Exploiting Comparative Studies Using Criteria: Generating Knowledge From An Analyst's Perspective

Daniel Salazar[1,*], Néstor Carrasquero[2], and Blas Galván[3]

[1] Postgrado de Investigación de Operaciones. Facultad de Ingeniería, Universidad Central de Venezuela. Caracas, 1041-A, Venezuela
daniel_salazar@cantv.net
[2] Grupo de Optimización Combinatoria Emergente (GOCE). Facultad de Ingeniería, Universidad Central de Venezuela. Caracas, 1041-A, Venezuela
nestor_carrasquero@cantv.net
[3] Instituto de Sistemas Inteligentes y Aplicaciones Numéricas en Ingeniería (IUSIANI)
División de Computación Evolutiva y Aplicaciones (CEANI)
Universidad de Las Palmas de Gran Canaria, Islas Canarias, España
bgalvan@step.es

Abstract. In this work the use of qualitative preferences for classifying and selecting MOEAs is introduced. The classical notions of the Analyst and the so called Prescriptive Analysis are introduced explicitly in EMO, identifying some difficulties in exploiting the results of the comparative studies performed by the current fashion. A methodology is developed that allows the analyst to translate DM's general preferences as well as quantitative benchmarking results into a practical tool for the comparison of MOEAs, facilitating the selection of the proper method and/or parameters for the MCDM problem at hand. A comparative experimentation is performed using well known state of the art functions, allowing drawing clear conclusions about the utility of the proposed methodology. The results are useful for research, practitioners and analysts involved in benchmarking, comparative studies and prescriptive analysis for EMO.

1 Introduction

When Multiple Criteria Decision Making (MCDM) is modelled, different stages and actors can appear as a part of the whole process. An actor is defined as any individual, group of individuals or entity, playing any role during the decision making process [1][2]. In this sense, besides the Decision Maker (DM), it is useful to identify another actor called the *analyst*. Arsham [3] lists a sequence -which allows feedback loops- of tasks accomplished by the analyst:

1. Understanding the Problem
2. Constructing an Analytical Model
3. Finding a Good Solution
4. Communicating the Results with the Decision-Maker

[*] The author is now also with 3.

Once the model is constructed, the analyst must choose a suitable method or solving technique. That is intended for "Finding a Good Solution". This stage (steps 2 and 3) is known as prescriptive analysis.

To our best knowledge, the notions of analyst and prescriptive analysis have not been introduced explicitly in Evolutionary Multiobjective Optimization (EMO) yet (despite they can be found in some way applied in practice, e.g. [4]), maybe due to these notions could seem more appropriate for decision making techniques which deal with problems not well defined (in terms of mathematical formulation) and where the preferences are articulated *a priori* or *interactively* (e.g. outranking methods). Even so, the lack of these concepts in EMO does not mean that EMO researchers and practitioners are not aware of them. Nevertheless, we do believe that considering prescriptive analysis explicitly in EMO can yield worthy results.

In this paper we concentrate in the case of *a posteriori* Multiple Objective Evolutionary Algorithms (MOEAs), which represent the most of available MOEA (Van Veldhuizen and Lamont [5] mention about 90%, and apparently this bias has not changed significantly in the last years). Consider now the role of an analyst who works with a posteriori MOEAs. Analyzing the No Free Lunch theorem, Knowles and Corne conclude in [6][7] that some multiobjectives optimizers are better than others. In consequence, it sounds reasonably that given a pool of MOEAs, the analyst should select the proper method for solving MCDM problem. The issue is *how* to do that; in fact, the selection could be far to be trivial in many cases.

In this research, we study a group of comparative analysis of some relevant state-of-art MOEAs, identifying some points which limit the ability of an analyst to choose a particular algorithm. Then, we developed a methodology, sustained in comparative studies, that allows the analyst to *translate* DM's preference as well as quantitative benchmarking results into a practical tool for the comparison of MOEAs. As a result, the influence of genetic parameters (crossover and mutation rates) and population size upon the overall performance were assessed empirically for three relevant MOEAs, and then interpreted, building qualitative preference maps which can help the analyst to achieve the prescriptive analysis. As the reader can intuit, the present work is relevant not only for persons interested in prescriptive analysis in EMO, but also for researchers and practitioners in general, involved in benchmarking and comparative studies.

The remainder of the paper is organized as follows. In the next section the concepts of analyst and prescriptive analysis are studied more deeply regarding the state of art in comparative studies in EMO. In section 3 the proposed methodology is developed step by step, showing how to exploit benchmarking by introducing preferences. Finally some concluding remarks are presented.

2 Background

In order to set this research in its context, we will discuss in this section how considering explicitly the notion of the analyst can be beneficial for EMO. To do that, let us begin considering the single MCDM/EMO model depicted in figure 1. In this proposed model, for the sake of simplicity, MOEAs are represented as black-box multiobjective optimizers. EMO is enclosed by a dashed line as is focused nowadays, i.e. with the three possible stages for DM's preference articulation [8]. At the right side of EMO's box, enclosed in a solid line, is the Decision Making stage, when the DM chooses a single alternative as a solution of the problem. Finally, at the left side of the figure is the first step of the whole process: the decision making model building.

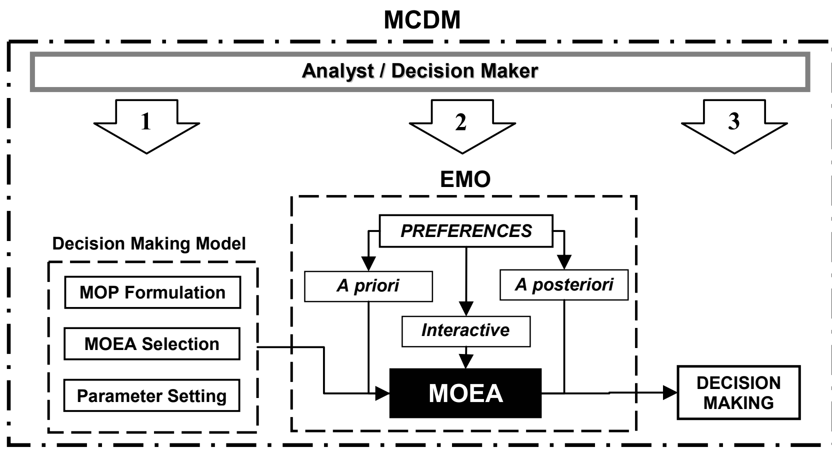


Fig. 1. A single MCDM model using MOEAs as a solution technique.

For a MDCM problem where EMO is employed as a solving technique, building the model comprises at least, three stages: Multiobjective Problem (MOP) formulation, MOEA selection and parameters setting. Generally speaking, model building requires the analyst to collect information (intelligence) [9] not only about the problem but also about the kind of solution the DM wants to reach. In other words, for constructing a model and selecting an adequate solving technique, it is good for the analyst to have a *general* idea of the solution wanted, despite the preferences were articulated a posteriori. For instance, suppose the case of a DM who doesn't have any notion about the technical aspects of a MOEA, but who is interested in solving a particular problem where the objective functions evaluation is time consuming. In this case, more than one exploration of the search space may be prohibitive; hence, any additional information is crucial for providing efficiency. If the analyst detects that the DM is more interested a particular group of aspects than in another, this interest should lead the selection of the MOEA. Thus, for our example, between two algorithms which none of them outperforms each other, in terms of convergence and diversity, the less sensitive to population size the better, because it means a reduced number of functions evaluation. On the other hand, even without articulating preferences, a DM could prefer convergence to diversity, because too many solutions complicate the final decision making.

Zeleny [10] remarks that "Decision makers are expected to have confidence in the analyst and the analyst's tools rather than in themselves". This is especially palpable, as we have just seen, when the DM doesn't have any conception of the technical aspects of EMO, and becomes a very relevant point when there is any matter that makes exploring the search space a difficult task. But let us focus now on the particular features an analyst must take into account when he works with a posteriori method, and the difficulties which take place for an analyst for the current EMO state of art.

Once the analyst has determined to use a posteriori method, the second step is to choose a particular MOEA, setting the proper parameters for the selected algorithm. The main tools to tackle this task are the published comparative studies or the studies that the analyst could make by his own. However, when analyzing the literature, is not easy to make solid conclusions about which method is better. Consider the problem of the parameters in the classification results. In [11] the authors found a "hierarchy of method" between eight different MOEAs for a determined group of parameters, but admitted that the situation may be different for other parameters settings and other test problems [11, page 193]. Such situations clearly generate doubts in the analyst. Furthermore, when introduced three well known methods (NSGA-II [12], SPEA2 [13], PESA-II [14]), each author reports their results using different parameters settings (crossover and mutation rates), and even different representation techniques (binary and real) with crossover operators distinct from each other. A brief review of more recent works, including new methods, applications and comparisons studies, ([15]-[21]) shows that the diversity in parameters settings, genetic representations and genetic operators still remains in EMO, not allowing the analyst to make direct comparisons and results extrapolation between researches (in fact, only [15] performed tune before comparing). Finally, even the traditional heuristic of mutating one bit per chromosome may prevent an algorithm to yield maximum efficiency in certain conditions [22][23].

The analyst may consider using auto adaptative or parameter-less algorithms ([24]-[27]) to avoid the problem of setting parameters, but these work schemata may present drawbacks [28] and employing them does not solve the problem completely but the selection of the proper algorithm remains open and features like genetic operators or internal parameters may be changed, possibly affecting the overall performance.

On the other hand, there is the issue of the performance metrics. When solving a MOP, most *a posteriori* MOEAs produce as outcome a set of non dominated solutions. These outcomes can be compared using unary or binary metrics. Some outperformance relationships have been introduced to analyze the quality of unary metrics [29] [30]. Latter in [31] a general framework for comparing outcomes is presented, demonstrating that unary metrics have theoretical limitations when no preferences information is used, and providing some indicators for binary metrics based comparisons. However, even when binary metrics do not have such theoretical limitations of unary metrics [31, page 127], in practice they present some difficulties, like they are not conceived for multiple runs, and some could lead to conclusions different from those which one could come by expressing some preference (e.g. see figures 2 and 3 in [32]).

In summary, selecting a particular solving technique is a decision making problem itself, where the analyst is the actor in charge of making the selection based on the desires, aspirations and general preferences of the DM. Some difficulties to accomplish this task are:

- The published comparative studies were performed employing different parameters settings, and analyzed with different metrics. Such a variety makes difficult to exploit the information presented in those works.
- Even when, for the sake of objectivity, researchers tend to work with metrics not based on DM's preferences, in practice at least some general information could be useful for working with unary and binary metrics ([29] and [33] are based on the assumption that some information is available)

In the following sections we present a methodology for the incorporation of DM's general preferences providing an aid to the analyst in the model building process. The methodology, as is presented, is based on a complete experimentation, which among other things, clarify some questions formulated before.

3 Exploiting Information using Criteria

The following subsections present the development of the methodology proposed from the generation of experimental data to their interpretation using DM general preferences.

3.1 Experiment design

In order to generate knowledge about the behaviour of the MOEAs when changing parameters settings, the experimental design was formulated as follows:

- SPEA2, PESA-II and NSGA-II were selected to be studied using functions ZDT2 to ZDT4 and ZDT6 [12] (see table 1).
- The chromosomes were represented with binary strings with one-point crossover and bitwise mutation (for the number of bits see table 1).
- Parameters were fixed as follows: six values of crossover rate (*pc*) and three for mutation (*pm*) rates were employed (pc = {0.5, 0.6, 0.7, 0.8, 0.9, 1}, pm = {0.001, 0.010, 0.100}). Population size – archive size ratios (M:N) were fixed to 1:1 and 2:1. The maximum number of generation was set to 250.
- For each function and parameters settings (pc, pm, M:N) ten runs were performed. The outcomes were analyzed comparing the mean value of metric *S* (for hypervolume covered) [30].
- Computational complexity was not considered.

Table 1. Multiobjective test problems selected for experimentation

Name	Formulation:	Domain	Optimal set
ZDT2	$f_1(x) = x_1$ $f_2(x) = g(x)\left(1-\left(x_1 / g(x)\right)^2\right)$ $g(x) = 1 + 9\left(\sum_{i=2}^{9} x_i\right)/(n-1)$	$x_i \in [0,1]$ n = 30	$x_1 \in [0,1]$ $x_i = 0$ $i \in \{2,...,n\}$
ZDT3	$f_1(x) = x_1$ $f_2(x) = g(x)\left(1-\sqrt{x_1 / g(x)} - \frac{x_1}{g(x)}\sin(10\pi x_1)\right)$ $g(x) = 1 + 9\left(\sum_{i=2}^{9} x_i\right)/(n-1)$	$x_i \in [0,1]$ n = 30	$x_1 \in [0,1]$ $x_i = 0$ $i \in \{2,...,n\}$
ZDT4	$f_1(x) = x_1$ $f_2(x) = g(x)\left(1-\sqrt{x_1 / g(x)}\right)$ $g(x) = 1 + 10(n-1) + \sum_{i=2}^{9}\left(x_i^2 - 10\cos(4\pi x_i)\right)$	$x_1 \in [0,1]$ $x_i \in [-5,5]$; $i \in \{2,...,n\}$ n = 10	$x_1 \in [0,1]$ $x_i = 0$ $i \in \{2,...,n\}$
ZDT6	$f_1(x) = 1 - \exp(-4x_1)\sin^6(4\pi x_i)$ $f_2(x) = g(x)\left(1-\left(x_1 / g(x)\right)^2\right)$ $g(x) = 1 + 9\left[\left(\sum_{i=2}^{9} x_i\right)/(n-1)\right]^{1/4}$	$x_i \in [0,1]$ n = 10	$x_1 \in [0,1]$ $x_i = 0$ $i \in \{2,...,n\}$

3.2 Building preferences maps

To aid the analyst during the formulation of the whole decision making model, preferences maps were conceived as the main tool to translate the quantitative information obtained from the metrics into qualitative information. Naturally, the selection of the metrics is related with the features considered important by the DM. In our case, we chose unary metrics because they allow statistic treatment. The selection of *S* is recommended in [30] for few objectives. Any Analyst/DM may base his maps in other metrics.

Figure 2 present the results of metric *S* for function ZDT4. The matrix layout allows comparisons between methods and between parameters for each method. The reader can easily note some variations in means values. At this stage, it is necessary to express some general criteria for classify the vectors of means (the curves). Hence, four preference relationships were proposed to build the classifications, they express: $\rhd$ *Strictly preference*, $\unrhd$ *Preference*, $\parallel$ *Indifference*, $\ntrianglerighteq$ *No preference.* They were inspired in the relationships formulated in [31], but not for express dominance but preference.

Finally, very simple and intuitive statements were formulated here as classification criteria (lexicographically):

- *if* one curve is closer the optimal than another, *then* that curve is strictly preferable, *else*
- *if* one curve is closer to the optimal or at the same level than another but it has lower variance, *then* that curve is preferable, *else*
- *if* one curve intersects another, *then* that curve is indifferent, *else*
- the curve is no preferable.

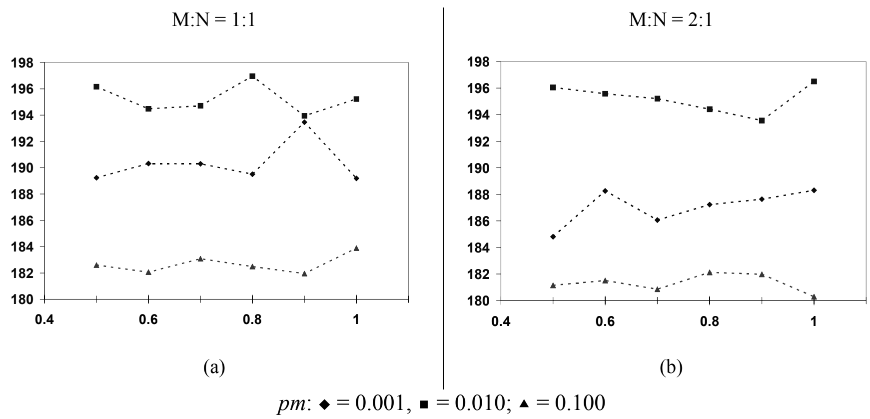


Fig. 2. Mean values of metric *S* vs. *pc* for SPEA2 over function ZDT4.

Substituting the if-then-else sentences any analyst/DM can build an appropriate classification for his/her criteria.

For the construction of the map, few topics were considered. First, the crossover rate was divided in two groups, low crossover (LC) with comprises *pc* from 0.5 to 0.7 and high crossover for *pc* from 0.8 to 1.0. Then, for each group and populations ratio, an assessment mutation rates were accomplished, based on the rules described above.

M:N	Metric S					
	1:1			2:1		
LC	10^{-3}	$\rhd$	$\not\rhd$	10^{-3}	$\rhd$	$\not\rhd$
	$\not\rhd$	10^{-2}	$\not\rhd$	$\not\rhd$	10^{-2}	$\not\rhd$
	$\rhd$	$\rhd$	10^{-1}	$\rhd$	$\rhd$	10^{-1}
HC	10^{-3}	$\rhd$	$\not\rhd$	10^{-3}	$\rhd$	$\not\rhd$
	$\not\rhd$	10^{-2}	$\not\rhd$	$\not\rhd$	10^{-2}	$\not\rhd$
	$\rhd$	$\rhd$	10^{-1}	$\rhd$	$\rhd$	10^{-1}

M:N	Global population-size ratio	
	1:1	2:1
1:1	---	$\not\rhd$
2:1	$\unrhd$	---

Fig. 3. SPEA2 preference map for metric *S* and function ZDT4. For reading always start vertically!

MOEA		ZDT2						ZDT3					
	M:N	1:1			2:1			1:1			2:1		
SPEA2	LC	10^{-3}	⋫	⋫	10^{-3}	⋫	⋫	10^{-3}	||	⋫	10^{-3}	||	⋫
		▷	10^{-2}	⋫	▷	10^{-2}	⋫	||	10^{-2}	⋫	||	10^{-2}	⋫
		▷	▷	10^{-1}	▷	▷	10^{-1}	▷	▷	10^{-1}	▷	▷	10^{-1}
	HC	10^{-3}	⋫	⋫	10^{-3}	⋫	⋫	10^{-3}	⋫	⋫	10^{-3}	⋫	⋫
		▷	10^{-2}	⋫	▷	10^{-2}	⋫	▷	10^{-2}	⋫	▷	10^{-2}	⋫
		▷	▷	10^{-1}	▷	▷	10^{-1}	▷	▷	10^{-1}	▷	▷	10^{-1}
		Global population-size ratio											
	1:1	---			||			---			||		
	2:1	||			---			||			---		
NSGA-II	LC	10^{-3}	▷	⋫	10^{-3}	▷	⋫	10^{-3}	⋫	⋫	10^{-3}	⋫	▷
		⋫	10^{-2}	⋫	⋫	10^{-2}	⋫	▷	10^{-2}	⋫	▷	10^{-2}	▷
		▷	▷	10^{-1}	▷	▷	10^{-1}	▷	▷	10^{-1}	⋫	⋫	10^{-1}
	HC	10^{-3}	▷	⋫	10^{-3}	||	⋫	10^{-3}	⋫	⋫	10^{-3}	||	▷
		⋫	10^{-2}	⋫	||	10^{-2}	⋫	▷	10^{-2}	⋫	||	10^{-2}	▷
		▷	▷	10^{-1}	▷	▷	10^{-1}	▷	▷	10^{-1}	⋫	⋫	10^{-1}
		Global population-size ratio											
	1:1	---			⋫			---			⋫		
	2:1	▷			---			▷			---		
PESA-II	LC	10^{-3}	||	⋫	10^{-3}	||	⋫	10^{-3}	||	⋫	10^{-3}	||	⋫
		||	10^{-2}	⋫	||	10^{-2}	⋫	||	10^{-2}	⋫	||	10^{-2}	⋫
		▷	▷	10^{-1}	▷	▷	10^{-1}	▷	▷	10^{-1}	▷	▷	10^{-1}
	HC	10^{-3}	||	⋫	10^{-3}	⋫	⋫	10^{-3}	||	⋫	10^{-3}	⊵	⋫
		||	10^{-2}	⋫	▷	10^{-2}	⋫	||	10^{-2}	⋫	⋫	10^{-2}	⋫
		▷	▷	10^{-1}	▷	▷	10^{-1}	▷	▷	10^{-1}	▷	▷	10^{-1}
		Global population-size ratio											
	1:1	---			⋫			---			⊵		
	2:1	⊵			---			⋫			---		

Fig. 4. Preference map for metric *S* and function ZDT2 and ZDT3.

MOEA	M:N	ZDT4						ZDT6					
		1:1			2:1			1:1			2:1		
SPEA2	LC	10^{-3}	$\triangleright$	$\ntriangleright$	10^{-3}	$\triangleright$	$\ntriangleright$	10^{-3}	$\ntriangleright$	$\ntriangleright$	10^{-3}	||	$\ntriangleright$
		$\ntriangleright$	10^{-2}	$\ntriangleright$	$\ntriangleright$	10^{-2}	$\ntriangleright$	$\triangleright$	10^{-2}	$\ntriangleright$	||	10^{-2}	$\ntriangleright$
		$\triangleright$	$\triangleright$	10^{-1}	$\triangleright$	$\triangleright$	10^{-1}	$\triangleright$	$\triangleright$	10^{-1}	$\triangleright$	$\triangleright$	10^{-1}
	HC	10^{-3}	$\triangleright$	$\ntriangleright$	10^{-3}	$\triangleright$	$\ntriangleright$	10^{-3}	$\ntriangleright$	$\ntriangleright$	10^{-3}	||	$\ntriangleright$
		$\ntriangleright$	10^{-2}	$\ntriangleright$	$\ntriangleright$	10^{-2}	$\ntriangleright$	$\triangleright$	10^{-2}	$\ntriangleright$	||	10^{-2}	$\ntriangleright$
		$\triangleright$	$\triangleright$	10^{-1}	$\triangleright$	$\triangleright$	10^{-1}	$\triangleright$	$\triangleright$	10^{-1}	$\triangleright$	$\triangleright$	10^{-1}
		Global population-size ratio											
	1:1	---			$\ntriangleright$			---			$\ntriangleright$		
	2:1	$\trianglerighteq$			---			$\trianglerighteq$			---		
NSGA-II	LC	10^{-3}	$\triangleright$	$\triangleright$	10^{-3}	$\triangleright$	$\triangleright$	10^{-3}	$\trianglerighteq$	$\ntriangleright$	10^{-3}	$\triangleright$	$\trianglerighteq$
		$\ntriangleright$	10^{-2}	||	$\ntriangleright$	10^{-2}	$\triangleright$	$\ntriangleright$	10^{-2}	$\ntriangleright$	$\ntriangleright$	10^{-2}	$\ntriangleright$
		$\ntriangleright$	||	10^{-1}	$\ntriangleright$	$\ntriangleright$	10^{-1}	$\triangleright$	$\triangleright$	10^{-1}	$\ntriangleright$	$\triangleright$	10^{-1}
	HC	10^{-3}	||	$\triangleright$	10^{-3}	$\triangleright$	$\triangleright$	10^{-3}	$\trianglerighteq$	$\ntriangleright$	10^{-3}	$\triangleright$	$\triangleright$
		||	10^{-2}	$\triangleright$	$\ntriangleright$	10^{-2}	$\triangleright$	$\ntriangleright$	10^{-2}	$\ntriangleright$	$\ntriangleright$	10^{-2}	$\ntriangleright$
		$\ntriangleright$	$\ntriangleright$	10^{-1}	$\ntriangleright$	$\ntriangleright$	10^{-1}	$\triangleright$	$\triangleright$	10^{-1}	$\ntriangleright$	$\triangleright$	10^{-1}
		Global population-size ratio											
	1:1	---			$\ntriangleright$			---			$\ntriangleright$		
	2:1	$\trianglerighteq$			---			$\trianglerighteq$			---		
PESA-II	LC	10^{-3}	$\triangleright$	$\trianglerighteq$	10^{-3}	$\triangleright$	||	10^{-3}	$\triangleright$	$\ntriangleright$	10^{-3}	$\trianglerighteq$	$\ntriangleright$
		$\ntriangleright$	10^{-2}	$\ntriangleright$	$\ntriangleright$	10^{-2}	$\ntriangleright$	$\ntriangleright$	10^{-2}	$\ntriangleright$	$\ntriangleright$	10^{-2}	$\ntriangleright$
		$\ntriangleright$	$\triangleright$	10^{-1}	||	$\triangleright$	10^{-1}	$\triangleright$	$\triangleright$	10^{-1}	$\triangleright$	$\triangleright$	10^{-1}
	HC	10^{-3}	$\triangleright$	$\ntriangleright$	10^{-3}	$\triangleright$	$\ntriangleright$	10^{-3}	$\trianglerighteq$	$\ntriangleright$	10^{-3}	$\trianglerighteq$	$\ntriangleright$
		$\ntriangleright$	10^{-2}	$\ntriangleright$	$\ntriangleright$	10^{-2}	$\ntriangleright$	$\ntriangleright$	10^{-2}	$\ntriangleright$	$\ntriangleright$	10^{-2}	$\ntriangleright$
		$\trianglerighteq$	$\triangleright$	10^{-1}	$\trianglerighteq$	$\triangleright$	10^{-1}	$\triangleright$	$\triangleright$	10^{-1}	$\triangleright$	$\triangleright$	10^{-1}
		Global population-size ratio											
	1:1	---			||			---			$\trianglerighteq$		
	2:1	||			---			$\ntriangleright$			---		

Fig. 5. Preference map for metric *S* and function ZDT4 and ZDT6.

Figure 3 shows an excerpt of the whole map for function ZDT6. Note that the numbers in the diagonals indicate the different mutation rates. The table is reading as follows, for assess the effect of *pc* and *pm*, the analyst choose one crossover rates group (LC or HC) and one populations ratio (1:1 or 2:1), then starting *vertically* for one *pm* value and finding the intersection with another *pm* value, he/she has the relationship that stands for this case. For instance, for LC and 1:1, we have $10^{-3} \ntriangleright 10^{-2}$ and $10^{-3} \triangleright 10^{-1}$, and in the same way $10^{-1} \ntriangleright 10^{-2}$ and $10^{-1} \ntriangleright 10^{-3}$. The same procedure is applied for the global M:N, but starting always *vertically*. In this case, $1{:}1 \trianglerighteq 2{:}1$ because in most of the cases the curves for 1:1 are above or at the same level of the curves for 2:1 for each *pm*.

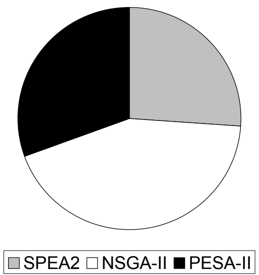
So far, we have not defined what *at the same level* means. It is clear that is not possible to expect numerical equality in the means, but there is space for statistical equality or intervals. Visual inspection of the results is possible in some cases, but impractical if a huge number of functions and metrics are considered during the experimentation and the analysis. Therefore, the preferences maps can be generated automatically employing the lexicographical rules and studying the statistical equivalence of the means for assure the conclusions are according with the preferences statements. Other options is to calculate the difference between means labeling the means as equal if the value belongs to an interval previously defined (by the analyst/DM), or if the amount of the difference is lower than a certain percentage of value between the upper and the lower means of the curves under comparison.

3.3 Extracting information from preferences maps

Figures 4 and 5 report the analysis of the metric *S* for the outcomes produced by SPEA2, NSGA-II and PESA-II over functions ZDT2, ZDT3, ZDT4 and ZDT6.

Note that the maps may help in selecting the proper parameters for a particular method, providing information about what dimension in *pm* and which population size ratio are preferable. However, more information can be obtained from the qualitative classification introducing weights as numerical equivalence of the relationships. For instance, the sensitivity to population size ratio is assessed assigning $\triangleright$=3, $\trianglerighteq$=2, and $\|$ =1 and counting the occurrences of each preference relationship in both maps (in *global population-size ratio* section), then a weighted sum is calculated. Figure 6 present the results. Notice that SPEA2 is less sensible to population size ratio than the other methods in terms of hypervolume covered (*S*), while NSGA-II is the most sensitive

An assessment of the sensitivity to the crossover rate is also possible. In this case, the number of changes (jumps) in the preference relationships when *pc* changes from LC to HC is tabulated in figure 7b. The jumps are labeled as short jumps when the preference relationship changes to an immediate one (e.g. from $\triangleright$ to $\trianglerighteq$), otherwise they are labeled as long jumps. Short jumps are weighted by 1 and long jumps by 2. The weighted sum is then presented graphically in Fig. 7a and numerically in Fig 7b. Note that all MOEAs are less or equal sensitive to *pc* when M:N ratio is 1:1.

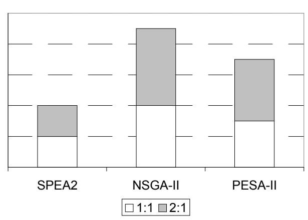


Preference Relationship	Weight	SPEA2	NSGA-II	PESA-II
⊳	3	0	2	0
⊵	2	2	2	3
||	1	2	0	1
Total (mean)		**1.5**	**2.5**	**1.75**

(a) (b)

Fig. 6. Sensitivity of metric *S* to M:N, for SPEA2, NSGA-II and PESA-II.

Notice in Fig 7 that there is a hierarchy of sensitivity to *pc*, which could lead to conclude SPEA2 ⊳ PESA-II ⊳ NSGA-II for the studied feature. In general, for any analyst/DM interested in selecting an algorithm, starting from the outcomes and calculating the metrics values with their statistics, preferences maps may be build according to the DM's general preferences. Then, by analyzing the maps, sensitivity information can be extracted. Finally, for making a decision about the algorithm to choose, a new structure of preferences must be expressed, indicating if-then-else sentences (if lexicographically) or weights (for weighted sum), that will identify the proper method. For example, if the DM's needs point to a search performed with small populations and as less sensitive as possible to *pc*, by examining the figures 6 and 7 the analyst can conclude that SPEA2 is the more appropriate to satisfy DM's desires. Nevertheless, the conclusions depend on the needs, preferences, aspiration levels, etc, expressed by the DM.



Function	Jump	Value	SPEA2		NSGA-II		PESA-II	
			1:1	2:1	1:1	2:1	1:1	2:1
ZDT2	short	1	0	0	0	0	0	0
	long	2	0	0	0	1	0	1
ZDT3	short	1	0	0	0	0	0	1
	long	2	1	1	0	1	0	0
ZDT4	short	1	0	0	0	0	0	1
	long	2	0	0	2	0	1	0
ZDT6	short	1	0	0	0	1	1	0
	long	2	0	0	0	0	0	0
Partials			2	2	4	5	3	4
Total			4		9		7	

(a) (b)

Fig. 7. Sensitivity of metric *S* to *pc* for SPEA2, NSGA-II and PESA-II.

In addition, consider that the maps allow identifying bounds for parameters settings. For instance, by examining figures 4 and 5 and making a census of strictly preference and preference relationships when *pm* changes from one value to another, an analyst can easily conclude that $[10^{-3},10^{-2}]$ is the best interval for fixing *pm* for

ZDT2-ZDT4 and $[10^{-2},10^{-1}]$ for ZDT6, when working with SPEA2 and PESA-II. It is worthy to note these intervals enclose the one-bit mutation rate value.

In summary, the methodology presented allows the construction of a simple and a useful tool that helps the analyst during the decision making modeling. With a small number of general preferences, the methodology guides to the classification of the numerical results from a DM/Analyst perspective (translation the data into preferences). Including more metrics in the preferences maps, it is possible to extract more information, like convergence or diversity/distribution, which can enrich the number of options to consider during the decision making model building.

3 Conclusions

In the presented work we have introduced explicitly the classical notion of the analyst into EMO, remarking the decision making modeling as one of the main steps or stages of a real MDCM problem. Then we have identified some difficulties which take place due to the diversity of values in parameters settings and metrics that can be found when analyzing the comparative studies already published. Finally, we have presented and described a methodology for the incorporation of DM's general preferences during the MOEA selection and the parameters settings. This methodology showed the ability to help the analyst in extracting and exploit the information obtained from comparative studies and translate them into preferences relationships, which can help to choose a particular technique and set the parameters, in a very easy way.

Some remarks of the proposed methodology are:

- An appropriate experimentation was carried out to assess the influence of the crossover and mutation rates, and the population sizes ratio in the behavior of the selected MOEAs.
- Employing a metric for measure covered hypervolume, the means values of several experiments were calculated.
- By means of simple and very general lexicographic rules, the quantitative results were classified in terms of preferences (translating into preferences). Four preference relationships for comparing the results were employed to express the lexicographical rules, then
- Preferences maps were introduced for representing the classification of the results in a compact way. These maps contain all the information generated by the preference translation for each test function.
- Finally, applying very simple techniques of weighting, some information of sensitivity were extracted from the maps. Additionally, brief examples of data extraction for mutation rates and for MOEA selection were presented.

References

1. Roy, B.: A French-English Decision Aiding glossary. Newsletter of the European Working Group "Multicriteria Aid for Decisions". Series 3, nº1, Spring, 2000.

2. Valls, A.: ClusDM: A Multiple Criteria Decision Making Method for Heterogeneous Data Sets. PhD. Thesis. Universitat Politècnica de Catalunya. September 2002.
3. Arsham, H.: Applied Management Science: Making Good Strategic Decisions. 1994. In http://home.ubalt.edu/ntsbarsh/Business-stat/opre/opre640.htm (last visited in October 2004).
4. Powell, D.: Multiobjective Optimization with Genetic Algorithms is Now Considered Mainstream. Evolutionary Computation in Industry. Workshop Proceedings, Tutorials, Late Breaking Papers, and Evolutionary Computation in Industry Track Presentations. Genetic and Evolutionary Computation Conference (GECCO-2004) (CD-ROM) X-CD Technologies. (2004).
5. Van Veldhuizen, D., and Lamont, G.: Multiobjective Evolutionary Algorithms: Analyzing the State-of-the-Art. Evolutionary Computation 8(2) (2000) 125-147.
6. Corne, D.W., Knowles, J.D.: No Free Lunch and Free Leftovers Theorems for Multiobjective Optimization Problems. Evolutionary Multi-Criterion Optimization (EMO 2003), Proceedings of the Second International Conference, Portugal (2003) 327-341.
7. Corne, D.W., Knowles, J.D.: Some Multiobjective Optimizers are Better than Others. Proceedings of the IEEE Congress on Evolutionary Computation. (2003) 2506-2512.
8. Horn, J.: F1.12: Multicriteria Decision Making and Evolutionary Computation. IlliGAL Report No. 9600X. University of Illinois. (1996).
9. Barba-Romero, S., Pomerol, J-C.: Decisiones Multicriterio. Fundamentos Teóricos y Utilización Práctica. Universidad de Alcalá. (1997).
10. Zeleny, M.: Multiple Criteria Decision Making. McGraw-Hill. (1982).
11. Zitzler, E., Deb, K., Thiele, L.: Comparison of Multiobjective Evolutionary Algorithms: Empirical Results. Evolutionary Computation **8**(2) (2000) 173-195.
12. Deb, K., Pratap, A., Agarwal, S., Meyarivan, T.: A Fast and Elitist Multi-Objective Genetic Algorithm: NSGA-II. KanGAL Report No. 200001. Kanpur Genetic Algorithms Laboratory (KanGAL). Indian Institute of Technology (2001).
13. Zitzler, E., Laummans, M., Thiele, L.: SPEA2: Improving the Strength Pareto Evolutionary Algorithm. TIK Report No. 103. Swiss Federal Institute of Technology (ETH). Computer Engineering and Networks Laboratory (TIK). (2001).
14. Corne, D.W., Jerram, N.R., Knowles, J.D., and Oates, M.J.: PESA-II: Region-based Selection in Evolutionary Multiobjective Optimization. Proceedings of the Genetic and Evolutionary Computation Conference (GECCO-2001), Morgan Kaufmann Publishers (2001) 283-290.
15. Khare, V., Yao, X., Deb, K.: Performance Scaling of Multi-objective Evolutionary Algorithms. Evolutionary Multi-Criterion Optimization (EMO 2003), Proceedings of the Second International Conference, Portugal (2003) 376-390.
16. Koch, T.E., and Zell, A.: Multi-Objective Clustering Selection Evolutionary Algorithm. Proceedings of Genetic and Evolutionary Computation Conference (GECCO-2002), Morgan Kaufmann, San Francisco, CA. (2002) 423-430.
17. Ishibuchi, H., and Shibata, Y.: An Empirical Study on the Effect of Matting Restriction on the Search Ability of EMO Algorithm. Evolutionary Multi-

Criterion Optimization (EMO 2003), Proceedings of the Second International Conference, Portugal (2003) 433-447.
18. Wanatabe, S., Hiroyasu, and T. Miki, M.: Multi-objective Rectangular Packing Problem and Its Applications. Evolutionary Multi-Criterion Optimization (EMO 2003), Proceedings of the Second International Conference, Portugal (2003) 565-577.
19. Greiner, D., Galván, B., and Winter, G.: Safety Systems Optimum Design by Multicriteria Evolutionary Algorithms. Evolutionary Multi-Criterion Optimization (EMO 2003), Proceedings of the Second International Conference, Portugal (2003) 722-736.
20. Shu, L-S., Ho, S-J, Ho, S-Y., Chen, J-H, and Hung, M-H.: A Novel Multi-objective Orthogonal Simulated Annealing Algorithm for Solving Multi-objective Optimization Problems with a Large Number of Parameters. Proceedings of Genetic and Evolutionary Computation Conference (GECCO-2004), Springer-Verlag, Germany. (2004) 737-747.
21. Deb, K., and Gupta, N. K.: Optimal Operating Conditions for Overhead Crane Maneuvering Using Multi-objective Evolutionary Algorithms. Proceedings of Genetic and Evolutionary Computation Conference (GECCO-2004), Springer-Verlag, Germany. (2004) 1042-1053.
22. Laumanns, M., Zitzler, E., and Thiele, L.: On The Effects of Archiving, Elitism, and Density Based Selection in Evolutionary Multi-Objective Optimization. En E. Zitzler et al. (eds.): Evolutionary Multi-criterion Optimization (EMO 2001), First International Conference, EMO 2001, Zurich, Switzerland, March 7-9 2001, Proceedings. Lecture Notes in Computer Science Vol. 1993, Springer. (2001) 181-196.
23. Ochoa, G.: Setting the Mutation Rate: Scope and Limitations of the 1/L Heuristics. Proceedings of Genetic and Evolutionary Computation Conference (GECCO-2002), Morgan Kaufmann, San Francisco, CA. (2002) 495-502.
24. Toscano, G., Coello, C.: The Micro Genetic Algorithm 2: Towards Online Adaptation in Evolutionary Multiobjective Optimization. Evolutionary Multi-Criterion Optimization (EMO 2003), Proceedings of the Second International Conference, Portugal (2003) 252-266.
25. Büche, D., Müller, S., Koumoutsatkos, P.: Self-Adaptation for Multi-objective Evolutionary Algorithms. Evolutionary Multi-Criterion Optimization (EMO 2003), Proceedings of the Second International Conference, Portugal (2003) 267-281.
26. Groşan, C.: An Evolutionary Approach for Multiobjective Optimization using Adaptive Representation of Solutions. Late Breaking Papers. Workshop Proceedings, Tutorials, Late Breaking Papers, and Evolutionary Computation in Industry Track Presentations. Genetic and Evolutionary Computation Conference (GECCO-2004) (CD-ROM) X-CD Technologies. (2004).
27. Salazar, D., Galván, B., and Winter, G.: Enhancing A Multiobjective Evolutionary Algorithm Through Flexible Evolution. Late Breaking Papers. Workshop Proceedings, Tutorials, Late Breaking Papers, and Evolutionary Computation in Industry Track Presentations. Genetic and Evolutionary Computation Conference (GECCO-2004) (CD-ROM) X-CD Technologies. (2004).

28. Laumanns, M., Rudolph, G., and Schwefel, H.-P.: Mutation Control and Convergence in Evolutionary Multi-Objective Optimization. In Matousek and Osmera (eds.): Proceedings of the 7th International Mendel Conference on Soft Computing (MENDEL 2001), Czech Republic, 2001. (2001) 24-29.
29. Hansen, P., Jaszkiewicz, A.: Evaluating the quality of approximations of the non-dominated set. Technical Report IMM-Rep-1998-7. Technical University of Denmark, Lyngby, Denmark (1998)
30. Knowles, J.D., Corne, D.W.: On Metrics for Comparing Non-Dominated Sets. In Proceedings of the 2002 Congress on Evolutionary Computation Conference (CEC02), IEEE Press (2002) 711-716
31. Zitzler, E., Laummans, M., Thiele, L. Fonseca, C. M., Grunert da Fonseca, V.: Performance Assessment of Multiobjective Optimizers: An Analysis and Review. IEEE Transactions on Evolutionary Computation **7**(2) (2003) 117-132.
32. Bosman, P., and Thierens, D.: The Balance Between Proximity and Diversity in Multiobjective Evolutionary Algorithms. IEEE Transactions on Evolutionary Computation **7**(2) (2003) 174-188.
33. Farhang-Mehr, A., and Azarm, S.: Minimal Sets of Quality Metrics. Evolutionary Multi-Criterion Optimization (EMO 2003), Proceedings of the Second International Conference, Portugal (2003) 405-417.