Landscape design and agricultural land-use allocation using Pareto-based multi-objective Differential Evolution

Jeroen C.J. Groot[a,b], **Walter A.H. Rossing**[a], **André Jellema**[a,c] **and Martin K. Van Ittersum**[b]

Wageningen University & Research Centre
[a] *Biological Farming Systems Group, Marijkeweg 22, 6709 PG Wageningen, The Netherlands*
[b] *Plant Production Systems Group, P.O. Box 430, 6700 AK Wageningen, The Netherlands*
[c] *Alterra Green World Research, P.O. Box 47, 6700 AA Wageningen, The Netherlands.*
E-mail: jeroen.groot@wur.nl

Abstract
Intensively managed agricultural areas in North-Western Europe are undergoing a shift from solely production oriented use to provision of multiple services and functions. Design of multifunctional agricultural landscapes can be supported by exploration of the potential to effectively combine economic performance, landscape identity, nature conservation and environmental quality. The Landscape IMAGES methodology enables spatially explicit exploration of options by multi-objective optimization, for multifunctional agriculture in landscapes at a scale of a few km². The framework has been developed to support stakeholder discussions and informed decision making. For simultaneous optimization of multiple objectives the evolutionary algorithm of Differential Evolution is employed. Selection pressure normal to the trade-off surface is exerted by Pareto-based ranking, while a crowding metric is used to provide tangential selection pressure. A large range of alternative configurations of a landscape representing the trade-off surface between the objectives was generated and explicit insight in the trade-off between the objectives was provided. Enriching the initial population of the optimization with extremes obtained from single objective optimizations resulted in an improvement of the quality of the obtained non-dominated solution set. A landscape visualization module enables rapid assessment of alternative landscape and land-use designs. In this paper, the methodology is elaborated and its functioning is illustrated with a hypothetical example of a grassland-based landscape with hedge rows bordering the fields.

Keywords: multifunctional agriculture; stakeholder discussions; optimization; scales.

1. INTRODUCTION

Over the last two decades, attention in policy, land-use planning and research directed at intensively managed agricultural areas has shifted from production to provision of multiple services and functions by agriculture. Examples of multifunctional land-use aims are maintenance or improvement of landscape structure, sustainable management of renewable natural resources, preservation of biodiversity and contribution to socio-economic viability of rural areas (OECD, 2001).

The required adjustments and innovation in landscapes and land-use systems can be characterized as complex, uncertain and value-laden issues, affecting various stakeholders. Therefore, systems approaches that integrate various issues, stakes of social actors, disciplines and scales are indispensable, and could be supported by methodologies and models to inform stakeholders and policymakers, by designing alternatives and by exploring scenarios for the future.

Existing spatially explicit, future-oriented land-use exploration approaches applied to agricultural landscapes dominated by cropping or grassland systems have focussed primarily on agro-ecological aspects of production, hydrology and nutrient loss abatement (e.g., O'Callaghan, 1995; Seppelt and Voinov, 2002; Matthews et al., 2006). Approaches for combined optimization of agricultural land-use and landscape elements configuration to improve habitat quality and nature conservation value are scarce.

Multi-objective optimization methods can be employed when there is a problem that incorporates objectives that conflict and trade-off must be accepted in compromise solutions (Anderson et al., 2005). The use of these techniques enables simultaneous optimization of multiple objectives without weighing or normalization. The dimensions can be expressed in their own units, and monetarisation of noneconomic functions can be avoided. In land-use exploration the decision variables in the optimization are the land-use options that have to be allocated to discrete land units. In particular when grid techniques are applied to sub-divide landscapes, thus resulting in large number of land units, usability of multi-objective optimization techniques can be limited by the dimensionality of the problem, which would lead to high required computation effort and time and uncertainty about the quality of the obtained solution sets. Such concerns can be partly alleviated when landscape units such as fields and their boundaries are represented as polygons with homogeneous land-use activities (Matthews, 2001).

In this paper we provide an illustration of a spatially explicit, GIS-based land-use optimization methodology named Landscape IMAGES (Interactive Multi-goal Agricultural Landscape Generation and Evaluation System) presented in Groot et al. (2006). This approach combines agronomic, economic and environmental indicators with biodiversity and landscape quality indicators. The paper describes the procedure to explore trade-off and gives an illustration. Moreover, we determine the effectiveness of adjustments in the initialization of the optimization procedure to improve the quality of the obtained solution sets.

2. METHODS

2.1 Conceptual Model

The assessment of the performance of a given farm or landscape can be based on multiple criteria, such as gross margin, nature value, landscape identity and nutrient losses. Different land-use activities make different contributions to the performance criteria and the activities on two or more spatial units may interact with respect to the performance criteria. Consequently, different configurations of activities result in different values of the performance criteria. The exploration of the trade-offs between performance criteria or objectives can be formulated as a multi-objective design problem, which can be generally stated as follows.

$$\text{Max } F(\mathbf{x}) = (F_1(\mathbf{x}),...,F_k(\mathbf{x}))^T \qquad (1)$$

$$\mathbf{x} = (x_1,...,x_n)^T \qquad (2)$$

Subject to i constraints:

$$g_i (\mathbf{x}) \le h_i \qquad (3)$$

Where, $F_1(\mathbf{x}),...,F_k(\mathbf{x})$ are the objective functions that are simultaneously maximized or minimized, and $(x_1,...,x_n)$ are the decision variables that represent the activities allocated to the n spatial units. The decision variables can take on values from a predefined array $\mathbf{x} \in S$, where S is the solution or parameter space. Constraints (Eq. 3) can arise from the problem formulation, for instance by limitations on the inputs or outputs related to the activities. Heuristic techniques such as genetic algorithms (GAs) and evolutionary strategies (ESs) can be employed to obtain approximations of the trade-off surfaces by a population of solutions, each representing a configuration of activities for the landscape.

2.2 Pareto-based Differential Evolution

The trade-offs between the objectives were explored with a multi-objective implementation of the ES algorithm of Differential Evolution (DE) developed by Storn and Price (1995). Currently, DE is widely used in the research community due to its simplicity, efficiency and robustness (Bergey and Ragsdale, 2005; Mayer et al., 2005). DE involves the iterative improvement of a set of solutions or genotypes. Each allele in the genotype is a real number. In our application, the genotypes represented alternative landscapes, and the alleles were decision variables in which the land-use of an individual field and the occupation of the field borders were encoded.

A genotype is a multi-dimensional vector $\mathbf{p}=(p_1,...,p_s)^T$ of s alleles. Each allele p_i is initialized as $p_{i,0}$ by assigning a random number within the allowed range:

$$p_{i,0} = L(p_i) + r_i \, (U(p_i) - L(p_i)) \qquad (4)$$

Where r_i denotes a uniformly distributed random value within the range [0,1] and L and U are the lower and upper values of the allowed range. A new generation x+1 is created by applying mutation and selection operators on the individuals in the population of genotype P of the current population x. The first step of the reproduction process is generation of a trial population P' that contains a counterpart for each individual in P, that is produced by parameterized uniform crossover (Spears and De Jong, 1991) of a target vector and a mutation vector. The mutation vector is derived from three mutually different competitors c_1, c_2 and c_3 that are randomly selected from the population P in the current generation x. The allele values are taken from the mutation vector with probability C_R:

$$p'_{i,x+1} = \begin{cases} c_3 + F \times (c_1 - c_2) & \text{if } r_i < C_R \\ p_{i,x} & otherwise \end{cases} \qquad (5)$$

The parameter $F \in [0,2]$ is a parameter that controls amplification of differential variations. After a mutation, the value of $p'_{i,x+1}$ can extend outside of the allowed range of the search space. For allele values that violate the boundary constraints the repair rule presented in Eq. 6 is applied. This rule implements a mechanism that can be denoted as 'back folding': the adjustment for the allele is calculated by interpolation into the allowed range from the boundary by a value that is proportional to the difference between the boundary and violation values:

$$p'_{i,x+1} = \begin{cases} L(p_i) - \dfrac{p'_{i,x+1} - L(p_i)}{F} & \text{if } p'_{i,x+1} < L(p_i) \\ U(p_i) - \dfrac{p'_{i,x+1} - U(p_i)}{F} & \text{if } p'_{i,x+1} > U(p_i) \\ p'_{i,x+1} & \textit{otherwise} \end{cases} \quad (6)$$

A trial genotype $p'_{i,x+1}$ replaces $p_{i,x}$ if it has a better ranking or is in a less crowded area of the search space (see below) than the parent genotype. Population size N is determined by the multiplication factor M (N=L×M). The last parameter is the number of generations G, which serves as the stopping criterion.

The first criterion for replacement of individuals by a trial solution is the pareto-based ranking. The ranking mechanism proposed by Goldberg (1989) is employed to evaluate the fitness of the individuals. Rank 1 is assigned to the non-dominated individuals and thus represents highest fitness values in the population. These individuals are removed from contention. A new set of non-dominated individuals in the rest of the population are ranked as 2 with next highest fitness values, and so forth until all of the individuals in the population are assigned a rank (Xue et al., 2003). An individual is replaced if the trial solution has a better ranking.

The second criterion for selection of trial solutions is the crowing distance metric proposed by Deb et al. (2002). This metric Θ represents the within-rank solution density and is calculated from the normalized distance for each objective between adjacent solutions in the search space, as follows:

$$\theta = \sum_{j=1}^{k} \frac{\left| d_i - \bar{d} \right|}{\left| B_j \right|} \quad (7)$$

In this equation, B_j is the boundary for objective j, which can be estimated from the difference between the minimum and maximum objective values along dimension j in the first rank. Parameter d_i denotes the Euclidian distance between two consecutive solutions within the Pareto front of a given rank. The parameter $\bar{d}$ is the average of these distances. An individual is replaced by a trial solution of the same rank if the latter has a higher value of Θ (Deb et al., 2002). This criterion promotes the spread of solutions within the objective space.

In the current maximization only problem, the distance of solutions from the origin should be maximized. Moreover, we aim to generate as wide a range of options as possible. Therefore, the size of the dominated space or hyper volume H (Zitzler, 1999) was used to evaluate the results of the DE optimization. H gives the volume enclosed by the union of area in the objective space where any point within this space is always dominated by at least one individual in the population P.

2.3 Landscape Optimization Problem

In the model agricultural land-use on the fields and the placement of hedges adjacent to the fields are allocated in an optimal manner, taking into account spatial heterogeneity and spatial interactions. In the current prototype implementation applied to regions dominated by dairy farming systems, the model seeks to maximize (i) gross margin from agricultural production, (ii) nature value of fields and borders and (iii) variation in the landscape in terms of species presence and hedge row allocation (half-openness). Constraints are applied to nutrient input and the proportion of herbage grazed. The landscape optimization problem and the calculation of indicators (see below) are described in detail by Groot et al. (2006).

Alternative land-use options that can be applied to fields were generated from simplified agro-ecological relations for grasslands and dairy production systems. To accommodate the implementation of discrete farm management choices and the possible inclusion of nature management packages, a discrete production activity generation approach was adopted (Van Ittersum and Rabbinge, 1997).

As indicator for the economic performance of farms, gross margin was used. The returns from production per field were calculated directly from the milk production and the milk price. Costs per field were separated into costs related to production (harvesting by grazing or mowing and fertilizer) and transport costs. The financial revenues from nature conservation packages were added to the value of the objective function for economic results. The applicability of conservation packages to individual fields was assessed on the basis of plant species abundance, and harvesting and fertilization regimes.

Species abundance in the grass swards and hedge rows was used as an indicator for nature value. The relationship between nutrient availability and

average species presence in grasslands was derived on the basis of data of Oomes (1992).
Landscape quality was related to variation in the landscape, calculated as the weighed sum of (1) the variance of the species number for each field and its adjacent fields and (2) the half-openness of the landscape, represented by the squared deviation from 50% occupation of the proportion of borders occupied by hedges.

2.4 Landscape

The methodology was applied to a hypothetical landscape (Figure 1). The majority of fields in this area belong to three farms, denoted A, B and C. Three fields were considered to represent the location of farm buildings. The other fields were buffer fields, which were not evaluated or updated throughout the optimization procedure. A gradient in soil fertility was assumed in the case study area (Figure 1), related to the nitrogen delivery capacity by the soil. This gradient was hypothetical with the purpose to illustrate the capability of the framework to deal with spatial variations in bio-physical circumstances. The ranges in nitrogen delivery capacity by the soil used here are actually observed in other case study areas on sandy soils.

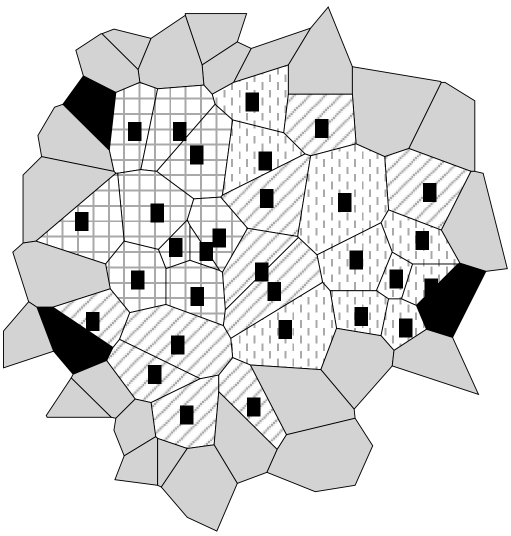

Figure 1. Hypothetical landscape with fields of three farms (different line patterns), location of farm buildings (black) and buffer fields (grey). Numbers indicate soil fertility level (0=140, 1=150, 2=160, 3=170, 4=180 kg N/ha/year).

2.5 Optimization experiments

Optimization experiments were conducted for 10,000 generations. The number of alleles per genotype was 60, i.e. 2 alleles for each of the 30 fields, one representing the land-use activity for the field and one encoding the border occupation with hedge rows. With a multiplication factor M of 10, the total DE-population comprised 600 genotypes.
The effect of enriching the DE-population with extremes from single objective optimizations on the explored volume of the solution space was tested. For each of the three objectives 1,000-generation minimization and maximization DE runs were carried out and the 25 best genotypes were selected. Thus, in total $3 \cdot 2 \cdot 25 = 150$ genotypes randomly replaced genotypes in the DE-population. Enriching was carried out after 100 generations, because at that stage constraint violating genotypes had been eliminated from the population.

3. RESULTS

The progress of the optimization and the effect of enriching the initial DE-population are presented for the trade-off between gross margin and nature value in Figure 2.
The spread of the genotypes within the objective space was considerably larger after enriching the initial population (Figure 2b). However, the hyper volume of the non-dominated front of the non-enriched population was the same as after enriching the initial DE-population ($1.61 \cdot 10^7$ versus $1.59 \cdot 10^7$), probably resulting from a better progress normal to the objective surface. In contrast, the size of the region of the objective space that was weakly dominated by the enriched front and not by the non-enriched front ($5.89 \cdot 10^5$) was larger than that dominated by the non-enriched front and not by the enriched front ($3.52 \cdot 10^5$). Therefore, it can be concluded that enriching the initial DE-population had resulted in improved quality of the solution.
In Figure 3 some examples of extreme landscapes generated from the optimization are presented. In the landscape with high gross margin (Figure 3a) the nature value was low, due to low plant species number in grassland associated with intensive management, and the low number of hedgerows. The reverse trend was observed for the landscape with high nature value (Figure 3b). The landscape in Figure 3c demonstrates high quality, here defined as variation in plant species number in adjacent fields and half-openness.

4. DISCUSSION

The optimization study with the Landscape IMAGES framework demonstrated that trade-offs between multiple objectives can be effectively explored in a spatially explicit land-use allocation problem. The solution sets contained a large range of possible configurations of the landscape in terms of land-use on fields and the placement of hedgerows on field borders. At a certain satisfaction level for a particular objective the potential 'window of opportunities' to improve on other objectives by selecting different production activities could be made explicit (Figure 2).

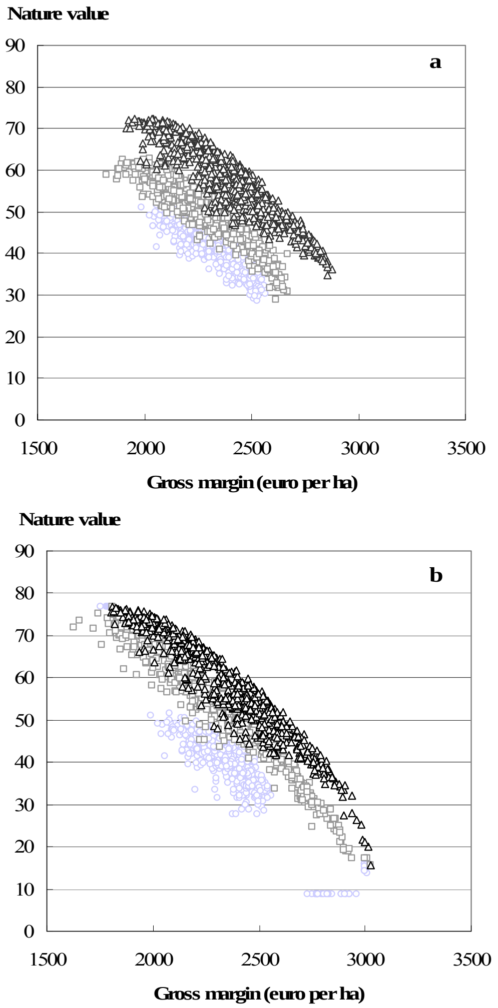

Figure 2. Landscape scale trade-off curves between gross margin and nature value after 100 (○), 1000 (□) and 10,000 (Δ) generations, without (a) and with (b) enriching of the initial DE-population.

Enriching the initial DE-population resulted in improved quality of the obtained non-dominated fronts representing the trade-offs between gross margin, nature value and landscape quality in landscapes, as indicated by hyper volume metrics (Zitzler,1999). However, for optimization during a fixed number of generations, improvement of the spread after enriching the initial DE-population came at the expense of progress in the direction normal to the objective surface. Increasing the number of generations could contribute to alleviating this drawback, at the cost of increased calculation effort.

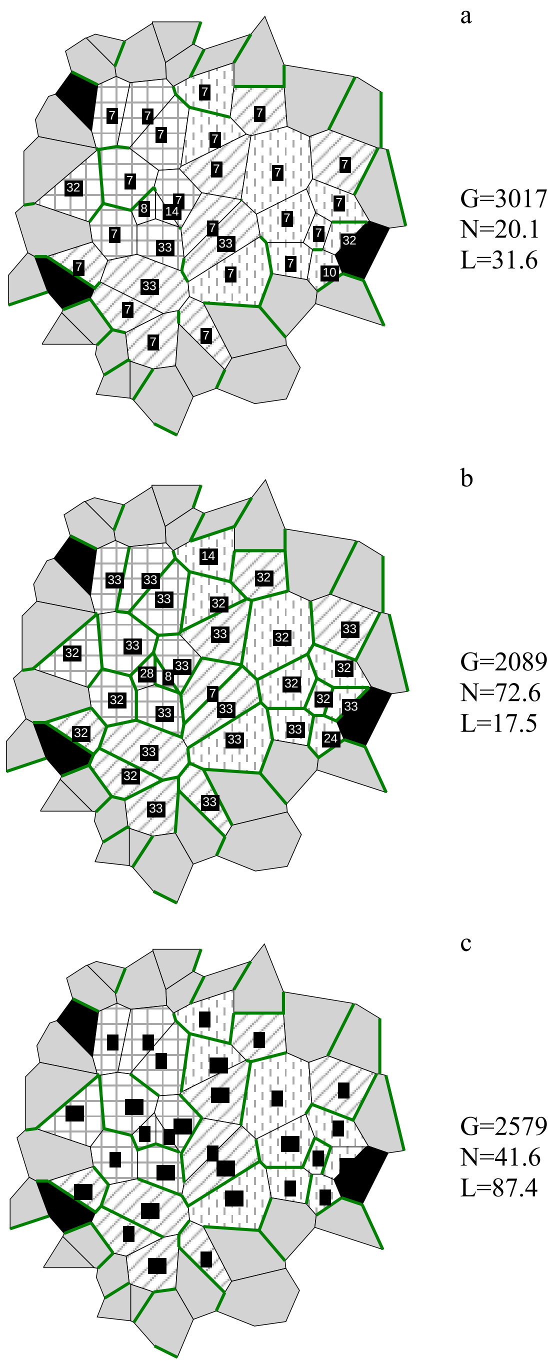

Figure 3. Example landscapes designed in the multi-objective optimization. Thick lines between fields indicate the presence of hedgerows; numbers denote the plant species number in grassland (per 25 m²). Objective values for the solutions are given: G=gross margin (euro per ha); N=nature value; L=landscape quality.

The generated alternatives as exemplified in Figure 3 offer ample opportunities for discussions with stakeholders on various topics. The current implementation with simplified agro-ecological relations illustrated that existing stakeholder questions can be addressed. The present version of the framework exhibits a number of requirements for effective model utilization in stakeholder discussions by, e.g., parameter, objective and constraint adjustment at the three relevant scales (field, farm and landscape), and selection of dimensions for visualization to enable interrogation of the results. These features enable the assessment of issues of mutual interest and explicit examination of different objectives and preferences. Moreover, the framework offers ample flexibility to adjust model functioning in consultation with stakeholders. Additional methods to effectively select alternatives that match the viewpoints of the respective stakeholders would further support stakeholder discussions.

5. ACKNOWLEDGEMENTS

The study was carried out as part of the project ‘Multifunctional Agriculture: From Farm Practice to Farm Design and Institutional Innovation’ in the MFA programme financed by the boards of WUR and INRA.

6. REFERENCES

Anderson, S.R., V. Kadirkamanathan, A. Chipperfield, V. Shafiri, and J. Swithenbank, Multi-objective optimization of operational variables in a waste inciration plant, *Computers and Chemical Engineering* 29, 1121–1130, 2005.

Bergey, P.K., and C. Ragsdale, Modified differential evolution: a greedy random strategy for genetic recombination, *Omega - International Journal of Management Science* 33, 255–265, 2005.

Deb, K., S. Agrawal, Pratap, and Meyarivan, A fast elitist non-dominated sorting genetic algorithm for multi-objective optimization: NSGA-II, *IEEE Transactions of Evolutionary Computation* 6, 182–197, 2002.

Goldberg, D.E., *Genetic Algorithms in Search, Optimization and Machine Learning*, Addison-Wesley, 412 pp, Reading, MA, 1989.

Groot, J.C.J., W.A.H. Rossing, A. Jellema, D.J. Stobbelaar, H. Renting, and M.K. Van Ittersum, Exploring multi-scale trade-offs between nature conservation, agricultural profits and landscape quality – a methodology to support discussions on land-use perspectives, *Agriculture, Ecosystems and Environment*, in press.

Matthews, K., *Applying genetic algorithms to multi-objective land-use planning*, PhD Thesis, The Robert Gordon University, Aberdeen, UK, URL: http://www.mluri.sari.ac.uk/LADSS/papers/keith-thesis.pdf , 2001.

Matthews, K.B., K. Buchan, A.R. Sibbald, and S. Craw, Combining deliberative and computer-based methods for multi-objective land-use planning, Agricultural Systems 87, 18–37., 2006.

Mayer, D.G., B.P. Kinghorn, and A.A. Archer, Differential evolution - an easy and efficient evolutionary algorithm for model optimisation, *Agricultural Systems* 83, 315–328, 2005.

O’Callaghan, J.R., NELUP: An Introduction, *Journal of Environmental Planning and Management* 38, 5–20, 1995.

OECD. *Multifunctionality – towards an analytical framework*, Organisation for Economic Co-operation and Development, Paris, 27 pp, 2001.

Oomes, M.J.M., Yields and species diversity of grasslands during restoration management, *Journal of Vegetation Science* 3, 271–274, 1992.

Seppelt, R., and A. Voinov, Optimization methodology for land use patterns using spatially explicit landscape models, *Ecological Modellling* 151, 125–142, 2002.

Spears, W. M., and K. A. De Jong, On the virtues of parameterized uniform crossover, paper presented at the 4th International Conference on Genetic Algorithms, 1991.

Storn, R., and K. Price, *Differential evolution - a simple and efficient adaptive scheme for global optimization over continuous spaces.* International Computer Science Institute, Berkeley USA, 12 pp. Technical Report TR-95-012, 1995.

Van Ittersum, M.K., and R. Rabbinge, Concepts in production ecology for analysis and quantification of agricultural input-output combinations, *Field Crops Research* 52, 197–208, 1997.

Xue, F., A.C. Sanderson, and R.J. Graves, Pareto-based multi-objective differential evolution, paper presented at the 2003 Congress on Evolutionary Computation (CEC2003).

Zitzler, E., *Evolutionary Algorithms for Multiobjective Optimization: Methods and Applications*. Swiss Federal Institute of Technology, Zürich, 120 pp. Dissertation ETH, No. 13398, 1999.