Applying Genetic Algorithms to Multi-Objective Land Use Planning

Keith B. Matthews
Susan Craw
Stewart Elder
Alan R. Sibbald
Iain MacKenzie

1 Macaulay Land Use Research Institute, Craigiebuckler, Aberdeen, Scotland, AB15 8QH, Scotland, UK.
+44 1224 318611
k.matthews, a.r.sibbald @ mluri.sari.ac.uk

2 The Robert Gordon University, Aberdeen, AB25 1HG, Scotland, UK.
+44 1224 262000
s.craw, s.elder @ rgu.ac.uk

Abstract

This paper explores the application of multi-objective Genetic Algorithms (mGAs) to rural land use planning, a spatial allocation problem. Two mGAs are proposed. Both share an underlying structure of: fitness assignment using Pareto-dominance ranking, niche induction and an individual replacement strategy. They are differentiated by their representations: a fixed-length genotype composed of genes that map directly to a land parcel's use and a variable-length, order-dependent representation making allocations indirectly via a greedy algorithm. The latter representation requires additional breeding operators to be defined and post-processing of the genotype structure to identify and remove duplicate genotypes. The two mGAs are compared on a real land use planning problem and the strengths and weaknesses of the underlying framework and each representation are identified.

1 INTRODUCTION

Land use planning, particularly in rural areas, is a specialised spatial resource allocation problem. The challenge for land managers and their advisors is to formulate complex, spatially and temporally interdependent patterns of land use to achieve multiple, non-commensurable and frequently conflicting goals. It is the need to be able to generate a range of alternative land use plans, (with predictable characteristics), to support the land manager in the decision making process that has led the authors to explore the potential of genetic algorithms (GAs) for land use planning tools (Matthews et. al. 1999).

The Land Allocation Decision Support System (LADSS) (Figure 1) is being developed to allow land managers to explore and evaluate alternative land use strategies, particularly for the development of management plans to target the compensation for productivity foregone due to environmental measures adopted.

The geographic information system (GIS) provides all spatially referenced data, the spatial analysis functionality required by environmental impact assessments and the means of visualising individual allocations. The land use systems modules make assessments of the suitability, productivity and financial returns for individual parcels of land, based on bio-physical data from the GIS, global and management parameters. Data from the land use systems modules are synthesised by the impact assessments to provide assessments for the whole management unit. Financial, social and environmental impacts are considered. The GAs are the core of an iterative system, generating alternative land use plans in a context set by the land manager/advisor. This context is currently defined by the fitness metrics used to evaluate land use plans. Individual plans generated by the GAs may be visualised within the GIS as maps and specific features (for example the returns generated from a specific land parcel) queried within the land use systems modules.

Figure 1. LADSS Components

This paper presents a comparison of two land-use planning tools based on multi-objective genetic algorithms (mGAs). Following a review of related work in Section 2, the contrasting representations of the two mGAs are presented in Section 3.1, the features common to both mGAs in Section 3.2 and the representation-
specific features in Section 3.3. Section 4 presents the experimental analysis of the performance of the two mGAs, with Section 5 summarising our conclusions.

2 RELATED WORK

Three aspects of mGA design distinguish them from conventional GAs (Davis, 1991): the need to transform the vector of fitness evaluations, returned by a multiple objective fitness function, into the scalar value required by a GA selection algorithm (Van Veldhuizen and Lamont, 1998); the search for a population of solutions defining the trade-off between objectives (Goldberg, 1989) and the mGAs replacement strategy (Zitzler and Thiele, 1999).

2.1 FITNESS ASSIGNMENT

One approach to mGA fitness assignment aggregates objectives into a single value by, for example, applying weightings to individual objectives. These were rejected primarily due to their dependence on secondary information; e.g. the expense of calculating individual objective optima or the uncertainty in finding ideal weighting schemes (Srinivas and Deb, 1995, Fonseca and Fleming, 1998, Zitzler and Thiele, 1998). A further disadvantage is that if the single solution found is unacceptable for some reason outwith the scope of the optimisation then the mGA provides no other information to help the decision-maker.

mGA fitness assignment methods based on the concept of Pareto-optimality (Goldberg 1989) are more promising. From Pareto-optimality is derived the dominance relationship, with a genotype dominating another only if superior in all objectives being optimised (Fonseca and Fleming 1998). GA fitness assignment may be accomplished using only genotype ranking information (Whitley 1989) so dominance-based ranking provides a robust means of assigning the scalar fitness values required by the GA. Several mGAs differing in the detail of their dominance ranking schemes have been implemented: MOGA (Fonseca and Fleming, 1998), NPGA (Horn, et. al. 1994), NSGA (Srinivas and Deb, 1995) and SPEA (Zitzler and Thiele, 1998). The mGAs within this paper employ rank-based fitness assignment using the ranking scheme of MOGA.

2.2 NICHING

Common to all Pareto-optimal mGA implementations is the view, that, in the absence of other preference information, there is no reason to distinguish between non-dominated solutions. The goal of an mGA is thus not to find a single solution but a population composed of non-dominated genotypes evenly distributed along the Pareto-front defining the trade-off between objectives. To achieve the even distribution of the population across the front, fitness sharing (or niche induction) methods are employed (Goldberg, 1989). Niching shares (reduces) fitness values of genotypes within a neighbourhood defined by the niche size (Fonseca and Fleming 1998, Horn, et. al. 1994, Srinivas and Deb 1995).

The setting of the ratio of niche size to population size can be problematical and have a strong influence on overall performance (Zitzler and Thiele, 1998). For the mGAs presented in this paper the guidelines provided by Fonseca and Fleming were used.

There is also disagreement on the domain in which niching occurs, genotypic (parameter) or phenotypic (fitness evaluation). Phenotypic niching is criticised as it does not explicitly protect parameter diversity and in any case cannot maintain simultaneously in the population different genotypes with the same function evaluation (Srinivas and Deb, 1995). Despite this, phenotypic niching is preferred here, as it is the even spread of genotypes across this space that the decision-maker cares most about (Horn et. al. 1994).

An alternative to niche-based sharing is proposed by Zitzler and Thiele (1998). Zitzler’s SPEA maintains a fixed size, external store of all non-dominated solutions found to date. Clustering is employed, when it is necessary to create a smaller subset of genotypes while preserving the characteristics of the original set (Zitzler and Thiele, 1998). While it is a flexible and arguably superior approach to niching, the SPEA approach has not as yet been found to be necessary for the land use planning application.

2.3 REPLACEMENT STRATEGY

All the mGAs (other than SPEA) mentioned so far have employed generational replacement schemes. For single objective GAs, individual replacement is a common alternative (Whitley, 1989, Davis, 1991, Matthews et. al., 1999). While Goldberg and Deb (1991) found no evidence for the superiority of individual replacement it has been argued that the successful employment of individual replacement depends on other aspects of the GA, such as the enforcement of genotype uniqueness criteria in the population to prevent premature convergence (Davis, 1991).

Comparison of the operation of learning classifier systems (LCS) and mGAs makes the case for adopting an individual (or at least elitist) replacement strategy. The goal of mGAs is similar to that of LCS, a population of co-adapted genotypes, LCS seeking a set of rules, individually encoded as genotypes and mGAs seeking a population of non-dominated genotypes evenly sampling the trade-off front. Individual replacement is the norm in LCS as there is the recognition that a co-adapted set of rules is found incrementally with individual genotypes replaced only when they cease to be fit in the context of the current population. This elitist strategy ensures that fit individuals cannot be lost by chance, with the mGA having to rediscover them subsequently. (Valenzuela-Rendon and Uresti-Charre, 1997)

The strongest evidence for the success of the non-generational approach is that when an elitist replacement
strategy was added to existing generational replacement mGAs their performance was significantly enhanced (Zitzler et. al., 1999).

3 MULTI-OBJECTIVE GAs

To investigate the potential of mGAs as land use planning tools two mGAs have been implemented, with contrasting representations but sharing a common underlying framework.

3.1 REPRESENTATIONS

The two representations were developed in the context of single-objective GA-based land use planning tools (Matthews et. al. 1999). The first representation directly maps the land uses of individual fields as individual genes in the genotype; this is the land block representation (LB) and is illustrated in Figure 2.

Figure 2. Land Block Representation

The genes in the second representation encode target land use percentages with the priority for allocating these land uses being determined by the order in which they appear on the genotype; this is the percentage and priority representation (P&P) and is illustrated in Figure 3.

Figure 3. Percentage and Priority Representation

The P&P representation is more complex than the LB as it has both order-based and messy (Goldberg et. al, 1993) features. This is reflected in the number of operators required to support the representations (Section 3.3). The representation may however, be more readily scaled as its size depends on the number of target land use percentages (typically less than 10) rather than the number of land blocks (typically more than 100).

Previous work applied the two representations to a single objective optimisation in land use planning (Matthews et. al., 1999). Both representations found acceptable solutions. Although the P&P GA found the solutions in fewer learning cycles this was offset by the additional computational effort required per cycle, particularly that required for fitness evaluation. One question raised by the paper is: does the use of an ordering function (productivity per unit area) as part of the greedy allocation process result in a significant bias in the GA search that would make the P&P representation unsuitable for use in a mGA?

3.2 SHARED mGA FEATURES

The following section details the features common to both mGAs with Section 3.3 addressing representation specific features. The mGAs operation is outlined in Figure 4.

The mGA’s populations are fixed size and unstructured with genotype uniqueness enforced. The enforcement of uniqueness maintains a higher level of population diversity thereby reducing problems of premature convergence. Population is sized based on the niche size chosen using Fonseca and Flemings formulation (1998).

The population is randomly initialised as standard but before the reproductive cycle begins the population is doped with solutions generated from heuristics, expert knowledge, the current pattern of land use or existing single-objective optimisations. The mGAs do not depend on the doping but it does speed up the search by adding extreme points that can then be recombined with other members of the population.
The vector of fitness values returned from the impact assessment component of LADSS (Figure 1) is translated into a single selection fitness value using the count of the number of genotypes dominating an individual. The ranking is translated into selection fitness using a linear normalisation function, providing the primary means of controlling the rate of convergence. Genotypes sharing the same rank have their selection fitness values averaged.

An individual replacement strategy is used with offspring from operators being inserted singly into the population, the population is then ranked using the dominance relation, the fitness values shared and the genotype with the lowest selection fitness eliminated.

3.3 REPRESENTATION SPECIFIC FEATURES

The complexity of the P&P is reflected in the number of operators required to provide appropriate exploration of the search space. Table 1 shows the operator sets for the two representations.

<table>
<thead>
<tr>
<th>Type</th>
<th>Land-block</th>
<th>Percentage and Priority</th>
</tr>
</thead>
<tbody>
<tr>
<td>Binary</td>
<td>Uniform Crossover</td>
<td>Uniform Crossover</td>
</tr>
<tr>
<td></td>
<td>Order Prioritised Crossover</td>
<td>Splice</td>
</tr>
<tr>
<td>Unary</td>
<td>Mutation</td>
<td>Type Mutation</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Non-Uniform Mutation</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Pair-Swap</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Insert-Gene</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Delete-Gene</td>
</tr>
</tbody>
</table>

3.3.1 Binary Operators

Both representations employ uniform crossover, implemented using a crossover mask, with the crossover proportion set to 0.5 to maximise exploratory power (Syswerda, 1989). Since the P&P representation has variable genotype length, the uniform crossover is performed only between the initial segment of the genotypes common to both parents.

The P&P representation also uses an order-prioritised crossover, (OPXO), based on uniform order-based crossover (Davis, 1991) since permutations of the gene order are significant. OPXO first selects a subset of genes using a crossover mask. Subsequently the gene subset in each parent is reordered to the order of the equivalent genes in the other parent. While permutations of component genes are important to the P&P GA, only a subset of genes will be common to the parent genotypes. OPXO reorders the common genes to their order in the other parent genotype.

The final P&P binary operator is splice, which concatenates two genotypes to form a double length genotype (Goldberg et. al. 1993). This provides a simple mechanism for increasing the number of elements in a candidate solution. The unary operator cut usually complements splice, but is not used with the P&P representation as it would result primarily in partial allocations, and these would be eliminated immediately by their inability to compete with existing completely allocated members of the population. Reductions in genotype length do, however, result from the elimination of parasitic genes and gene pairs. The number of genes within genotypes may also be modified by the unary mutation style operators gene-insert and gene-delete.
3.3.2 Unary Operators

For the land-block representation a single mutation operator is required. This replaces the current value of the land use of a gene with one chosen at random from the remaining suitable land use options.

The more complex P&P representation requires operators to mutate the land use, its target percentage and its priority. The land use is mutated by changing the class of the gene while preserving its position on the genotype and its value. This is termed type-mutation. The target percentage is mutated using non-uniform mutation (Michalewicz, 1992) of the real-coded value of the gene. The range of possible mutation values is constrained by upper and lower bounds and a granularity parameter carried by the individual genes. The granularity parameter allows the accuracy of specification for candidate solutions to be matched to the ability of the LADSS to discriminate between them in its evaluation. Initial testing has used a granularity of 5%, though finer granularities will require experimentation. The mutation of priorities is achieved by swapping the location of a pair of genes on the genotype.

Two additional mutation operators change genotypes. The insert-gene operator inserts a randomly generated gene into the genotype providing a mechanism to increase the number of land uses present in the land use plan. The delete-gene operator deletes a single randomly selected gene allowing the promotion of lower-priority land use percentages “blocked” from achieving high fitness solutions by higher priority land uses.

4 EXPERIMENTAL RESULTS

The goal of the testing was to establish the relative performance of the two representations, using a typical land allocation problem. The application chosen to evaluate the mGAs performance was the optimisation of the pattern of land use for a farming research station in a disadvantaged area of the Scottish uplands. The research station comprises 90 land blocks with 9 possible land uses.

A two-objective test problem was formulated with the objectives chosen to be typical of the non-commensurable financial and environmental objectives that have to be reconciled by decision-makers. The financial fitness-metric was net present value, in £ over 60 years, and the environmental metric the Shannon-Wiener index (Forman and Godron, 1986) which measures the diversity and evenness of land use. These two metrics were chosen as it was known that for the particular application no single utopian solution, best in both objectives, would exist (a cattle mono-culture having been found to result in the optimum financial returns (Matthews et. al. 1999) while scoring zero for diversity and evenness). The mGAs ability to find a population of Pareto-optimal land allocations evenly spread across the trade-off front between the two objectives could thus be tested.

4.1 EVALUATION CRITERIA

To measure the effectiveness of the mGAs on the individual objectives two criteria were recorded: the maximum net present value for the population (MaxNPV) and the maximum value for the Shannon-Wiener index (MaxS-W).

The criteria employed to compare the multi-objective performance of mGAs are more complex than for single objective GAs with several factors influencing the quality of solutions (Zitzler et. al. 1999). As our goal is a population of non-dominated genotypes evenly spread across the trade-off front, three criteria are apparent: the fraction of the population that is non-dominated; the evenness of coverage (both Srinivas and Deb, 1995) and maximising the extent of coverage (Zitzler and Thiele, 1998). Non-dominance (NonDom) is evaluated simply by the count of the non-dominated genotypes in the population. Evenness (Even) is measured by summing for each dimension in the objective space the absolute value of the difference between the expected and actual number of genotypes per niche. Extent of coverage (Cover) is measured for two objective test problems by the union of the areas of all the rectangles formed by the origin and the co-ordinates defined by the two objectives. All these criteria are evaluated online over the course of the GA run. Finally it is useful to compare the terminal land block and P&P populations from a series of runs to establish the degree to which the populations of one representation dominate the other. For each land block population an individual dominance (IDom) score is calculated as the mean proportion of P&P populations dominated. The process is repeated for the P&P populations relative to the land block. The means of each set of IDom scores are then used as the dominance criteria (Dom).

To investigate the effectiveness of the learning, four further metrics were recorded: the number of fitness gaining events (Gain); the number of events where no fitness gain was made (NoGain); the number of events where reproduction resulted in an existing genotype (Dup) and the CPU time used in each reproductive event. The first three provide useful diagnostics indicating how efficiently the mGA operating. The CPU time metric accounts for the effort expended by the mGA and the fitness-function evaluation performed within LADSS. Both require to be taken account of as the land-block representation operates simply by looking up values within LADSS while the P&P GA is required to perform a series of more computationally intensive sorting operations in addition to the lookups.

4.2 RESULTS AND DISCUSSION

The evaluation criteria for the two mGAs are each compiled from 50 runs of the test application. Both mGAs had a niche size of 0.1 and a population size of 21. The mating distance was set to 0.25. The mGAs were terminated after 200 gaining events or 20 consecutive events without making a gain.
The mean values for the criteria are tabulated in Table 2. The differences (Diff) between the representations for each of the metrics were calculated and their significance tested (Sig). A Monte-Carlo sample-difference test (Noreen, 1989) was used, as some of the metrics are strongly skewed due to the fixed upper limit of their value.

Table 2. mGA Evaluation Criteria.

<table>
<thead>
<tr>
<th>Metric</th>
<th>Land-block</th>
<th>P&amp;P</th>
<th>Diff</th>
<th>Sig</th>
</tr>
</thead>
<tbody>
<tr>
<td>MaxNPV</td>
<td>3.415</td>
<td>3.416</td>
<td>-0.001</td>
<td>0.00</td>
</tr>
<tr>
<td>(£m)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MaxS-W</td>
<td>2.1</td>
<td>2.2</td>
<td>-0.1</td>
<td>0.00</td>
</tr>
<tr>
<td>NonDom</td>
<td>20.8</td>
<td>20.5</td>
<td>0.3</td>
<td>0.06</td>
</tr>
<tr>
<td></td>
<td>99%</td>
<td>98%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cover</td>
<td>61</td>
<td>57</td>
<td>4e6</td>
<td>0.00</td>
</tr>
<tr>
<td>(10^6)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Even</td>
<td>17</td>
<td>30</td>
<td>-13</td>
<td>0.00</td>
</tr>
<tr>
<td>Dom</td>
<td>0.47</td>
<td>0.11</td>
<td>0.37</td>
<td>0.00</td>
</tr>
<tr>
<td>All</td>
<td>453</td>
<td>561</td>
<td>-108</td>
<td>0.00</td>
</tr>
<tr>
<td>Gain</td>
<td>200</td>
<td>115</td>
<td>85</td>
<td>0.00</td>
</tr>
<tr>
<td></td>
<td>44%</td>
<td>21%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NoGain</td>
<td>248</td>
<td>232</td>
<td>16</td>
<td>0.17</td>
</tr>
<tr>
<td></td>
<td>55%</td>
<td>41%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dup</td>
<td>4</td>
<td>214</td>
<td>38%</td>
<td>0.00</td>
</tr>
<tr>
<td></td>
<td>1%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CPU (sec)</td>
<td>1042</td>
<td>1256</td>
<td>-214</td>
<td>0.00</td>
</tr>
</tbody>
</table>

For the individual fitness criteria (MaxNPV and MaxS-W) the doping has ensured that both mGA populations have near optimal genotypes (within 1% of the known optima). Between the representations the differences are small but statistically significant, with P&P outperforming land block. The difference is in the consistency with which the mGAs find the single objective optima, with the P&P representation more consistent in this respect.

In terms of the proportion of non-dominated solutions in the population (NonDom) there is no significant difference between the two representations. The replacement strategy ensures that for both mGAs it is rare for a dominated solution to remain in the population as they are replaced first. This is desirable, as it is the recombination of existing non-dominated genotypes that will ensure even coverage across the trade-off front.

For the evenness (Even) and coverage (Cover) criteria the land block representation outperforms the P&P. The terminal land-block mGA populations also dominate nearly half the genotypes of the P&P populations (Dom).

The poorer evenness performance of the P&P representation may be explained by the use of a 5% granularity for the target land use percentages of the P&P genes. While the 5% granularity was chosen as it represented a level of detail of solution desired by land managers, it may be necessary to reduce this to ensure that the even distribution of genotypes across the trade-off front is not compromised.

The superiority of the land block representation in coverage and dominance is the result of its greater ability to find intermediate solutions. This is well illustrated by plots of the terminal populations (Figures 5 and 6). In these it can be observed that while both mGAs have found allocations across the range of the trade-off front, those of the land block representation mGA are more consistently concentrated closer to the Pareto-optimal front. Reasons for this poorer performance are apparent when the criteria used to measure the effectiveness of the mGA learning are considered.
While the land block mGA is terminating due to reaching the maximum gain event limit of 200, the P&P mGA is terminating on average at gain event 115 as it has exceeded the consecutive non-gaining event limit (20). (Note that while we differentiate between no gain and duplicating events for the mGA evaluation criteria both contribute to the count of non-gaining events that can cause the mGA to terminate). While it would be possible to relax the non-gaining limit the P&P mGA would still be less efficient. The problem of excessive numbers of duplicates is the result of insufficient genetic diversity in the population due to the small size of the population. This is particularly a problem for the P&P mGA as the genotypes with high economic fitness will tend to be near monocultures represented by genotypes composed of one or two genes. Thus while the niche size may be chosen to provided an acceptable visualisation it can conflict with the operation of the mGA.

It is also possible that the P&P mGA is suffering as a result of sub-optimal parameterisation. While the LB mGA with only two operators is effective using default single objective values for parameterisation the P&P mGA with its larger number of operators may be vulnerable. In particular the effectiveness of the operators used to increase the length of the genotype needs to be established, as fitness criteria with optima defined by larger numbers of genotypes (such as Shannon-Wiener) will otherwise prove difficult to analyse.

5 CONCLUSIONS

This paper has presented two mGAs sharing an underlying structure of fitness assignment using Pareto-dominance ranking; niche induction and an individual replacement strategy. The two mGAs are differentiated by their representations, one using a direct mapping of land parcels to genes and the other making allocations indirectly using a greedy algorithm parameterised by the values held in the genotype. From analysing their performance on a typical land allocation problem it was possible to conclude that mGAs provide a useful means of establishing the structure of the trade-off between objectives. Both mGAs consistently found solutions close to the individual optima pointing to the usefulness of doping the population after initialisation with solutions known to have high fitness values for the individual objectives. The high proportion of non-dominated solutions in both mGAs populations highlighted the success of the individual replacement strategy. The poorer performance of the P&P mGA emphasises the need for care when setting niche and population sizes. Too small a population, based on a desirable niche size, may leave insufficient genetic diversity for the mGA to exploit depending on the representation adopted.

Multi-objective GAs applied to rural land use planning have significant potential for assisting land managers in tackling complex resource allocation problems with conflicting non-commensurable objectives. Their ability to search for populations of solutions defining the structure of the trade-off between objectives not only allows them to be used directly as decision support tools but may also allow land managers to become better informed about the interactions between components of their land management plans.

6 REFERENCES


