but they appear to be better suited at tackling multi-objective problems with conflicting objectives.

This paper presents a new methodology for automating the land partitioning process which integrates GIS and GAs (Goldberg 1989, Openshaw and Openshaw 1997), framed both as a single and a multiple objective optimisation problem subject to a set of constraints. This methodology has been operationalised through a module called LandParcelS (Land Parcelling System) which is the final sub-system of LACONISS (LAndCONsolidation Integrated Support System) for the preparation of land reallocation plans as presented in Demetriou (2013). This new module was tested for both single- and multi-objective optimisations using two land blocks from a case study area in Cyprus as shown previously in Figure 1. The results indicate that the algorithm is heading in the right direction although further improvements are still required to produce operationally acceptable results.
2. Land partitioning as an optimisation problem

As noted earlier, the land partitioning problem involves a series of parameters and constraints. Land consolidation, as with most spatial planning processes, is a semi-structured problem that inherently involves some assumptions. Thus, the role of planners in using their judgement can be critical. As a result, solving a complex spatial problem generally involves defining some premises so as to simplify certain aspects of the problem especially during the initial stages of the problem-solving process. Then, if the first modelling efforts succeed, the complexity of the solution may gradually increase as these premises are removed. In simplifying this particular problem and the subsequent calculations, only the most important parameters are taken into account, i.e. parcel shape, parcel size, parcel land value and the road access to the land parcels. The main objective is to generate parcels with regular shapes subject to three main constraints: parcels should have access to a road and be of a predefined size and value. Deb (2001) distinguishes between hard and soft constraints. Whilst a hard constraint cannot be violated without making the solution infeasible, a soft constraint permits a range of variation within which a solution is feasible or, alternatively, a maximum variation can be specified. Thus, road accessibility is a hard constraint whereas the size and the value of each parcel are soft constraints, with an acceptable maximum variation of around ±10%. A fundamental principle from the land consolidation legislation is that each landowner shall be granted a property of an aggregate value that is the same (after deducting the land value corresponding to each landowner for public infrastructure) after land consolidation as the value of the property owned prior to the consolidation. However, it is sometimes difficult to apply this principle in practice for many reasons (FAO, 2003) and thus the acceptable variation stated above is used in practice in Cyprus.

2.1. Single-objective land partitioning

Based on the previous considerations, land partitioning can be modelled as a single-objective problem aimed at optimising the shape of parcels subject to the set of constraints noted earlier. Although other attempts at measuring shapes have been carried out (Boyce and Clark 1964, MacEachren 1985, Wentz 2000), these methods are not appropriate for evaluating rectangular shapes. In particular, area-perimeter ratios, which in essence represent the compactness of a shape and the fractal dimension, they both present significant weaknesses and are therefore not appropriate for parcel shape analysis as discussed in Demetriou et al. (2013). Thus, a new index referred to as the PSI (parcel shape index) has been developed (Demetriou et al. 2013) which outperforms existing indices. The PSI takes values between 0 and 1, representing the two extremes, i.e. the worst and the optimum parcel shape, respectively. In particular, the PSI takes into account the following six parcel shape parameters: length of sides, acute angles, reflex angles, boundary points, compactness and regularity. The application of the PSI in a case study area has shown that it offers good consistency, reliability and accuracy. The first five shape parameters are represented by value functions (Beinat 1997) which are mathematical representations of human judgement (in this case nonlinear), which are able to convert a score for each parameter and each shape (automatically calculated by a GIS) to a value between 0 and 1, representing the degree to which each parameter is close (in terms of value) to the corresponding parameter of an optimum shape. Both value functions and optimum shape have been constructed and defined by a group of land consolidation experts. In particular, an optimum shape has been defined as a rectangle with a length: breadth ratio of 2:1. The sixth parameter is represented...
by a linear transformation method. Based on this consideration, the main objective of the problem involves the following minimisation function:

\[ \min \sum_{i=1}^{N} 1 - PSI_i \]

Ideally, the above function (called F1) equals zero if all parcels of a block \((i = 1-N)\) have the optimum shape, which is a rectangle with a length:breadth ratio of 2:1. Alternatively, as shown in Demetriou et al. (2013), a shape is regular or near regular if it has a PSI of 0.7–0.9 and optimum or near optimum if the PSI is greater than 0.9.

The three problem constraints (i.e. the size, land value and the accessibility of a parcel from a road) would be easily manageable in the context of optimisation only if a mechanism for generating feasible solutions was available, i.e. solutions which do not violate any constraint. In such cases, the optimisation process only needs to find the optimum or near optimum solution in terms of parcel shape. However, this is not the case and the problem is more complicated since the generation of parcels with a predefined size and/or land value is a part of the problem. As a result, both parameters must be incorporated into the optimisation process. In other words, the two soft constraints (size and land value) can be treated as objective functions; hence land partitioning can also be treated as a multi-objective optimisation problem as outlined in the next section.

### 2.2. Multi-objective land partitioning

Land partitioning can be formulated as a multi-objective problem with the three objective functions representing shape, size and land value. Shape is represented by the function F1 noted above and the size and land value by the functions F2 and F3, respectively.

\[ \min \sum_{i=1}^{N} |dArea|_i \]

\[ \min \sum_{i=1}^{N} |dValue|_i \]

In addition, the problem is subject to the following accessibility constraint, \(R\):

\[ \sum_{i} R_i = 0 \]

In the above expressions, \(dArea\) and \(dValue\) are the percentage differences between the desired and proposed size and land value of a parcel \((i)\), respectively. The function \(R\) equals 0 or 1, respectively, when a parcel has access to a road or not. As noted earlier, according to legislation, all parcels of a land block (after land consolidation) should have access via a road. A parcel may have access from only one of its sides (which is the most common situation) or a parcel may have access from more than one road. In our case study, the external boundaries of the two land blocks utilised for running the algorithm coincide with the edge of a road. Topologically, the boundary of a parcel should coincide (touch) the
line representing the one side of the road. This is an equally hard constraint which can
be used as a penalty function. It equals the number of parcels without accessibility in a
given land block and it can be added to the overall fitness function. The use of a penalty
function in order to penalise solutions that violate one or more objectives is a popular
constraint handling strategy, although it may distort the objective function and hence lead
to a sub-optimal solution (Deb 2001).

An overall fitness function can then be generated by combining the above four func-
tions. The sum of the fitness will equal zero if all the parcels included in a land block
have an optimum shape (F1) with the desired size (F2), land value (F3) and access from a
road (R).

\[
\text{Fitness} = \left( \sum_{i=1}^{N} (1 - PSI_i)^* w_1 + \sum_{i=1}^{N} |d\text{Area}_i| * w_2 + \sum_{i=1}^{N} |d\text{Value}_i| * w_3 \right) + \sum_{i}^{N} R_i \tag{1}
\]

where \( w_1, w_2 \) and \( w_3 \) are the weights for each objective function and sum to 1.

As noted earlier, multi-objective problems with conflicting objectives involve a set
of solutions that are all optimal by varying degrees of trade-off between the objectives.
Graphically, these optimal solutions lie on a curve called the Pareto-optimal front. In par-
ticular, if all objective functions are to be minimised, this front lies close to the bottom-left
corner of the search space.

Therefore, the task in multi-objective optimisation is to find the Pareto-optimal solu-
tions which are also called non-dominated solutions because none of these solutions is
the best with respect to all objectives unless the importance of each objective can be
defined. In other words, in the case where there is confidence regarding the weights of
the objective functions, there is no reason to find other trade-off solutions (Deb 2001) and
the multi-objective problem can then be converted into a single-objective solution by utilis-
ing an appropriate vector of weights for objective functions. Weights are currently defined
empirically (and not scientifically) by land consolidation experts who are responsible for
evaluating the final land partitioning plans before their publication for inspection by the
landowners concerned. As a result, they are able to alter the weights and decide where to
assign more importance, i.e. in optimising shape, size or the land value of the parcels.
Although the definition of the weights in such models could be considered a research
topic on its own, there is a great value in the human judgement of planners who have
a considerable experience in this area and the outcome is, therefore, the result of group
decision-making and long-standing experience.

3. The land partitioning GA

A raster-based representation was originally set out, which was aimed at adaptation by
a GA (Demetriou et al. 2012b). However, the effort was abandoned early on because
the process of crossover between two raster solutions presented significant weaknesses
when executed on a pixel-by-pixel basis as explained below. First, it was inherently very
time consuming and, in addition, extra time was needed for the calculation of the various
parameters involved in the fitness function which required the conversion from raster to
vector because of the limitations of the former structure. Second, the crossover operator
was tested, and resulted in completely infeasible solutions at times in terms of the number
of parcels generated. Third, crossover resulted in parcels with nonlinear boundary sides
and thus an additional constraint would have been necessary, resulting in much more time
needed for optimisation. Fourth, the accuracy of a vector representation is much higher than the raster (centimetres vs. metres). Finally, the vector format is fundamentally the representation utilised in CAD systems where the actual task of land partitioning is normally carried out in practice. On the positive side, the advantage of using a raster representation for the particular problem is that it would be easier to reach the desired size and land value of a parcel because it provides a detailed cell-based representation of space in contrast to a vector structure that does not divide space into cells.

Thus, taking into account the above considerations, it was decided that a vector-based structure would be utilised for representing the land partitioning problem. This involves carrying out the calculations on a block-by-block basis and further reduces the complexity of the problem. Specifically, in evolutionary terms, a land block represents an individual which is evolved during the optimisation process. A land block is divided into parcels representing chromosomes. A chromosome encodes the characteristics that define an individual such as shape, size, land value and accessibility to a road. Moreover, shape is further represented through the PSI by the six aforementioned variables. Each chromosome has a core gene, namely, a centroid which is defined by a set of X and Y coordinates. Together, the individuals comprise the population. Therefore, the GA has the following hierarchical vector-based structure: population-individuals-chromosomes-genes representing, respectively: a set of subdivision solutions for a land block; one subdivision solution for a land block; land parcels and the centroids of the parcels. A graphical representation of this structure is provided in Figure 2.

Based on the above structure, the genetic process is as follows: initially, a random population of individuals is created using Thiessen polygons (or Voronoi diagrams) that divide the two-dimensional (2D) Euclidean space into a number of regions equal to the number of points provided (Dong 2008, Gong et al. 2011). In this system, the centroids are provided by another module of LACONISS called the LandSpaCES design module (Demetriou et al. 2011) which integrates GIS and ES for solving the land redistribution problem noted earlier. An example output is shown for land blocks I and II in Figure 3. Each point represents the approximate centroid location of each new parcel and the number above it denotes the landowner ID. Two parcels should be created, where double numbering is evident. The coordinates of each centroid are based on the centroids of the existing parcel(s) owned by a landowner in a certain land block according to some heuristic rules. A set of random individuals in the population is generated by moving the centroids to new, needed for optimisation. Fourth, the accuracy of a vector representation is much higher than the raster (centimetres vs. metres). Finally, the vector format is fundamentally the representation utilised in CAD systems where the actual task of land partitioning is normally carried out in practice. On the positive side, the advantage of using a raster representation for the particular problem is that it would be easier to reach the desired size and land value of a parcel because it provides a detailed cell-based representation of space in contrast to a vector structure that does not divide space into cells.

Thus, taking into account the above considerations, it was decided that a vector-based structure would be utilised for representing the land partitioning problem. This involves carrying out the calculations on a block-by-block basis and further reduces the complexity of the problem. Specifically, in evolutionary terms, a land block represents an individual which is evolved during the optimisation process. A land block is divided into parcels representing chromosomes. A chromosome encodes the characteristics that define an individual such as shape, size, land value and accessibility to a road. Moreover, shape is further represented through the PSI by the six aforementioned variables. Each chromosome has a core gene, namely, a centroid which is defined by a set of X and Y coordinates. Together, the individuals comprise the population. Therefore, the GA has the following hierarchical vector-based structure: population-individuals-chromosomes-genes representing, respectively: a set of subdivision solutions for a land block; one subdivision solution for a land block; land parcels and the centroids of the parcels. A graphical representation of this structure is provided in Figure 2.

Based on the above structure, the genetic process is as follows: initially, a random population of individuals is created using Thiessen polygons (or Voronoi diagrams) that divide the two-dimensional (2D) Euclidean space into a number of regions equal to the number of points provided (Dong 2008, Gong et al. 2011). In this system, the centroids are provided by another module of LACONISS called the LandSpaCES design module (Demetriou et al. 2011) which integrates GIS and ES for solving the land redistribution problem noted earlier. An example output is shown for land blocks I and II in Figure 3. Each point represents the approximate centroid location of each new parcel and the number above it denotes the landowner ID. Two parcels should be created, where double numbering is evident. The coordinates of each centroid are based on the centroids of the existing parcel(s) owned by a landowner in a certain land block according to some heuristic rules. A set of random individuals in the population is generated by moving the centroids to new,
random locations. An example showing six different random solutions for land block I is illustrated in Figure 4. The first three solutions are feasible (all parcels have access to a road on the polygon perimeter), whilst the other three are not. It should be noted that existing boundaries of original parcels are not the basic input since land consolidation involves a drastic restructuring of land parcels (i.e. creating larger new parcels), which is not necessarily based on the existing boundaries. However, this restricting may involve a constraint that is not yet taken into account by the algorithm.

Each random solution is then evaluated using Equation (1), which is applied at two levels: focal and zonal referring to the parcel (chromosome) and block (individual) levels, respectively. The fitness function may vary depending on the number of terms (objective functions) it contains, which then defines the land partitioning as a single or multi-objective
problem. Then an iterative process begins, where each iteration is referred to as a generation. A tournament selection method is employed (Goldberg and Deb 1991) to fill the mating pool with the same number of individuals as found in the initial population based on their fitness value.

New individuals (or offspring) are then created by applying the crossover operator to two randomly selected parent individuals from the mating pool. In particular, the BLX – a Blend Crossover operator (Deb 2001, Gwiazda 2006), which was introduced by Eshelman and Schaffer (1993), was utilised and involves the combination of genes (i.e. the X, Y coordinates of a centroid) between two corresponding chromosomes (i.e. land parcels) from two different parents. An advantage of this operator is that it may search outside of the line that connects the centroid of the two parent solutions. This operator therefore follows an adaptive search strategy, which involves searching the entire space early on while also maintaining a focused search when the population tends to convergence in some region of the search space. The way in which the operator works for the X coordinate is shown in Figure 5. Similarly, it can be applied to the Y coordinate. The crossover operator probability, \( P_c \), is set to 1.0, because an elitist operator (discussed later) is also used, which directly transfers a percentage of the best parents into the next generation.

The offspring values for \( X_{\text{new}} \) and \( Y_{\text{new}} \) are calculated based on Equations (2) and (3), respectively:

\[
X_{\text{new}} = (1 - \gamma_i) \times X_{\text{parent1}} + (\gamma_i \times X_{\text{parent2}})
\]

\[
Y_{\text{new}} = (1 - \gamma_i) \times Y_{\text{parent1}} + (\gamma_i \times Y_{\text{parent2}})
\]

where

\[
\gamma_i = (1 + 2a)u_i - a
\]

and \( u_i \) is a random number between 0 and 1, and \( a = 0.5 \).

Finally, small changes are introduced into the genetic code of an individual by mutation. Although the fitness may be worse than before mutation, the process is necessary to maintain the diversity in the population and to avoid premature convergence to local optima. Mutation involves a random change or displacement of a gene (X and Y coordinates) of a chromosome (a parcel) of an individual (land block) in a new location with probability \( P_m \), equal to 0.05. It can be applied at two levels (Delahaye 2001), namely, at the level of the parcel, i.e. in just one chromosome of an individual that is randomly selected, or at the level of a block, where all the chromosomes of an individual are randomly selected and subject to the mutation operator.

![Figure 5. The BLX – a crossover operator for the X coordinate (adapted from Deb 2001).](image-url)
The new offspring is then evaluated using the fitness measure chosen. If the termination criterion is met, then the iterative process ends and the best solution at the end of that process is returned. Otherwise, a new population is created by keeping a percentage ($e\%$) of the best individuals from the previous population to be placed directly into the new mating pool. This latter process is called an elite preserving operator and aims to speed up the convergence of the GA, thus enhancing the possibilities for creating better offspring. Although it has been proven that an elitist operator is important for the success of a GA (Rudolph 1996), attention should be paid to define the appropriate $e$ value; a commonly used value is 10%.

4. Application of the algorithm

The minimum number of parcels to be created within a land block in the employed case study area is 1 and the maximum is 26. Thus, two typical land blocks were selected (i.e. they are surrounded by roads and they have quite regular shapes) of a small and medium size so as to investigate the behaviour and the performance of the algorithm. The two land blocks, which are shown in Figure 1, reflect differing degrees of complexity of land partitioning, which is defined by the number of parcels that will be created, the size of the search space and the shape of the block. In particular, block I involves six parcels and the size is about 3 ha while block II involves 10 parcels with a size of around 5 ha. The tests that follow discuss the results of the algorithm in solving the land partitioning problem using both single and multi-objective approaches.

4.1. Shape optimisation

When partitioning is first carried out utilising the Thiessen polygons tool before optimisation, the initial subdivision is shown in Figure 6 for blocks I and II. As a result, the parcel shape depends entirely upon the layout of the centroids and, therefore, they are neither necessarily regular nor optimum. The relevant metrics for the three objective functions and the constraint for blocks I and II are the following: $F_1$ (0.264), $F_2$ (0.634), $F_3$ (0.621), $R$ (0) and
Figure 7. Subdivision of land blocks I and II by utilising the GA.

F1 (0.221), F2 (0.957), F3 (0.610), R (1.0) meaning that the solution is not feasible since the parcel with ID 159 has no access to a road.

If the land partitioning problem is then treated using single-objective optimisation, the result of the best subdivision for each block is illustrated in Figure 7. In addition to the considerable visual improvement in the parcel shapes of both land blocks, the overall fitness values for both cases are very close to zero, i.e. 0.073 and 0.019, which represents a respective improvement of 72.4% and 91.35% compared to that of the initial subdivisions. Although the algorithm fails to draw perfect rectangular parcels, which clearly depends on the external boundaries of the block and, therefore, cannot be altered; it is clear that the algorithm can successfully create polygons with regular shapes using Thiessen polygons as the starting point.

In order to undertake a more in-depth investigation into the behaviour of the algorithm and the effects of the main operators in the evolutionary process, the GA was run five times to produce five different sets of parameters denoted as Cases I to V for both land blocks. All runs involved a population size of 40 which was found to be the most appropriate in terms of convergence and computational time. It should be noted that before defining the size of the initial population to 40, some trials with a smaller population size, i.e. 20, and larger sizes, i.e. 60 and 80, were carried out. The former was too small and hence the algorithm could not converge whilst the latter increased the computational time too much. Therefore, the population size was set to 40 for all optimisation cases presented here. A similar population size has been utilised in other spatial problems as well (Krzanowski and Raper 2001, Datta et al. 2006). A detailed representation of the behaviour of the GA for blocks I and II is illustrated in Figures 8a–e and 9a–d, respectively, showing four evolutionary statistics: minimum, maximum, mean values of F1 and the overall fitness for each generation. The latter is involved only when the penalty function is added to the fitness measure. Otherwise, the overall fitness equals the mean F1. Case I for block II is not represented due to a weakness in the algorithm that means it did not converge in a reasonable time frame.

Some interesting findings can be extracted from these tests. For example, the fastest convergence time of the algorithm was 4.8 hours (Case III) for block I, which was achieved in the 18th generation and in 42 generations (or 12 hours) for block II (Case III). In both
cases, an elitist operator was utilised to speed up the process by 67.3% and 32.3%, respectively, indicating the importance of this operator. It seems that the mutation operator may have an influence on the evolutionary process in some cases, e.g. block I, but not in others, i.e. block II, depending on the particular features of the problem and the values of the other optimisation parameters. However, the mutation operator is always useful for maintaining the diversity of a population from generation to generation, especially if the crossover used does not have this ability. Another notable outcome is that when the penalty function was included in the fitness for block I (Case V), this increased the computational time fivefold compared to that achieved without the function for the best convergence time. In contrast, the penalty function was necessary for the convergence to occur in block II in order to steer the algorithm away from creating infeasible solutions.

The computational time needed to achieve the convergence is very high compared to the time a human expert could design near optimum subdivisions in terms of parcel shape. This happens for many complex problems related to spatial planning or engineering design because the evaluation of the fitness function is time consuming (Renner and Ekart 2003, Stewart et al. 2004). For example, in a multi-objective problem related to land use management (Datta et al. 2006), the algorithm needed 5000 generations and took 3.82 days to converge. In addition, the simulated annealing algorithm of Tourino et al. (2003) needed
4.2. Shape and size optimisation

The best subdivision outcomes for shape and size optimisation (F1 and F2) are presented in Figures 10a and 11a for blocks I and II, respectively, along with the corresponding sets of Pareto-optimal fronts, which are marked as a dashed line in Figure 12a and b. The latter figures include the solution with the minimum overall fitness (final population) marked as a triangle and a few other selected populations having a fitness value close to the minimum. Solutions that lie on the Pareto-optimal front are all feasible, whilst solutions that fall within the non-Pareto optimum front region can be either feasible or infeasible. Taking into account that the earlier results showed that the algorithm can satisfactorily produce regular shapes, the performance of the algorithm was tested to minimise objective F2 and, therefore, assigned (in Equation (5)) a high weight value of 0.8 for F2, a low weight of 0.2 for F1 and ignoring F3. The penalty function for infeasible solutions is also invoked.
In the case of block I, the best solution resulted in an F1 of 0.181 and in F2 of 0.094, and an overall fitness of 0.112 meaning that F1 has been improved by 31.4% and F2 by 85.2% compared to the initial subdivision. Furthermore, the shape (F1) and the size (F2) of the parcels are on average only 18.1% and 9.4% far from the optimum, respectively. These are very encouraging results because the PSI is on average very close to the optimum (0.819) and the variation of parcel size is within the acceptable range in practice (±10%), suggesting that if a guidance operator was utilised to create the parcels, then the Pareto-optimal front would be shifted even closer to the origin point of the two axes, hence to the optimum solution.

In the case of block II, the results are slightly worse compared to those of block I because of the higher complexity of the former block. In particular, the best solution has an overall fitness of 0.298, F1 of 0.089 and F2 of 0.35. This represents an improvement of 59.7% and 63.4% in F1 and F2, respectively, compared with the initial subdivision. It can also be said that the shape (F1) and the size (F2) of the parcels are on average far from the optimum by 8.9% and 35.0%, respectively. The latter outcome regarding the size of the parcels exceeds the desirable variation noted above, which emphasises the need to improve the performance of the algorithm for more complex land partitioning problems.

4.3. Shape and land value optimisation

Similar to the outcome for minimising the shape and the size of the parcels, the results for minimising the shape and the land value for land block I are encouraging. In particular,
4.4. Shape, size and land value optimisation

The best solutions for simultaneously optimising shape, size and land value of parcels for blocks I and II are shown in Figures 10c and 11c, respectively. In addition, Figure 14a and b show, in a three-dimensional (3D) plane, the projection of the set of solutions with
A set of solutions and the Pareto-optimal front for land block I.

A set of solutions and the Pareto-optimal front for land block II.

Figure 13. Shape (F1) and land value (F3) optimisation and the Pareto-optimal front for land blocks I and II.

Figure 14. A set of solutions for simultaneous optimisation of parcel shape (F1), size (F2) and land value (F3) for land blocks I (a) and II (b).

respect to the three objective functions, F1, F2 and F3 for both blocks I and II, respectively. The best solution for land block I resulted in the following metrics: fitness (0.143), F1 (0.138), F2 (0.176) and F3 (0.113), which produced an average improvement of F1, F2 and F3 by 47.7%, 72.2% and 81.8%, respectively. In other words, F1, F2 and F3 are on average...
away from the absolute optimum by 13.8%, 17.6% and 11.3%, respectively. Despite the complexity of simultaneously optimising three objectives, the results are very encouraging. It is obvious from Figure 14a that several trade-off solutions from the cloud of points representing the overall fitness (large points) are close to the origin of the three axes that reflect the optimum solution.

In the case of block II, the outcome is worse than expected although it is, in general terms, moderate. In particular, the best solution gave the following results: overall fitness (0.332), F1 (0.193), F2 (0.355) and F3 (0.378). This represents an improvement in each optimisation parameter compared with the initial subdivision of 12.7%, 62.9% and 38.0%, respectively. In other words, F1, F2 and F3 are on average 19.3%, 35.5% and 37.8% away from the absolute optimum, respectively. This result is reflected graphically in Figure 14b, where the cloud of points representing the overall fitness is quite far from the origin of the three axes that reflect the optimum solution.

The comparison of the algorithm outcomes with those given by experts shown in Figure 1 is not entirely fair since the experts took into account all of the noted constraints and allocated all of the block area whilst the algorithm did not. Moreover, the decisions taken by the LandSpaCes module (Demetriou et al. 2011) may differ from those taken by the experts although the former may be better (as proved in Demetriou et al. 2012c). In addition, the two land partitioning studies noted in the introduction lack quantitative evaluation information regarding the outputs so as to compare them with the results of the algorithm. Although results provided here are clearly encouraging, there is undoubtedly a need to improve the performance of the algorithm in terms of reaching optimum solution(s) for both single and multi-objective land partitioning, but especially for the latter case. This is due to the fact that the genotype of the algorithm involves two input variables, i.e. X and Y of the centroid of each polygon, that indirectly define optimisation parameters, i.e. shape, size and land value. As a result, the latter are not involved directly in the optimisation process. Therefore, the improvement of the performance of the algorithm can be achieved either by developing a new generic space partitioning algorithm or by introducing so-called guidance (or a learning or local optimiser) into the current optimisation process. In the former case, the algorithm should take as input parameters that the geometric features of the shapes through the PSI and the size/land value of parcels that will then be optimised through LandParcelS. In the latter case, the size and land value will be considered as constraints and the guidance operator will try to satisfy them during both initialisation and optimisation process through a kind of hill-climbing process.

5. Conclusions

The integration of GIS with GAs and multi-objective decision-making (MODM) methods for solving the land partitioning process has produced encouraging results, indicating a step forward in solving this complex spatial problem. In particular, in the case of single optimisation (i.e. the shape of the parcels), the results are near optimum for both land blocks and hence the algorithm was shown to be capable of generating polygons with regular shapes using the Thiessen polygons method. This approach overcomes the inherent limitation of generating polygons that are only based on the unique property that any location within a polygon is closer to its associated point than to the point of any other polygon.

On the other hand, in the case of multi-objective optimisation with two or three objectives, the results present a different picture depending upon the complexity of the block. In particular, for the block with the lower complexity, the outcome is fairly close to the optimum whilst for the block with the higher complexity, the outcomes are further from
the optimum in the case of size and land value. These findings suggest that, although the results are promising, further research is needed to improve the algorithm both in terms of the accuracy of the results and the computational time, and some suggestions have been provided. In addition to the evaluation of the results, several interesting findings were noted regarding the behaviour of the algorithm when the values of various operators are changed. In particular, an elitist operator is very important to speed up the process and a mutation operator may not have an influence on the evolutionary process in some cases. In addition, a penalty function considerably increases the computational time and it should be used only if it is necessary for the convergence of the algorithm.

The contribution of this research extends land partitioning and space partitioning in general, since these approaches may have relevance to other spatial processes that involve single or multi-objective problems that could be solved in the future by spatial evolutionary algorithms.

References


