MULTIOBJECTIVE EVOLUTIONARY OPTIMIZATION OF GREENHOUSE VEGETABLE CROP DISTRIBUTIONS

A. L. Márquez[1], F. Manzano-Agugliaro[2], C. Gil[1], R. Cañero-León[3], F. G. Montoya[2] and R. Baños[1]
[1]*Department of Computer Architecture and Electronics*
[2]*Department of Rural Engineering*
[3]*Department of Management*
University of Almeria, Carretera de Sacramento S/N, Almeria, Spain

Keywords: MultiObjective optimization, Greenhouse crop distribution, NSGA-II, msPESA, Risk management.

Abstract: Multiobjective evolutionary algorithms (MOEAs) are known for their ability to optimize several objective functions simultaneously to provide a representative set of the Pareto front, which is a set of problem solutions representing a trade-off between the best values of each one of the objectives. This characteristic is specially interesting for the optimization of many real world problems, such as the allocation of land resources to maximize profit while reducing the economical risks associated to different distributions of crops in southern Spain, which has one of the largest concentrations of greenhouses in the world.

1 INTRODUCTION

While the purpose of any optimization procedure is to find the best possible solution to a certain problem, many problems have several objectives to optimize simultaneously. Therefore, the mission of Multi-Objective Optimization Problems (MOPs) is to find trade-off solutions instead of a single one.

Since the formulation of the problem of maximizing profit and reducing risks in crop area distribution involves optimizing two conflicting objectives at the same time, the aim is to obtain a set of solutions as an approximation to the Pareto-optimal set (Fonseca and Flemming, 1993). This set holds the best trade-off solutions found by the problem solver. Generating the Pareto-optimal set in complex problems is computationally expensive, and often infeasible, so a large number of heuristic approaches such as simulated annealing, tabu search, evolutionary algorithms, memetic algorithms, etc. have been proposed in the past.

In this study, the problem of planning greenhouse crops has been approached from an economic point of view with the aim of maximizing the profit and minimizing the risk. Therefore, this paper evaluates the performance of a multi-objective approach to solve this optimization problem with two different algorithms.

2 MULTI-OBJECTIVE OPTIMIZATION: CONCEPTS AND TECHNIQUES

This section introduces some multi-objective concepts, that are of key importance to understand the motivations and usefulness of the procedures presented on this paper. They are necessary because it is impossible to exactly describe what a good approximation to the Pareto Front is, in terms of a number of criteria such as closeness to the Pareto set, diversity, etc (Deb, 2002; Coello et al., 2002). Therefore, there is a need to introduce some Multi-Objective Optimization concepts:

Definition 1. Multi-objective Optimization is the process of searching for one or more decision variables that simultaneously satisfy all constraints, and optimize an objective function vector that maps the decision variables to two or more objectives.

$$minimize/maximize(f_k(s)), \quad \forall k \in [1,K]$$

Each decision vector or solution $s=\{(s_1, s_2, .., s_m)\}$ represents accurate numerical qualities for a MOP. The set of all decision vectors constitutes the *decision space*. The set of decision vectors that simultaneously satisfies all the constraints is called *feasible set (F)*.

Márquez A., Manzano-Agugliaro F., Gil C., Cañero-León R., Montoya F. and Baños R. (2009).
MULTIOBJECTIVE EVOLUTIONARY OPTIMIZATION OF GREENHOUSE VEGETABLE CROP DISTRIBUTIONS.
In *Proceedings of the International Joint Conference on Computational Intelligence*, pages 218-223
DOI: 10.5220/0002317002180223

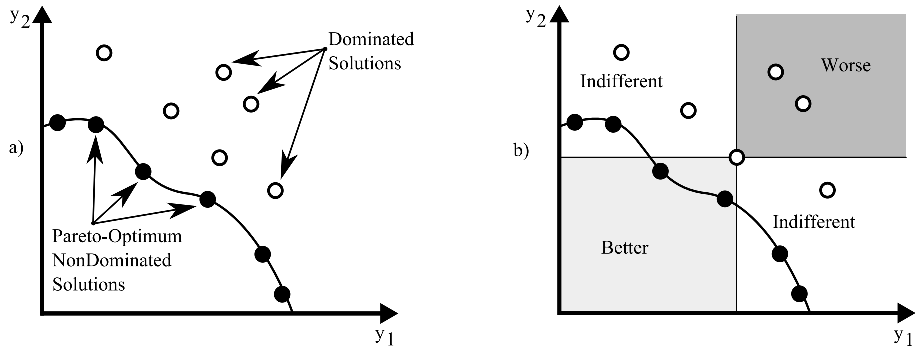

Figure 1: Pareto-dominance relations in a two-objective problem.

The objective function vector (f) maps the decision vectors from the decision space into a K-dimensional objective space $Z \in \Re^K$, $z=f(s)$, $f(s)=\{f_1(s), f_2(s),..., f_K(s)\}$, $z \in Z$, $s \in F$.

Given a MOP with $K \geq 2$ objectives, instead of giving a scalar value to each solution, a partial order is defined according to Pareto-dominance relations, as we detail below.

Definition 2. Order Relation between Decision Vectors. Let s and s' be two decision vectors. The dominance relations in a minimization problem are:

$$s \textit{ dominates } s' \ (s \prec s') \quad \textit{iff}$$
$$f_k(s) < f_k(s') \wedge f_{k'}(s) \not> f_{k'}(s'),\ \forall k' \neq k \in [1,K]$$
$$s, s' \text{ are } \textit{incomparable } (s \sim s') \quad \textit{iff}$$
$$f_k(s) < f_k(s') \wedge f_{k'}(s) > f_{k'}(s'),\ k' \neq k \in [1,K]$$

Definition 3. Pareto-optimal Solution. A solution s is called *Pareto-optimal* if there is no other $s' \in F$, such that $f(s') < f(s)$. All the Pareto-optimal solutions define the *Pareto-optimal set*.

Definition 4. Non-dominated Solution. A solution $s \in F$ is *non-dominated* with respect to a set $S' \in F$ if and only if $\nexists s' \in S'$, verifying that $s' \prec s$.

Definition 5. Non-dominated Set. Given a set of solutions S', such that $S' \in F$ and $Z'=f(S')$, the function $ND(S')$ returns the set of non-dominated solutions from S':

$$ND(S') = \{\forall s' \in S' | s' \text{ is non-dominated by any other } s", s" \in S'\}$$

Figure 1 graphically describes the Pareto-dominance concept for a minimization problem with two objectives (f_1 and f_2). Figure 1(a) shows the location of several solutions. The filled circles represent non-dominated solutions, while the non-filled ones symbolize dominated solutions. Figure 1(b) shows the relative distribution of the solutions in reference to a solution s. There exist solutions that are *worse* (in both objectives) than s, *better* (in both objectives) than s, and *incomparable* (better in one objective and worse in the other).

2.1 NSGA-II

The first algorithm we have used is the *Non-dominated Sorting Genetic Algorithm II* (Deb et al., 2000) which is an extended version of NSGA (Srinivas and Deb, 1994). NSGA-II creates a non-dominance based hierarchy by using its non-dominated sorting procedure, which is combined with a crowded comparison operator that keeps diversity by using the euclidean distance between solutions to specify the fitness of each one. In its operation (see (Deb et al., 2000)) non-dominated solutions are added to the external archive by using the dominance hierarchy previously created, until it is completely filled. The remaining solutions are discarded.

2.2 msPESA

The second algorithm we have used is msPESA (Gil et al., 2007), a hybrid algorithm that combines aspects of PESA (Corne et al., 2000) and NSGA-II. It uses a small internal population and a larger external population (archive), implements a variant of the archive-update strategy of PESA, and takes some ideas of the selection mechanism used in NSGA-II.

When a candidate solution enters the archive, instead of removing the solutions it dominates from the archive, msPESA keeps them for the sake of improving genetic variability. It can perform a local search procedure using the mutation operator over the candidate solution, to try to find more good solutions in its neighbourhood. After that, the non-dominated solutions generated from the local search are stored within the archive. Once the archive is full, the squeeze factor is used (see below) to remove the excess of solutions from the archive.

The squeeze factor is calculated with the help of the hyper-grid strategy. In msPESA the objective space is divided in a $(N-1)$ dimensional grid, which can be used to achieve a finer resolution with the same memory requirements. It even allows using as many divisions of the search space as subjects can be stored into the Archive, therefore allowing the procedure to obtain an evenly distributed Pareto front, since it tries to keep a maximum of one subject on each cell of the hyper-grid.

3 OPTIMIZING CROP DISTRIBUTION

Plastic-covered greenhouses have undergone rapid expansion in recent years, covering a surface of over 1.600.000ha worldwide (Espi et al., 2006), which are mainly distributed in two geographical areas: The Far East (whose maximum contributors are China, Japan and Korea) with almost 80% of the surface (Jiang et al., 2004) and the Mediterranean area with about a 15% of the world's greenhouse-covered extensions (Pardossi et al., 2004).

Table 1: Main economical data obtained from the different greenhouse crop options (€$/m^2$).

	2004-2005		*2005-2006*	
Crop	**Revenue**	**Cost**	**Revenue**	**Cost**
Pe	3.85€	1.70€	5.62€	2.12€
Pe-M	5.35€	2.27€	4.43€	2.82€
Pe-W	5.55€	2.51€	4.48€	4.02€
T-M	6.90€	2.23€	6.18€	3.76€
T-W	6.69€	1.99€	6.14€	2.56€
T-T	9.07€	3.91€	8.21€	4.88€
GB-GB	5.40€	2.68€	8.90€	2.74€
GB-M	5.17€	2.06€	6.16€	1.98€
GB-W	3.55€	1.89€	6.20€	2.18€
Cu	10.63€	1.88€	5.83€	2.58€
Cu-M	6.80€	1.92€	4.64€	2.53€
Cu-W	7.01€	2.16€	4.69€	2.73€
Cu-GB	7.03€	2.53€	7.14€	3.09€
Co-Co	8.58€	2.14€	5.12€	2.22€
Co-W	5.06€	1.79€	6.29€	2.03€
Co-M	4.86€	1.55€	6.24€	1.83€

The greenhouse crop surface at the province of Almeria (south-eastern Spain) is about 30.000ha with an estimated production of $3x10^9$kg of produce at an approximate value of $1,384 \cdot 10^6$€ (IEC, 2009), where 80% of the cultivated crop varieties are: pepper, tomato, green beans, cuccumber, courgette, watermelon and melon (Manzano-Agugliaro, 2007). The distributions of these crop surfaces change every year, sometimes causing a decrease in average prices compared with the previous year. The behaviour of average prizes is inversely proportional to the quantities produced (IEC, 2009).

In previous works, weighted goal programming using utility functions has been used as a methodology for the analysis and simulation of agricultural systems (Sumpsi et al., 1997; Amador et al., 1998). These techniques have been used in a decision-making process in planning crops (Berbel and Gomez-Limon, 1997). This is consistent with how classical optimization methods suggested converting the multi-objective problems into a single-objective formulation by combining the objectives in only one mathematical function (Fonseca and Flemming, 1993).

3.1 Problem Information

The seven main crop varieties mentioned above are combined in sixteen vegetable crop alternatives that are the object of this study. The data fed to the solver has been obtained from an accountancy tracking of 46 and 49 greenhouses in 2004-2005 and 2005-2006 seasons respectively. There is additional information for each crop to be dealt with, such as production, revenue, fertilisers, needed water per m^2, or fixed costs (depreciation, soil disinfection, etc). With this information we can obtain the Gross Margin (GM) (see table 1) for each crop option and each year (revenue minus variable costs), irrigation water consumption (m^3/m^2), nitrate consumption (Fertilizing unit per m^2). For the actual crop distributions we have observed a gross margin of $84,666$€ and risk value of $465 \cdot 10^6$.

3.2 Objectives to Optimize

There are sixteen variables in this problem which correspond to the surface distribution for each one of the sixteen crop alternatives mentioned before, there are two restrictions and two main objectives to optimize that are of the highest interest to farmers:

Maximizing Profit. To do this, the gross margin (GM) is maximized for the various options proposed:

$$GM = \sum_{i=1}^{n} (GM_i \cdot X_i) \qquad (1)$$

Where GM_i is the gross margin of option i per surface area unit (€$/m^2$) and X_i is the surface area that the crop alternative covers.

Minimizing Risk. Since agricultural production may be affected by random conditions, the risk is a factor that influences the choices of the crop options. To calculate risk, the variance and

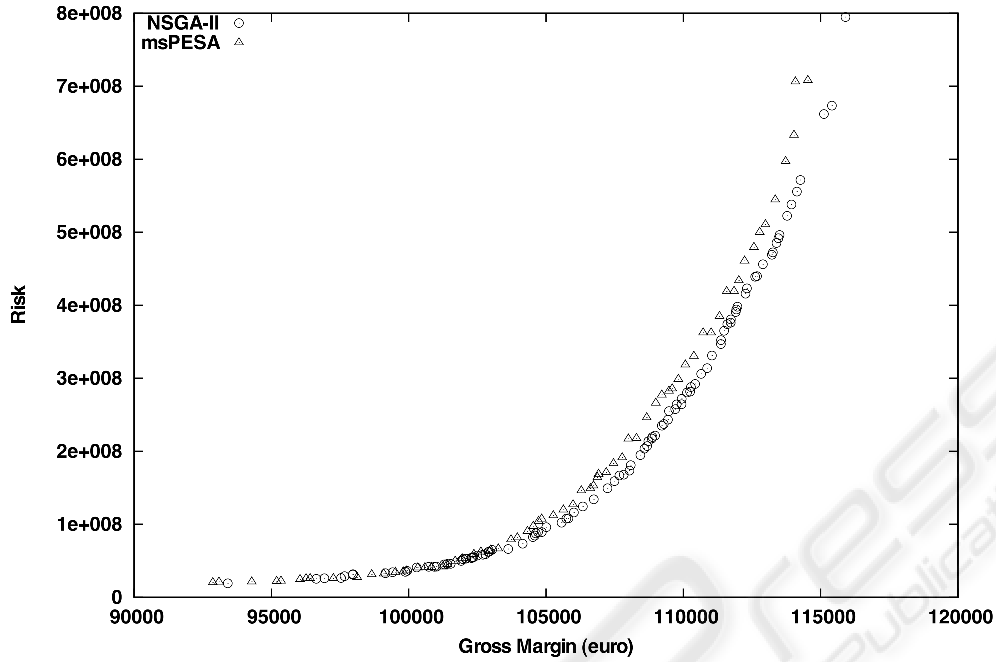

Figure 2: Pareto front of Gross Margin (GM) vs. Risk for NSGA-II and msPESA.

covariance matrix is used for the gross margins of the different crop options, based on the above data.

$$R = \overline{X_i}[cov]\overline{X_i} \quad (2)$$

Since these objectives are on conflict, the use of Multi-Objective Evolutionary Algorithms for generating a Pareto Front with the distributions of the crop alternatives seems appropriate, especially considering that with enough abundance of widely distributed solutions over the objective space we could offer the greenhouse owners the choice to select the most appropriate distribution of crop alternatives considering his financial needs or the risks that the farmer is able to withstand.

There are also a couple of constraints that are to be taken into account:

- The total surface area is limited to 2.5ha, which is the average greenhouse surface.

$$\sum_{i=1}^{n} X_i = 25,000m^2 \quad (3)$$

- The maximum surface area for a certain crop should never be higher than the 40% of the total surface as a restriction, because it produces a market flood of the product, leading to an important drop in the prices, therefore reducing the revenues for that crop to the point that it could even generate a money loss to the owners of the affected lands.

3.3 Evolutionary Operations

A floating point numeric representation for each one of the sixteen variables has been chosen, because it is the most natural representation for this problem. Furthermore, it is necessary to implement specific operators in MOEAs to be able to work with this problem, such as mutation, crossover and a chromosome repair procedure. While in other problems there might not arise such need, it is necessary to implement a chromosome repair mechanism for the crop surface optimization procedure to guarantee that the generated solutions fall within the problem restrictions, therefore making them feasible to use in the next generation of solutions for each iteration of the MOEA.

Mutation. The mutation operation is the simplest of all three, since it just iterates through all the chromosome variables applying a random change of up to a $\pm$25% of its initial value.

Crossover. We applied a multipoint chromosome crossover, where each one of the values of the genotype for each parent chromosome has a 50% of probabilities to ending up as part of the genotype for the chromosome of the next generation.

Chromosome Repair Mechanism. Its function is to normalize the total surface represented by each one of the chromosome variables to a surface of 2.5ha. Once it has been done, then we proceed to calculate the total surface belonging to each one of

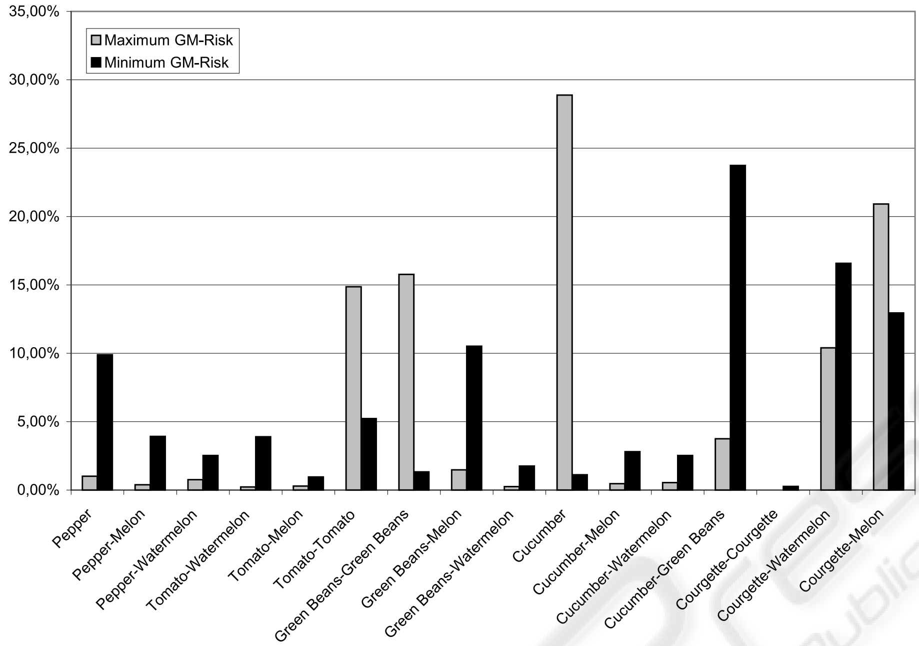

Figure 3: Comparison between surfaces on the extreme opposite solutions in the Pareto Front for a single run of NSGA-II.

the seven crops, to check if any of them extends over a 40% of the total surface. If that is the situation, we proceed to proportionally redistribute the excess surface between the rest of the crops. This means that if, for example, the surface of cucumber is distributed between 3 crop alternatives, each one of them would be substracted a proportional amount of the excess surface, meaning that the greatest contributor to the situation is the alternative that will experience a greater surface reduction. This also ensures that the crop reduction is not greater than the available surface for the affected alternative.

4 EXPERIMENTAL RESULTS

Since Multi-Objective Evolutionary Algorithms are highly configurable procedures (Goldberg, 1989), we have chosen to use similar parameters for both of them, always considering the particularities of each alternative. For both algorithms we chose to execute the same number of function evaluations (20,000) to allow them to perform enough evaluations to generate good solutions. The parameters chosen for each procedure are:

NSGA-II. This algorithm works at its best when both the internal and external population have the same size.

- Population and External population: 100 individuals.
- Mutation rate: 0.1

msPESA. For msPESA, working with a small internal population allows the procedure to generate a greater number of generations. Since all the best solutions are always kept in the archive, this allows it to generate a good approximation Pareto Front.

- Population: 10 individuals.
- External Population (Archive): 100 individuals.
- Mutation rate: 1.0

The figure 2 represents a typical execution of NSGA-II and msPESA for the problem at hand. As we can observe, this graphic shows how the non-dominated solutions obtained as approximation to the Pareto-optimal front have a clear definite shape, meaning that there is a strong relation between risk and the gross margin, and that they are opposite objectives, since the ideal situation would yield a maximum gross margin with a low risk. It is worth noting that both algorithms yield similar results, being NSGA-II the algorithm that generates the best Pareto-optimal set of solutions using the proposed experimental parameters.

5 CONCLUSIONS

The nature of the problem of crop surface optimization allows it to be efficiently represented and developed as a Multi-Objective problem that can be solved using any of the current algorithms. This allows the user to search and analyze a wide range of possible situations to choose the simulated solution that may be of the best interest for the situation of a particular farmer.

As figure 3 shows, the differences between a higher profit-higher risk simulation and a lower profit-lower risk one are significant, meaning that green beans are a relatively safe value to use, while cucumber is the most profitable crop to plant, but its associated risks are high enough to be considered.

The new distribution of crops obtained with this method, shows better Gross Margin and lower Risk in the minimum point of the Pareto Front than the real situation. This means that the crop distributions may be optimized in order to maximize the benefits for the greenhouse farmers.

ACKNOWLEDGEMENTS

This work has been financed by the Spanish Ministry of Innovation and Science (TIN2005-00447), Ministry of Science and Technology (AGL2002-04251-C03-03) and the Excellence Project of Junta de Andalucía (P07-TIC02988), in part financed by the European Regional Development Fund (ERDF).

REFERENCES

Amador, F., Sumpsi, J. M., and Romero, C. (1998). A non-interactive methodology to asses farmers' utility functions: An application to large farms in andalusia, spain. *European Review of Agricultural Economics*, 25:92,109.

Berbel, J. and Gomez-Limon, J. (1997). The impact of water-pricing policy in spain: an application to irrigated farms in southern spain. *European Journal of Operational Research*, 107:108–118.

Coello, C., Veldhuizen, V., D.A., and Lamont, G. (2002). *Evolutionary Algorithms for solving Multi-Objective Problems*. Kluwer Academic Publishers.

Corne, D., Knowles, J., and H.J., O. (2000). The Pareto-Envelope based Selection Algorithm for Multiobjective Optimisation. *Lecture Notes in Computer Science*, 1917:869–878.

Deb, K. (2002). *Multi-Objective Optimization using Evolutionary Algorithms*. John Wiley & Sons.

Deb, K., Agrawa, S., Pratap, A., and Meyarivan, T. (2000). A fast elitist non-dominated sorting genetic algoritm for multiobjective optimization: Nsga-ii. *Lecture Notes in Computer Science*, 1917:849–858.

Espi, E., Salmeron, A., Fontecha, A., García-Alonso, Y., and Real, A. (2006). Plastic films for agricultural applications. *Journal of Plastic Film and Sheeting*, ss:85–102.

Fonseca, C. and Flemming, P. (1993). Genetic algorithms for multiobjective optimization: Formulation, discusion and generalization. In Forrest, S., editor, *Proceedings of the Fifth International Conference on Genetic Algorithms*, San Mateo, California.

Gil, C., Márquez, A., Baños, R., Montoya, M., and Gómez, J. (2007). A hybrid method for solving multi-objective global optimization problems. *Journal of Global Optimization*, 28(2):265–281.

Goldberg, D. (1989). *Genetic Algorithms in Search, Optimization and Machine Learning*. Addison Wesley, New York.

IEC (2009). Análisis de la campaña hortofrutícola de Almería, campaña 2007-2008 (greenhouse crop analisys in Almeria for the 2007/2008 season).

Jiang, W., Qu, D., Mu, D., and Wang, L. (2004). China's energy-saving greenhouses. *Chronica Horticulturae*, 44:15–17.

Manzano-Agugliaro, F. (2007). Gasification of greenhouse residues for obtaining electrical energy in the South of Spain: Localization by gis. *Interciencia*, 32(2):131–136.

Pardossi, A., Tognoni, F., and Incrocci, L. (2004). Mediterranean greenhouse technology. *Chronica Horticulturae*, 44(2):28,34.

Srinivas, N. and Deb, K. (1994). Multiobjective optimization using nondominated sorting in genetic algorithms. *Evolutionary Computation*, 2 (3):221–248.

Sumpsi, J. M., Amador, F., and Romero, C. (1997). On farmers' objectives: A multi-criteria approach. *European Journal of Operation Research*, 96:64–71.