MILP-based clustering method for multi-objective optimization: Application to environmental problems

Diego Gabriel Oliva, Gonzalo Guillén-Gosálbez *, Josep Maria Mateo-Sanz, Laureano Jiménez-Esteller

Departament d'Enginyeria Química, Universitat Rovira i Virgili, Av. Països Catalans 26, 43007 Tarragona, Spain

A B S T R A C T

Multi-objective optimization (MOO) has recently emerged as a useful technique in environmental engineering. One major limitation of this approach is that its computational burden grows rapidly with the number of environmental objectives, which causes difficulties regarding the computation and visualization of the Pareto solutions. In this work we present several theoretical and algorithmic developments for grouping environmental objectives into clusters on the basis of which the multi-objective optimization can be performed, thereby facilitating the computation and analysis of the Pareto solutions. Our method is based on a novel mixed-integer linear program (MILP) that identifies in a systematic manner groups of objectives that behave similarly. We test the capabilities of our approach using several examples, in which we compare it against other well-known clustering methods.

© 2013 Elsevier Ltd. All rights reserved.

1. Introduction

In the last years, environmental engineering has gained wider interest (Jacquemin, Pontalier, & Sablayrolles, 2012; Pieragostini, Mussati, & Aguirre, 2012). The ultimate goal in this field is to mitigate the impact that human activities cause in the environment. Among the tools available to accomplish this, multi-objective optimization (MOO) has emerged as a useful technique, as it allows treating environmental aspects as additional objectives to be optimized rather than as constraints imposed on the system (Brunet, Guillén-Gosálbez, & Jiménez, 2012; Kostin, Guillén-Gosálbez, Mele, & Jiménez, 2012; Salcedo, Antipova, Boer, Jiménez, & Guillén-Gosálbez, 2012). This permits identification of solutions where significant environmental savings can be attained at a marginal increase in cost.

Unfortunately, the need to consider several environmental objectives in the analysis leads to complex multi-objective models whose solutions are difficult to generate and interpret. Objective reduction methods allow reducing the computational burden of MOO models by omitting redundant objectives. In a seminar work, Deb and Saxena (2005) proposed a method for reducing the number of objectives in MOO based on principal component analysis (PCA). Pozo, Ruíz-Femenia, Caballero, Guillén-Gosálbez, & Jiménez (2012) studied also the use of Principal Component Analysis (PCA) to identify redundant environmental metrics in MOO that can be omitted without disturbing the main features of the problem, thereby reducing the associated complexity. They showed that, besides its numerical usefulness, the use of PCA coupled with MOO provides valuable insights on the relationships between environmental indicators of concern for decision-makers.

Brockhoff and Zitzler (2006a) were the first to propose a rigorous definition of conflict between objectives and define necessary and sufficient conditions under which certain objectives can be omitted. They also introduced the problem of finding a minimum subset of objectives that does not exceed a given approximation error and proposed two algorithms to tackle such a problem.

Zhou, Zheng, Li, and Lv (2009) developed a large dimensionality reduction algorithm based on the least square method. This algorithm fits every objective function to a line and compares the corresponding slopes to identify redundant objectives.

Singh, Isaacs, and Ray (2011) proposed another method for dimensionality reduction based on the analysis of a representative set of solutions produced by a conventional multi-objective evolutionary algorithm (MOEA). Their method, termed Pareto corner search evolutionary algorithm (PCSEA), searches for the corners of the Pareto front rather than for the complete Pareto front. The solutions obtained by means of this algorithm are then used for dimensionality reduction in order to identify relevant objectives.

* Corresponding author.
  E-mail addresses: diegogabriel.oliva@urv.cat (D.G. Oliva), gonzalo.guillen@urv.cat (G. Guillén-Gosálbez), josepmaria.mateo@urv.cat (J.M. Mateo-Sanz), Laureano.Jimenez@urv.cat (L. Jiménez-Esteller).

Nomenclature

Sets/indices

C	set of clusters indexed by c
J	set of combinations indexed by j
O	set of objectives indexed by o
F	set of objectives in the original problem
F'	set of objectives in the reduced space
$OJ(j,k)$	set of objectives contained in cluster k in combination j

Parameters

error (j)	approximation error associated with combination j
k	desired number of clusters in the problem
n	number of objectives in the set F
δ error	approximation error

Variables

error$v$	continuous variable that computes the maximum error in an active clustering option
$x(o,c)$	binary variable that denotes whether objective o belongs to cluster c in an active clustering option or not
$z(j)$	binary variable that takes a value of one if the approximation error associated with combination j is accounted for and zero otherwise

Guillén-Gosálbez (2011) proposed a systematic method for reducing the number of objectives in MOO that relies on a mixed-integer linear program (MILP) that can be solved efficiently by standard branch and bound techniques. Based on this work, Copado-Méndez, Guillén-Gosálbez, and Jiménez (2012) introduced a decomposition algorithm for reducing the number of objectives in MOO that exploits the mathematical structure of the aforementioned MILP.

The contributions mentioned above focused on reducing the number of objectives in MOO. In contrast, the analysis of objectives with similar features in the context of an MOO problem and their classification into clusters have both received much less attention to date. Clustering tools are useful when data reduction is required. They allow grouping data with similarities into clusters, which simplifies the analysis of a system. While there are many clustering algorithms available in the literature, very few have been applied to MOO problems. Taboada and Coit (2008) applied a k-means clustering algorithm to identify promising Pareto solutions of a multi-objective scheduling problem. In another context, Gutiérrez, Lozano, Moreira, and Feijoo (2010) applied PCA and clustering methods to uncover and visualize the structure of large multidimensional data sets in two applications: the environmental assessment of wastewater treatment plants, and the cultivation, processing and consumption of mussels.

In this work, we study how to group objectives of an MOO problem into meaningful clusters with the property that the minimization of any objective within a cluster will result in the minimization of the rest of objectives within the same cluster. Particularly, we present advances in the following fronts. First, we introduce the concept of clustering error in a MOO problem to quantify how the original Pareto structure of a problem changes when we perform the optimization in a reduced domain of objectives defined according to some clustering. Then, we present a rigorous MILP that identifies in a systematic manner clusters of similar objectives in MOO models. This MILP makes use of the concept of approximation error introduced by Brockhoff and Zitzler (2006a). Third, we perform an extensive computational analysis of

our clustering algorithm, comparing it with well known statistical clustering methods (Jain, Murty, & Flynn, 1999; Mucha, Bartel, & Morales-Merino, 2013).

Our final goal is to aid decision-makers in complex MOO problems with a large number of environmental objectives. Our clustering method provides valuable insight into the problem and facilitates the calculation and analysis of the Pareto solutions.

2. Mathematical background

We provide first some general concepts on MOO (further details can be found in Brockhoff and Zitzler (2006a)). We consider the following MOO problem MO(X) defined as follows:

$$MO(X) = \min_x \{f(x) := (f_1(x), ..., f_n(x)) : x \in X\} \quad (1)$$

with n objective functions $f_i := X \to R$, $1 \le i \le n$, where the vector function $f := (f_1, ..., f_n)$ maps each solution $x \in X$ to an objective vector $f(x) \in R^n$.

The underlying dominance structure considered here (assuming that we aim to minimize all of the objectives simultaneously) is given by the weak Pareto dominance relation which is defined as follows: $\preceq F' := \{(x, y)|x, y \in X \land \forall f_i \in F' : f_i(x) \le f_i(y)\}$, where F' is a set of objectives with $F' \subseteq F := \{f_1, ..., f_n\}$. Our analysis will use the following definitions:

Definition 1. x weakly dominates y with respect to the objective set $F'(x \preceq F' y)$ if $(x, y) \in \preceq F'$.

Definition 2. $x* \in X$ is called Pareto optimal if there is no other $x \in X$ that dominates x^* with respect to the set of all of the objectives.

3. Problem statement

The problem we aim to solve can be formally stated as follows. We are given an MOO problem that optimizes a set F of n objectives. Our goal is to group the n objectives into k subgroups (i.e., clusters) with the property that the simultaneous minimization of one member of each cluster will produce the same (or very similar) results as the minimization of the original set of objectives.

Our clustering method groups objectives according to their similarity and so as to minimize an error metric. To define this clustering error, we make use of the concept of approximation error introduced by Brockhoff and Zitzler (2006a). To clarify the fundamentals of this approximation error, we introduce next a simple example that considers 3 Pareto solutions (x1, x2 and x3) and 4 objectives (f1, f2, f3 and f4). Fig. 1 is a parallel coordinates plot that
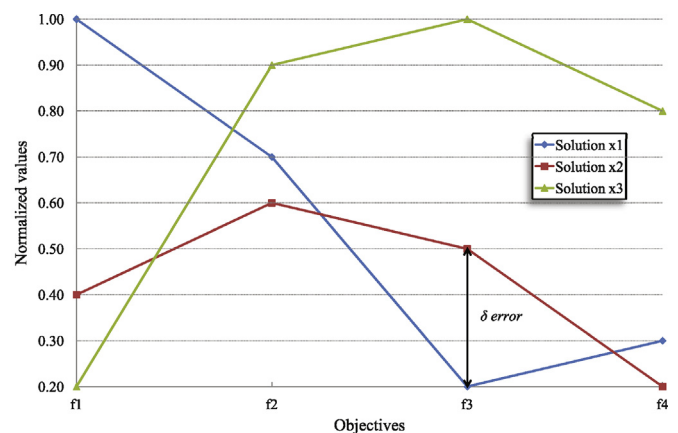

Fig. 1. Pareto solutions. The Pareto structure of the problem changes when objective f3 is omitted, since solution x2 dominates x1 in the reduced domain but not in the original one.
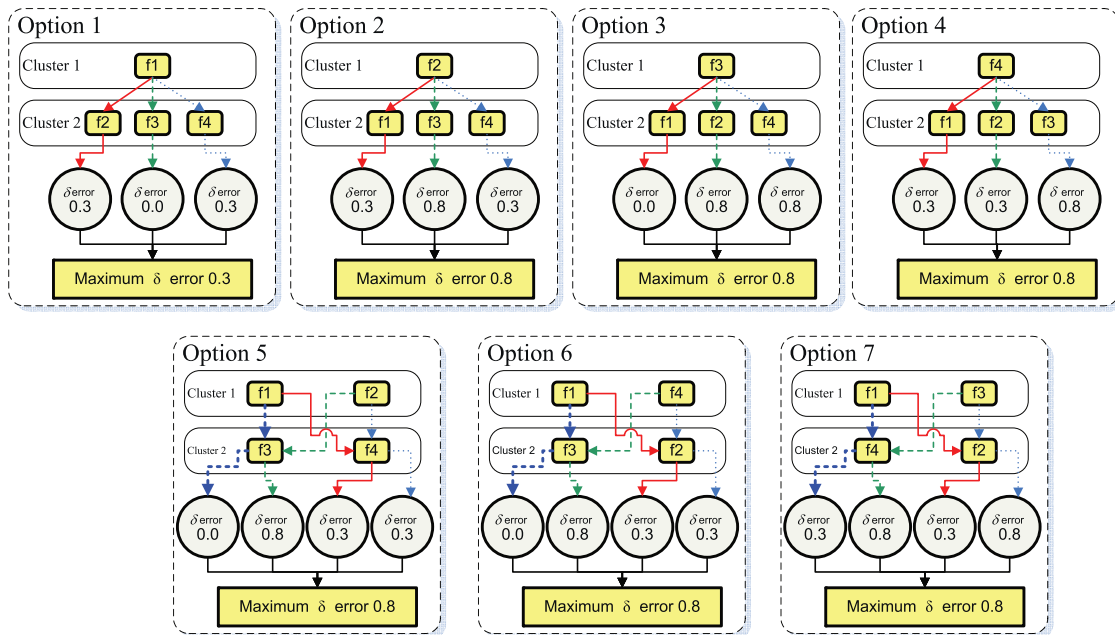
Fig. 2. Possible combinations of two objectives and associated δ error.

depicts in the horizontal axis the set of objectives and in the vertical axis, the normalized value of each objective. As observed, every line in the plot represents a single solution, and all of the lines intersect in at least one point, since no solution dominates any of the others.

As seen, it is possible to omit objective f4 from the set while still keeping the dominance structure. That is, all the solutions that are optimal in the original set (f1, f2, f3 and f4), are also optimal in the reduced set (f1, f2 and f3). If we omit objective f3 instead of f4, then the dominance structure changes, since solution x2 dominates solution x1 in the reduced set (f1, f2, f4), but not in the original one. Hence, solution x1 will be lost if we optimize in the reduced domain.

Note that if solutions x1 and x2 showed the same value in objective f3, then the dominance structure would be kept in both cases (sets f1, f2, f4 and f1, f2, f3, f4). The amount that we have to subtract from the score of solution x2 in the omitted objective such that this solution dominates x1 not only in the original set but also in the reduced one is called δ error, and it is used to quantify the difference between the dominance structure of both sets. A large error means that there is a big change in the dominance structure, while a small delta error implies the converse. The reader is referred to the works by Brockhoff and Zitzler (2006a) and Guillén-Gosálbez (2011) for further mathematical details on this metric. In this work, we make use of this approximation error to develop our clustering algorithm.

Particularly, to determine the error of a given classification of objectives into clusters, we proceed as follows. We first determine the set of combinations of objectives obtained by selecting one criterion from every cluster (i.e., group of objectives), and for each of them we calculate the associated approximation error. The error of the clustering alternative (i.e., groups of objectives) is then defined as the maximum approximation error over all these combinations of objectives. The motivation for this procedure is that once the error has been calculated, we can guarantee that if we choose one objective from each cluster and optimize in the resulting reduced domain, the approximation error will not exceed the one calculated following the above-mentioned procedure.

To further clarify this concept, we revisit the example presented before. Suppose we want to group the objectives into a desired number of clusters. Fig. 2 shows all the combinations of two clusters and the associated δ error. Note that clustering option 1 is the

best one, since the clustering error is the minimum possible. Fig. 3 depicts an example on how to calculate the δ error for all possible combinations of objectives associated with clustering option 2. For three clusters, the best clustering option is number 2 (see Fig. 4). Note that the error of the clustering decreases as we increase the number of clusters considered.

4. Proposed approach

We present next a rigorous approach to cluster objectives in MOO models that is based on the δ error definition introduced by Brockhoff and Zitzler (2006a). Fig. 5 summarizes our approach, which comprises five mayor steps.

Assume we want to group n objectives of an MOO model into k clusters. The first step to take is to generate a set of Pareto solutions of the MOO model on the basis of which the clustering will be performed. To this end, we can use any MOO algorithm available in the literature.

After this, we define the number of clusters we want to generate (let us say k clusters), and then generate all of the k-combinations of n objectives (all subsets of k distinct elements of the original
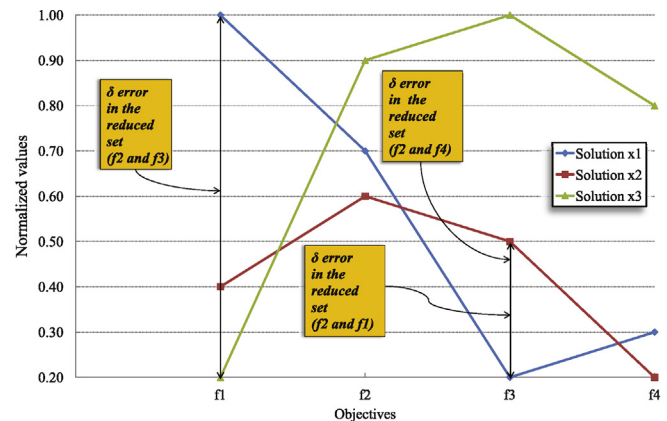

Fig. 3. Graphical representation of the δ error for clustering option 2.
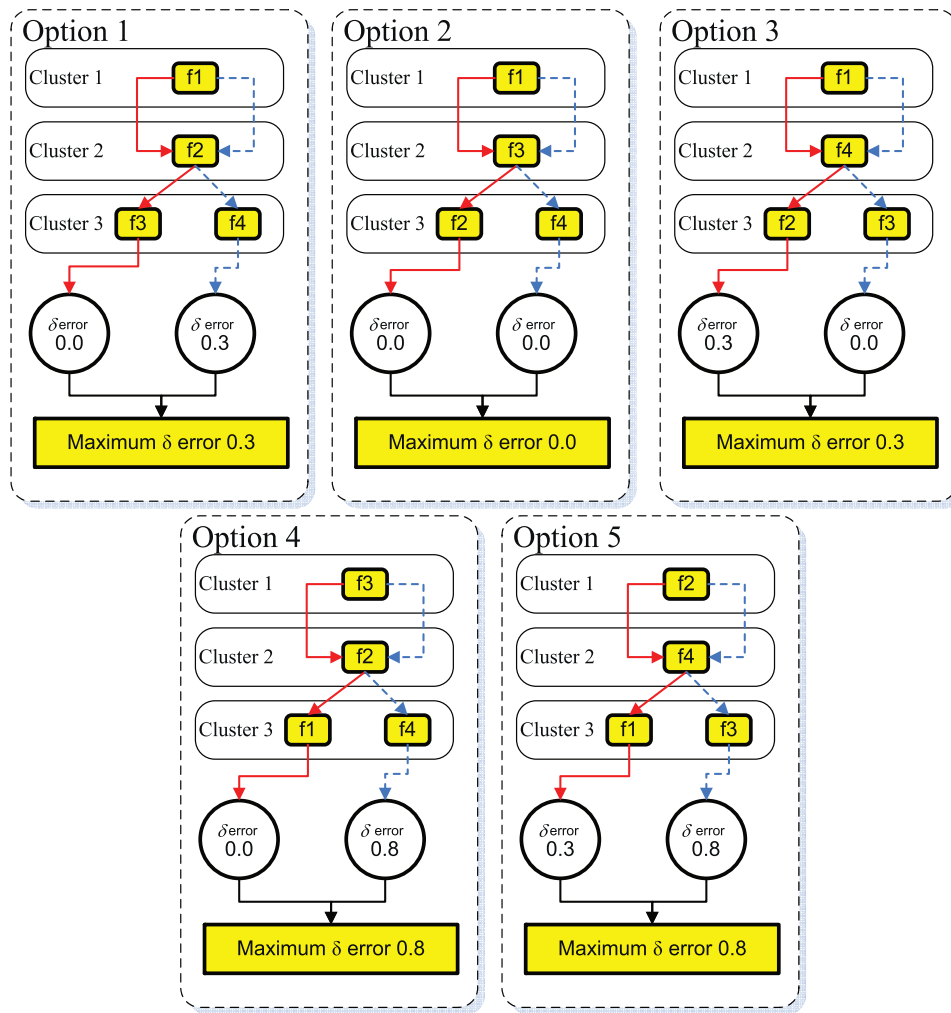
Fig. 4. Possible combinations of three objectives and associated δ error.

set). As an example, Table 1 shows all of the possible combinations for the first example described before. Note that if the problem has n elements, then the number of *k*-combinations is equal to the binomial coefficient $\binom{n}{k}$.
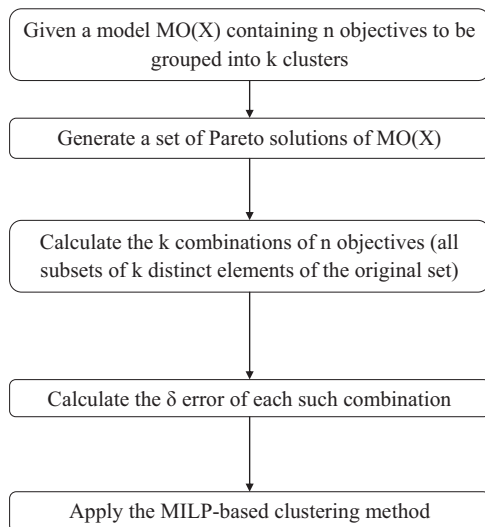
Fig. 5. Algorithm.

We then calculate the approximation error (δ error) that would be obtained if the optimization was conducted in the reduce space corresponding to each combination of objectives. To this end, we can either use the exhaustive algorithm proposed by Brockhoff and Zitzler (2006a), which is available online free of charge (Brockhoff and Zitzler, 2006a, 2006b, 2006c, 2007a, 2007b, 2007c, 2009a, 2009b; Brockhoff, Saxena, Deb, & Zitzler, 2007; Brockhoff, 2009, 2010), or in general any other algorithm of the same type, like the one introduced by Guillén-Gosálbez (2011). The results obtained from this analysis as applied to the first example are presented in Table 2.

This information will be used to construct an MILP model that will identify in the best manner to group objectives into clusters (see Supplementary Material for further details).

Table 1
All possible combinations to classify the 4 objectives of the first example in 3 clusters (i.e., 3-element subsets of the original set containing 4 elements).

Combination id	Cluster		
	1	2	3
1	f1	f2	f3
2	f1	f2	f4
3	f1	f3	f4
4	f2	f3	f4

Table 2
Error calculated using the external java code developed by Brockhoff and Zitzler (2006a). The delta error is calculated for each combination of objectives.

Combination id	Delta error
1	0.00
2	0.30
3	0.00
4	0.80

Supplementary material related to this article can be found in the online version at http://dx.doi.org/10.1016/j.compchemeng.2013.05.016.

The MILP we propose makes use of binary variable $x(o, c)$, which is one if objective o belongs to cluster c and zero otherwise. Every objective must belong to at least one cluster, a condition that is enforced via the following constraint:

$$\sum_{c=1}^{k} x(o, c) = 1 \quad \forall o \tag{2}$$

Furthermore, every cluster must contain at least one objective. This is accomplished by adding the following equation:

$$\sum_{o=1}^{n} x(o, c) \geq 1 \quad \forall c \tag{3}$$

The error of a solution is calculated via Eq. (4), which accounts for the approximation error associated with a given combination of objectives provided such a combination is active in the clustering solution being assessed:

$$1 - k + \sum_{c=1}^{k} \sum_{o \in OJ(j,k)} (x(o, c)) \leq z(j) \quad \forall j \tag{4}$$

In this equation, $z(j)$ is a binary variable that will take the value of 1 if combination j belongs to the active clustering (i.e., to the clustering alternative encoded by the binary solution), and 0 otherwise, while $OJ(j, k)$ is the set of objectives contained in cluster k in combination j. Hence, the error associated with combination j will be accounted for if and only if it emerges from the active clustering. Note that in this case the summation of x will be equal to k, so the left hand side of the inequality will be 1, which will force z to take a value of 1 as well. Conversely, if the classification of objectives into clusters is not the same as in combination j, then the summation will be less or equal than zero. Note that since the error is being minimized, the model will force variable z to take a zero value in this last case. Table 3 illustrates how this equation would be applied for the case of 2 clusters in the example presented before.

Table 4
Values of error(j) for two clusters.

Combination j	OJ(j,k)		error(j)
	$k = 1$	$k = 2$	
1	f1	f2	0.3
2	f2	f1	0.3
3	f1	f3	0.0
4	f3	f1	0.0
5	f1	f4	0.3
6	f4	f1	0.3
7	f2	f3	0.8
8	f3	f2	0.8
9	f2	f4	0.3
10	f4	f2	0.3
11	f3	f4	0.8
12	f4	f3	0.8

Finally, the error of the active clustering is calculated as follows:

$$\text{error}\nu \geq z(j)\text{error}(j) \tag{5}$$

where error(j) is the approximation error of combination j, which is calculated according to the definition by Brockhoff and Zitzler (2006a), while errorν is a continuous variable denoting the maximum approximation error in the active clustering option. Table 4 illustrates the values of error(j) for the case of 2 clusters in the example presented before.

The optimization problem is finally expressed in compact form as follows:

$$\begin{aligned} \min \quad & \text{error}\nu \\ \text{s.t.} \quad & \text{constraints} \ (2)\text{–}(5) \end{aligned} \tag{6}$$

This model provides as output a clustering solution where the objectives are distributed among k clusters.

In this work, we propose to use this model iteratively in order to identify clustering alternatives with different approximation errors. To this end, after solving the model, we add an integer cut that exclude from the search space the solution identified in the last iteration. Hence, a new integer cut is added every time the model is solved, and we repeat this procedure iteratively until the necessary number of clustering solutions is identified. Hence, let $x(o, s)^{it}$ be the binary solution obtained in iteration it after solving model (6). The following integer cut is added to exclude this solution from the

Table 3
Illustrative example of Eq. (4) for two clusters.

Combination j	OJ(j,k)		Eq. (4)	Values of the binary variable $z(j)$						
	$k = 1$	$k = 2$		Option 1	Option 2	Option 3	Option 4	Option 5	Option 6	Option 7
1	f1	f2	$1 - 2 + (x(f1,\text{cluster } 1) + x(f2,\text{cluster } 2)) \leq z(1)$	1	0	0	0	0	1	1
2	f2	f1	$1 - 2 + (x(f2,\text{cluster } 1) + x(f1,\text{cluster } 2)) \leq z(2)$	0	1	0	0	0	0	0
3	f1	f3	$1 - 2 + (x(f1,\text{cluster } 1) + x(f3,\text{cluster } 2)) \leq z(3)$	1	0	0	0	1	1	0
4	f3	f1	$1 - 2 + (x(f3,\text{cluster } 1) + x(f1,\text{cluster } 2)) \leq z(4)$	0	0	1	0	0	0	0
5	f1	f4	$1 - 2 + (x(f1,\text{cluster } 1) + x(f4,\text{cluster } 2)) \leq z(5)$	1	0	0	0	1	0	1
6	f4	f1	$1 - 2 + (x(f4,\text{cluster } 1) + x(f1,\text{cluster } 2)) \leq z(6)$	0	0	0	1	0	0	0
7	f2	f3	$1 - 2 + (x(f2,\text{cluster } 1) + x(f3,\text{cluster } 2)) \leq z(7)$	0	1	0	0	1	0	0
8	f3	f2	$1 - 2 + (x(f3,\text{cluster } 1) + x(f2,\text{cluster } 2)) \leq z(8)$	0	0	1	0	0	0	1
9	f2	f4	$1 - 2 + (x(f2,\text{cluster } 1) + x(f4,\text{cluster } 2)) \leq z(9)$	0	1	0	0	1	0	0
10	f4	f2	$1 - 2 + (x(f4,\text{cluster } 1) + x(f2,\text{cluster } 2)) \leq z(10)$	0	0	0	1	0	1	0
11	f3	f4	$1 - 2 + (x(f3,\text{cluster } 1) + x(f4,\text{cluster } 2)) \leq z(11)$	0	0	1	0	0	0	1
12	f4	f3	$1 - 2 + (x(f4,\text{cluster } 1) + x(f3,\text{cluster } 2)) \leq z(12)$	0	0	0	1	0	1	0

search space in subsequent iterations:

$$\sum_{(o,c)\in M_{it}} x(o,c)^{it} - \sum_{(o,c)\in Q_{it}} x(o,c)^{it} \leq |M_{it}| - 1$$

$M_{it} = \{(o,c)|x(o,c)^{it} = 1 \text{ for the solution obtained in iteration it}\}$

$Q_{it} = \{(o,c)|x(o,c)^{it} = 0 \text{ for the solution obtained in iteration it}\}$

(7)

In the following section, we present examples taken from previous works and perform comparisons between our MILP and a standard statistical clustering method.

5. Numerical examples

The capabilities of the approach presented are tested through its application to several case studies, in which we compare the results of the MILP with those produced by a statistical method. The statistical method consists of a standard clustering algorithm. This algorithm is applied to the Spearman correlation matrix obtained from a set of Pareto solutions of the original problem (Lehmann, 1998; Seber, 2004; Späth, 1985). In essence, the statistical method relies on a hierarchical clustering. To perform this clustering, we use Matlab (MathWorks, n.d.), and more precisely the available functions *pdist, linkage and dendrogram*, which generate a dendrogram plot of the hierarchical cluster tree (Mucha et al., 2013). A dendrogram contains many U-shaped lines connecting objects in a hierarchical tree. The height of each U-shaped line represents the distance between the two objects being connected. The number of clusters depends on the error that the decision maker is willing to accept.

The MILPs were implemented in GAMS 23.5.2 and solved with CPLEX 12.2.0.0 (Brooke, 1997) on a computer with an AMD Phenom 8600B Triple-Core Processor working at 2.29 GHz and 1.75 Gb of RAM. The time spent in solving the MILPs varied depending on the instance being solved, but was always in the range of 1–2 CPU seconds.

The first case study deals with the design of petrochemical supply chains considering several environmental indicators simultaneously (Pozo et al., 2012). A set of 16 Pareto solutions were generated using the epsilon constraint method. These solutions were normalized before applying the clustering method. For the case of the net present value (NPV), we use the following equation:

$$nf_o(s) = \frac{\overline{f_o} - f_o(s)}{\overline{f_o}} \quad \forall s, o = \text{NPV}$$

(8)

where $\overline{f_o}$ denotes the maximum value of objective o among all of the Pareto solutions, $f_o(s)$ is the value of objective o in the Pareto solution s, and $nf_o(s)$ represents the normalized value of objective o in solution s. The normalized values of the environmental indicators were calculated as follow:

$$nf_o(s) = \frac{f_o(s) - \underline{f_o}}{\overline{f_o}} \quad \forall s, o \neq \text{NPV}$$

(9)

where $\underline{f_o}$ is the minimum value of objective o among all of the Pareto solutions.

Table 5 shows the normalized Pareto solutions obtained following this procedure, while Fig. 6 depicts the associated parallel coordinates plot. At a first sight, this plot suggests that objectives 2 and 3 are equivalent and should therefore be grouped into the same cluster.

As this example has 4 objectives, we consider only 2 and 3 clusters. Table 6 displays the clustering solutions obtained with the rigorous MILP method. As observed, objectives 2 and 3 are highly correlated. Using the dendrogam approach, we obtain the same results as those produced by the rigorous MILP method (see Fig. 7). Table 7 shows the square matrix used to obtain the dendrogram, which considers distances between objectives based on the Spearman correlation. In the comparison and hereinafter, we use the best clustering option produced by the MILP (i.e., the one with the minimum error) as reference solution in all of the cases. The comparison between methods is shown in Table 8. Note that the results associated with option 1 produced by the rigorous method are the same as those obtained with the dendrogram approach. Indeed, option 1 (for three clusters) is the same solution as the one obtained by Pozo et al. (2012) using the PCA method.

The second case deals with the design of bio-ethanol supply chains in Argentina considering 6 objectives (Kostin et al., 2012). Table 9 displays the normalized Pareto solutions where the objectives have been numbered for simplicity, while Fig. 8 depicts the associated parallel coordinates plot. These solutions were generated via the epsilon constraint method, and normalized in the same manner as in the first example.

Table 10 shows the results of the MILP for the cases of 2, 3 and 4 clusters. Note that for the case of 2 clusters, there are only 2 meaningful clustering options. Particularly, the second clustering option gives a clustering error equal to 1, which is the maximum error possible. This implies that every clustering option identified in subsequent iterations will lead to the same maximum clustering error, so it can be neglected for practical purposes. For 3 and 4 clusters, the maximum clustering error is 0.08 and 0, respectively. Similarly, as in the previous example, it is possible to reduce the

Table 5
Normalized Pareto solutions for the first example (petrochemical supply chains).

Pareto solution	Objective			
	1. Net present value	2. Ecosystem quality	3. Human health	4. Damage to natural resources
1	0.1650	0.0000	0.0000	0.0000
2	0.0840	0.0160	0.0171	0.0157
3	0.0751	0.0177	0.0191	0.0268
4	0.0749	0.0178	0.0192	0.0269
5	0.0554	0.0288	0.0310	0.0313
6	0.0353	0.0355	0.0382	0.0462
7	0.0352	0.0356	0.0382	0.0466
8	0.0332	0.0409	0.0436	0.0470
9	0.0131	0.0533	0.0563	0.0627
10	0.0129	0.0533	0.0562	0.0630
11	0.0129	0.0541	0.0573	0.0627
12	0.0121	0.0543	0.0573	0.0636
13	0.0129	0.0543	0.0575	0.0627
14	0.0046	0.0710	0.0758	0.0707
15	0.0044	0.0715	0.0764	0.0709
16	0.0000	0.0888	0.0955	0.0783
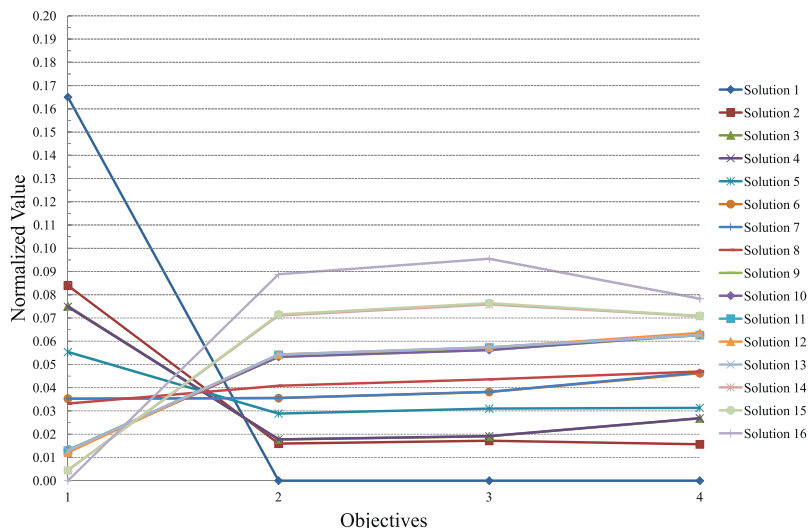
Fig. 6. Parallel coordinates plot for the first example (petrochemical supply chains).

Table 6
Results of the rigorous MILP method for the first example (petrochemical supply chains).

Number of clusters	Cluster	Objectives inside the clustering option					
		Option 1	Option 2	Option 3	Option 4	Option 5	Option 6
2	1	1	4	2	3	1,2	3,2
	2	2,3,4	1,2,3	1,3,4	1,2,4	3,4	1,4
	Clustering error	0.010	0.170	0.170	0.170	0.170	0.170
3	1	4	3	4	1	3	4
	2	1	1	3	2	2	2
	3	2,3	2,4	1,2	3,4	1,4	1,3
	Clustering error	0.000	0.010	0.170	0.170	0.170	0.170
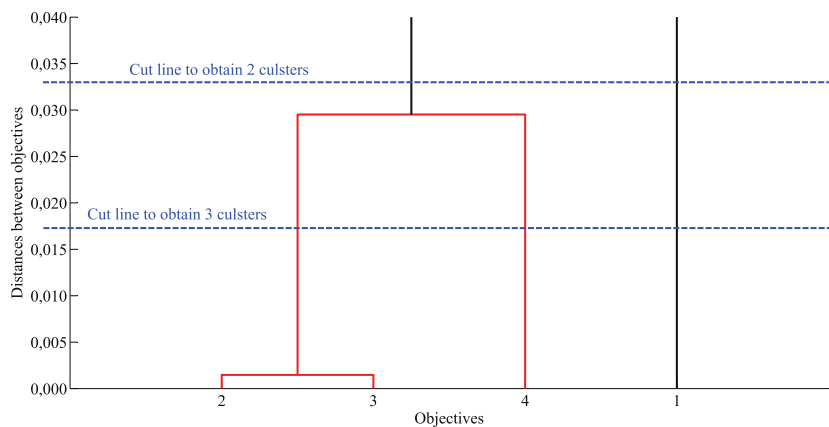
Fig. 7. Dendrogram scheme for the first example (petrochemical supply chains).

number of objectives while still preserving to a large extent the Pareto structure of the problem.

Fig. 9 shows the results obtained with the statistical method, while Table 11 displays the square matrix employed in the generation of the dendrogram, which contains the distances between objectives using the Spearman metric. We consider again 2, 3 and 4 clusters using 3 cut lines in the dendrogram scheme. Note that for this example both methods produce again the same results.

Table 7
Distances between objectives (one minus the Spearman correlation) for the first example (petrochemical supply chains).

Objectives	Objectives			
	1	2	3	4
1	0.0000	1.9919	1.9845	1.9912
2	1.9919	0.0000	0.0015	0.0258
3	1.9845	0.0015	0.0000	0.0332
4	1.9912	0.0258	0.0332	0.0000

Table 8
Comparison between the rigorous MILP and the statistical method for the first example (petrochemical supply chains).

Number of clusters	Cluster	Objectives inside the cluster	
		MILP methodology	Dendrogram approach
2	1	1	1
	2	2,3,4	2,3,4
	Clustering error	0.010	0.010
3	1	4	4
	2	1	1
	3	2,3	2,3
	Clustering error	0.000	0.000

The third example addresses the multi-objective optimization of hydrogen supply chains for vehicle use in Spain considering several Life Cycle Assessment impacts (Sabio, Kostin, Guillén-Gosálbez, & Jiménez, 2012). Table 12 shows the Pareto solutions produced by applying the epsilon constraint method to the MILP introduced in Sabio et al. (2012) considering eight objectives. Fig. 10 depicts the parallel coordinates plot of these Pareto solutions, which were normalized by dividing each objective by its maximum value attained over all the solutions. Specific details on the case study, including the type of environmental indicators used have been omitted here due to space limitations but can be found in the original publications.

We applied our MILP-based clustering method to the set of Pareto solutions considering 2, 3 and 4 clusters. Table 13 displays the results obtained. Note that for simplicity we show only the first 16 clustering alternatives identified by the algorithm. For 2 clusters, there are different alternatives with the same clustering error (0.68). For 3 clusters, option 1 is the optimum, with a clustering error equal to 0. For 4 clusters, we identify several clustering options with a clustering error equal to 0. As observed, significant reductions in the number of objectives are possible by applying our systematic approach.

We applied next the statistical method to the same problem. The results of this approach are depicted in the dendrogram tree of

Table 9
Normalized Pareto solutions for the second example (bio-ethanol supply chain in Argentina).

| Solution	Objective					
	1. Net present value	2. GWP100 Global warming potential	3. Eco-indicator 99	4. Damage to human health	5. Damage to eco-system quality	6. Damage to resources
1	1.0000	1.0000	0.0000	0.0000	0.0000	0.0000
2	1.0000	1.0000	0.0000	0.0000	0.0000	0.0000
3	1.0000	1.0000	0.0000	0.0000	0.0000	0.0000
4	1.0000	1.0000	0.0000	0.0000	0.0000	0.0000
5	1.0000	1.0000	0.0000	0.0000	0.0000	0.0000
6	0.9999	1.0000	0.0000	0.0000	0.0000	0.0000
7	0.6340	0.7239	0.1681	0.1679	0.1689	0.1667
8	0.6340	0.7309	0.1749	0.1749	0.1758	0.1667
9	0.6204	0.6936	0.1690	0.1684	0.1696	0.1797
10	0.6199	0.6928	0.1694	0.1688	0.1700	0.1802
11	0.6191	0.6917	0.1700	0.1694	0.1707	0.1809
12	0.5122	0.4609	0.3309	0.3304	0.3325	0.3333
13	0.4740	0.5308	0.3380	0.3380	0.3392	0.3294
14	0.4735	0.5303	0.3388	0.3388	0.3400	0.3300
15	0.4734	0.5303	0.3389	0.3389	0.3401	0.3300
16	0.4683	0.5373	0.3545	0.3549	0.3560	0.3333
17	0.4369	0.4062	0.3388	0.3363	0.3384	0.4060
18	0.4360	0.4049	0.3396	0.3372	0.3392	0.4071
19	0.4341	0.4020	0.3413	0.3389	0.3409	0.4094
20	0.3697	0.2813	0.4571	0.4553	0.4575	0.5000
21	0.3346	0.3431	0.5076	0.5075	0.5088	0.5000
22	0.3345	0.3438	0.5082	0.5081	0.5094	0.5000
23	0.3345	0.3440	0.5083	0.5083	0.5096	0.5000
24	0.3342	0.3548	0.5174	0.5177	0.5188	0.5000
25	0.3001	0.2552	0.5082	0.5059	0.5078	0.5700
26	0.2994	0.2546	0.5092	0.5070	0.5088	0.5708
27	0.2984	0.2539	0.5105	0.5083	0.5101	0.5718
28	0.2891	0.1016	0.6674	0.6668	0.6692	0.6667
29	0.2434	0.1016	0.6772	0.6762	0.6784	0.6941
30	0.2431	0.1016	0.6776	0.6766	0.6788	0.6939
31	0.2422	0.1016	0.6787	0.6777	0.6800	0.6942
32	0.2126	0.1880	0.6534	0.6527	0.6540	0.6667
33	0.1922	0.1573	0.6776	0.6765	0.6778	0.7024
34	0.1918	0.1571	0.6782	0.6771	0.6784	0.7031
35	0.1914	0.1568	0.6788	0.6777	0.6790	0.7037
36	0.0998	0.0858	0.8421	0.8422	0.8428	0.8333
37	0.0918	0.0686	0.8470	0.8466	0.8472	0.8557
38	0.0914	0.0683	0.8476	0.8472	0.8478	0.8563
39	0.0913	0.0682	0.8478	0.8474	0.8480	0.8566
40	0.0000	0.0000	1.0000	1.0000	1.0000	1.0000
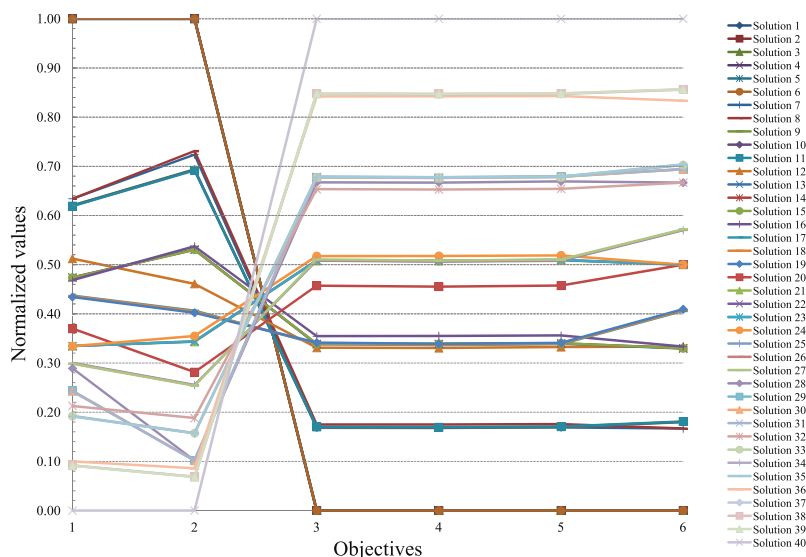
Fig. 8. Parallel coordinates plot for the second example (bio-ethanol supply chain in Argentina).

Table 10
Results of the rigorous MILP method for the second example (bio-ethanol supply chain in Argentina).

Number of clusters	Cluster	Objectives inside the clustering option															
		Option 1	Option 2	Option 3	Option 4	Option 5	Option 6	Option 7	Option 8	Option 9	Option 10	Option 11	Option 12	Option 13	Option 14	Option 15	Option 16
2	1	1.2	1														
	2	3,4,5,6	2,3,4,5,6														
	Clustering error	0.098	1.000														
3	1	3,4,5,6	2,3,4,6	3,4,5	2,3,5,6	4,5,6	2,3,5	2,3,4,5	2,3,6	2,4,5	3,5,6	2,4,5,6	3,4,6	2,4,6	2,5,6	2,3,4	1,2
	2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	3,4
	3	2	5	2.6	4	2.3	4.6	6	4.5	3.6	2.4	3	2.5	3.5	3.4	5.6	5.6
	Clustering error	0.080	0.086	0.086	0.086	0.086	0.086	0.086	0.086	0.086	0.086	0.086	0.086	0.086	0.086	0.086	0.098
4	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	2	2	2	2	2	2	2	2	5	3	6	3	5	4	3	5	6
	3	6	4.5	3	4.6	3.4	4	5	6	5	2.4	2.5	2.6	6	2.6	2.4	2.3
	4	3,4,5	3.6	4,5,6	3.5	5.6	3,5,6	3,4,6	2,3,4	2,4,6	3.5	4.6	3.4	2,3,5	4.5	3.6	4.5
	Clustering error	0.000	0.076	0.076	0.078	0.078	0.078	0.078	0.086	0.086	0.086	0.086	0.086	0.086	0.086	0.086	0.086

Fig. 11. Table 14 shows the square matrix used to obtain the dendrogram, which contains the distances between objectives quantified according to the Spearman metric. A comparison between the MILP and the statistical method is given in Table 15. Note that for this example the results obtained with the statistical method are rather similar to those produced by the rigorous MILP for the case of two clusters, but not for 3 and 4 clusters, in which the statistical method fails to identify a cluster with zero error. Hence, while the MILP guarantees the minimum possible error, the statistical method leads to solutions that may or may not be optimal
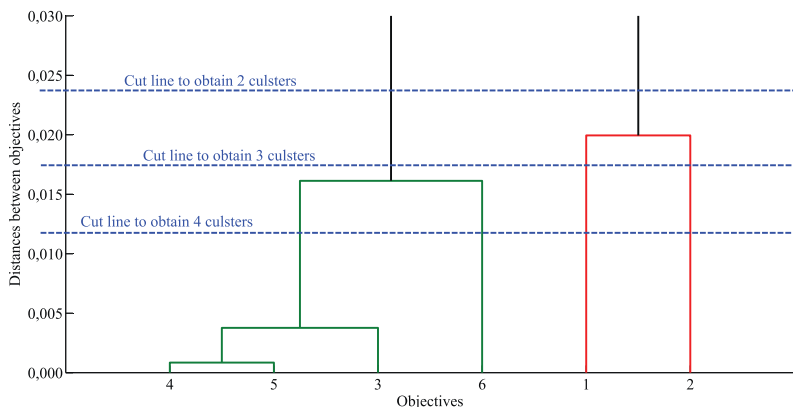

Fig. 9. Dendrogram scheme for the second example (bio-ethanol supply chain in Argentina).

Table 11
Distances between objectives (one minus the Spearman correlation) for the second example (bio-ethanol supply chain in Argentina).

Objectives	Objectives					
	1	2	3	4	5	6
1	0.0000	0.0200	1.9912	1.9832	1.9826	1.9952
2	0.0200	0.0000	1.9765	1.9700	1.9738	1.9887
3	1.9912	1.9765	0.0000	0.0039	0.0037	0.0104
4	1.9832	1.9700	0.0039	0.0000	0.0008	0.0191
5	1.9826	1.9738	0.0037	0.0008	0.0000	0.0189
6	1.9952	1.9887	0.0104	0.0191	0.0189	0.0000

depending on the case being solved. Let us note that the PCA approach used in the original work (Sabio et al., 2012) produces the same clustering as the MILP method for the case of three clusters.

The examples given before dealt with environmental problems. In this last example, we study an artificial dataset with 10 objectives and 41 points generated randomly. The goal is to test the performance of the statistical method when the objectives are poorly correlated. We assume that the elements inside the matrix are Pareto solutions of a given problem. These points have been in fact generated randomly (i.e., they are random numbers between 0 and 1) (see Table 16). Fig. 12 depicts the parallel coordinates plot associated with these points.

In Table 17 we present the results obtained using the rigorous MILP method (only the first 4 options are shown). As seen, the clustering error is high, mainly because of the stochastic nature of the data. For 2, 3 and 4 clusters, option 1 yields the minimum error, while for 5–8, all of the options shown lead to the same approximation error. With regard to the statistical method, it leads to clusters with larger errors, as seen in Table 18.

Table 12
Normalized Pareto solutions for the third example (hydrogen supply chain for vehicle use in Spain).

Pareto solution	Objective							
	1. Carcinogenic substances	2. Damage to human health caused by respiratory effects	3. Damage to human health caused by climate change	4. Damage to human health caused by ozone layer depletion	5. Damage to ecosystem quality caused by ecotoxic substances	6. Damage to ecosystem quality caused by acidification and eutrophication	7. Damage to minerals	8. Damage to fossil fuels
1	1.0000	0.3325	1.0000	0.9993	0.9998	1.0000	0.9759	1.0000
2	0.9996	0.4969	0.8548	0.9992	0.9997	0.9535	0.9829	0.8208
3	0.9994	0.6160	0.7450	1.0000	1.0000	0.9174	0.9881	0.6906
4	0.9992	0.6699	0.7095	0.9994	0.9998	0.9094	0.9902	0.6331
5	0.9990	0.7387	0.6344	0.9992	0.9998	0.8812	0.9932	0.5564
6	0.9990	0.7574	0.6175	0.9990	0.9997	0.8756	0.9939	0.5359
7	0.9988	0.8166	0.5643	0.9991	0.9997	0.8584	0.9965	0.4713
8	0.9986	0.8991	0.4900	0.9991	0.9997	0.8343	1.0000	0.3812
9	0.9402	1.0000	0.3684	0.9406	0.9414	0.7624	0.9480	0.2432
10	0.8709	0.3044	0.9568	0.8696	0.8703	0.9143	0.8498	0.9691
11	0.8709	0.3044	0.9568	0.8697	0.8704	0.9143	0.8498	0.9691
12	0.8703	0.3043	0.9566	0.8691	0.8698	0.9139	0.8493	0.9690
13	0.8707	0.3044	0.9568	0.8695	0.8702	0.9142	0.8497	0.9690
14	0.8706	0.3044	0.9568	0.8721	0.8701	0.9143	0.8496	0.9691
15	0.8700	0.9848	0.3450	0.8703	0.8711	0.7160	0.8795	0.2265
16	0.7606	0.9610	0.3085	0.7611	0.7616	0.6437	0.7727	0.2004
17	0.7412	0.2763	0.9136	0.7401	0.7405	0.8287	0.7233	0.9382
18	0.7414	0.2769	0.9131	0.7402	0.7408	0.8285	0.7236	0.9376
19	0.7413	0.2763	0.9136	0.7401	0.7406	0.8286	0.7234	0.9382
20	0.7406	0.2762	0.9134	0.7395	0.7400	0.8283	0.7227	0.9380
21	0.7412	0.2763	0.9136	0.7420	0.7406	0.8286	0.7233	0.9382
22	0.6569	0.9385	0.2738	0.6572	0.6578	0.5751	0.6716	0.1756
23	0.6119	0.2529	0.8661	0.6106	0.6110	0.7417	0.5973	0.9021
24	0.6117	0.2482	0.8704	0.6112	0.6110	0.7431	0.5969	0.9073
25	0.6109	0.2480	0.8701	0.6097	0.6101	0.7424	0.5961	0.9071
26	0.6118	0.2482	0.8704	0.6122	0.6110	0.7430	0.5970	0.9073
27	0.6119	0.2482	0.8704	0.6106	0.6110	0.7430	0.5971	0.9073
28	0.5810	0.9220	0.2485	0.5813	0.5817	0.5249	0.5975	0.1575
29	0.4824	0.2203	0.8269	0.4811	0.4814	0.6573	0.4707	0.8761
30	0.4824	0.2201	0.8271	0.4810	0.4814	0.6574	0.4707	0.8764
31	0.4823	0.2201	0.8273	0.4885	0.4813	0.6575	0.4707	0.8765
32	0.4823	0.2201	0.8271	0.4817	0.4813	0.6574	0.4706	0.8764
33	0.4812	0.2199	0.8268	0.4811	0.4803	0.6567	0.4696	0.8761
34	0.4014	0.8830	0.1886	0.4017	0.4019	0.4061	0.4222	0.1146
35	0.3529	0.1920	0.7839	0.3516	0.3518	0.5717	0.3444	0.8455
36	0.3529	0.1920	0.7839	0.3517	0.3517	0.5717	0.3443	0.8455
37	0.3529	0.1920	0.7840	0.3519	0.3517	0.5718	0.3443	0.8455
38	0.3529	0.1985	0.7781	0.3515	0.3518	0.5698	0.3447	0.8384
39	0.3515	0.1917	0.7835	0.3502	0.3504	0.5708	0.3430	0.8451
40	0.2235	0.1638	0.7407	0.2219	0.2221	0.4861	0.2180	0.8145
41	0.2218	0.8440	0.1286	0.2219	0.2221	0.2873	0.2469	0.0718

Table 13
Results of the rigorous MILP method for the third example (hydrogen supply chain for vehicle use in Spain).

Number of clusters	Cluster	Objectives inside the clustering option															
		Option 1	Option 2	Option 3	Option 4	Option 5	Option 6	Option 7	Option 8	Option 9	Option 10	Option 11	Option 12	Option 13	Option 14	Option 15	Option 16
2	1	1	1,3	1,3,6	1,3,6,8	1,3,8	1,6	1,6,8	1,8	2,4,5,7	3	3,6	3,6,8	3,8	6	6,8	8
	2	2,3,4,5,6,7,8	2,4,5,6,7,8	2,4,5,7,8	2,4,5,7	2,4,5,6,7	2,3,4,5,7,8	2,3,4,5,7	2,3,4,5,6,7	1,3,6,8	1,2,4,5,6,7,8	1,2,4,5,7,8	2,4,5,7	1,2,4,5,6,7	1,2,3,4,5,7,8	1,2,3,4,5,7	1,2,3,4,5,6,7
	Clustering error	0.680	0.680	0.680	0.680	0.680	0.680	0.680	0.680	0.680	0.680	0.680	0.680	0.680	0.680	0.680	0.680
3	1	2	2	2	3	3,8	2,7	2,7	2,7	7	1	2	2	6	7	7	7
	2	1,3,4,5,6,7	1,4,5,6,7	1,4,5,6,7,8	1,2,4,5,6,8	1,2,4,5,6	1,3,4,5,6	1,4,5,6	1,4,5,6,8	1,2,3,4,5,6	3,4,5,6,7,8	4,5,6,7,8	3,4,5,7	1,2,3,4,5,8	2,4,5,6,8	2,3,4,5	1,2,4,5
	3	8	3,8	3	7	7	8	3,8	3	8	2	1,3	1,6,8	7	1,3	1,6,8	3,6,8
	Clustering error	0.000	0.124	0.124	0.228	0.228	0.228	0.228	0.228	0.228	0.343	0.343	0.343	0.343	0.343	0.343	0.343
4	1	2	1,4,5,7	1,3,4,6	1,3,5,6,7	8	8	8	8	3,4	8	2	2	1,4,5,6,7	1	2	3,4,5,6
	2	8	2	5,7	8	4,6	1,3,7	4,5,7	2	8	2	4,5,6,7	8	8	2	8	1,7
	3	1,4,7	3,6	8	2	1,3,5,7	4,5,6	2	1,6	2	7	1,3	5,6,7	3	8	4,7	8
	4	3,5,6	8	2	4	2	2	1,3,6	3,4,5,7	1,5,6,7	1,3,4,5,6	8	1,3,4	2	3,4,5,6,7	1,3,5,6	2
	Clustering error	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000

Table 14
Distances between objectives (one minus the Spearman correlation) for the third example (bio-ethanol supply chain in Argentina).

| Objectives	Objectives							
	1	2	3	4	5	6	7	8
1	0.0000	0.3475	0.7975	0.0109	0.0070	0.1039	0.0206	0.8601
2	0.3475	0.0000	1.3795	0.3337	0.3261	0.6620	0.2904	1.4225
3	0.7975	1.3795	0.0000	0.8334	0.8476	0.4376	0.9065	0.0104
4	0.0109	0.3337	0.8334	0.0000	0.0072	0.1184	0.0222	0.8915
5	0.0070	0.3261	0.8476	0.0072	0.0000	0.1277	0.0139	0.9026
6	0.1039	0.6620	0.4376	0.1184	0.1277	0.0000	0.1634	0.4953
7	0.0206	0.2904	0.9065	0.0222	0.0139	0.1634	0.0000	0.9598
8	0.8601	1.4225	0.0104	0.8915	0.9026	0.4953	0.9598	0.0000
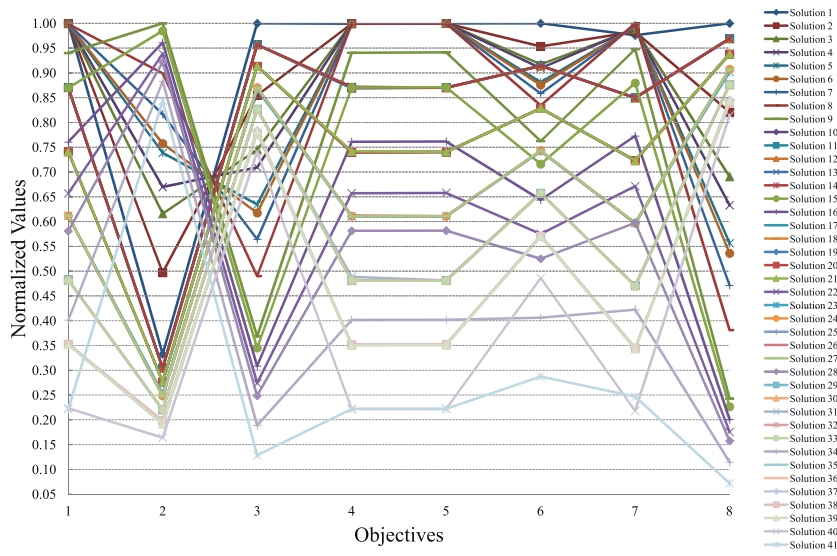
Fig. 10. Parallel coordinates plot for the third example (hydrogen supply chain for vehicle use in Spain).
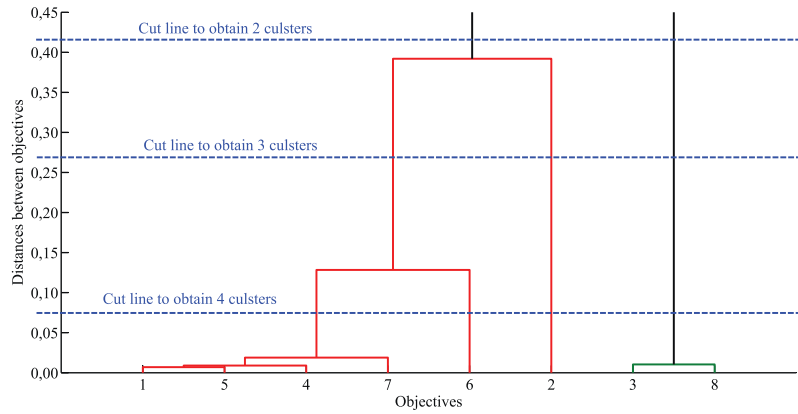

Fig. 11. Dendrogram scheme for the third example (hydrogen supply chain for vehicle use in Spain).
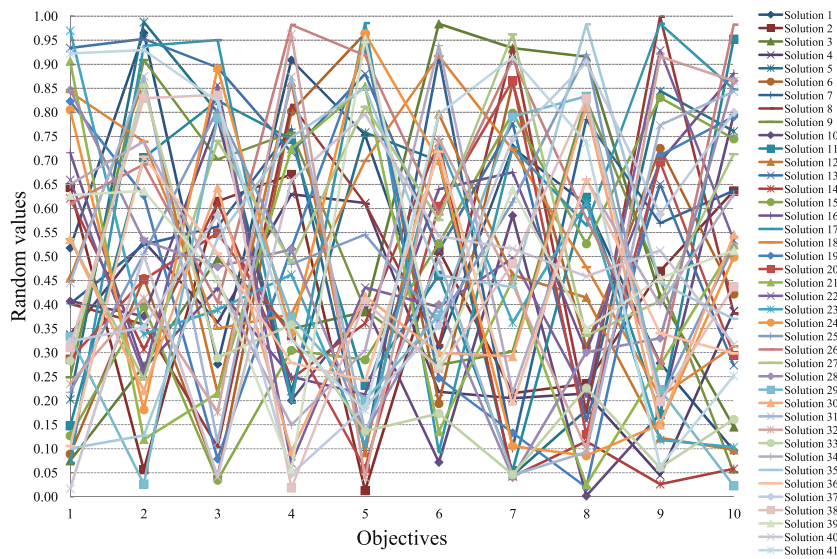

Fig. 12. Parallel coordinates plot for the fourth example (random source of data).

Table 15
Comparison between the rigorous MILP and the statistical method for the third example (hydrogen supply chain for vehicle use in Spain).

Number of clusters	Cluster	Objectives inside the cluster	
		MILP methodology	Dendrogram approach
2	1	3,8	3,8
	2	1,2,4,5,6,7	1,2,4,5,6,7
	Clustering error	0.680	0.680
3	1	2	2
	2	1,3,4,5,6,7	1,4,5,6,7
	3	8	8,3
	Clustering error	0.000	0.124
4	1	8	6
	2	2	2
	3	3,6	3,8
	4	1,4,5,7	1,4,5,7
	Clustering error	0.000	0.124

Table 16
Pareto solutions for the fourth example (random data).

Pareto solution	Objective									
	1	2	3	4	5	6	7	8	9	10
1	0.5177	0.9654	0.2763	0.9089	0.7530	0.3123	0.7264	0.6084	0.2783	0.0963
2	0.6434	0.0564	0.6150	0.6710	0.0129	0.5128	0.2154	0.2352	0.4696	0.6365
3	0.0750	0.2722	0.4957	0.3495	0.3850	0.9841	0.9336	0.9157	0.4170	0.1455
4	0.3376	0.5311	0.3630	0.6295	0.6113	0.2184	0.2050	0.2149	0.0445	0.3863
5	0.2028	0.9880	0.7999	0.2199	0.7582	0.6988	0.0446	0.1803	0.8447	0.7607
6	0.0889	0.4546	0.0420	0.8016	0.9640	0.1941	0.8618	0.3114	0.7251	0.4220
7	0.4016	0.5247	0.5620	0.7652	0.0966	0.9164	0.1011	0.7847	0.5694	0.6370
8	0.4017	0.3585	0.1036	0.8155	0.6100	0.3156	0.9350	0.1753	0.9992	0.3817
9	0.2489	0.9090	0.7021	0.7578	0.3892	0.2744	0.3025	0.8265	0.4513	0.0509
10	0.4071	0.3758	0.7867	0.2004	0.3973	0.0719	0.5853	0.0019	0.1721	0.7953
11	0.1476	0.7056	0.8244	0.7440	0.2323	0.4612	0.0549	0.6234	0.1714	0.9525
12	0.4552	0.8628	0.2165	0.8629	0.0916	0.7966	0.4612	0.4147	0.1226	0.0978
13	0.9340	0.9528	0.8932	0.7168	0.8800	0.5123	0.7771	0.1583	0.6508	0.2737
14	0.6209	0.3067	0.6117	0.2455	0.3609	0.7376	0.0424	0.1154	0.0263	0.0581
15	0.1269	0.3959	0.0347	0.3047	0.2851	0.5269	0.7984	0.5265	0.8306	0.7448
16	0.7158	0.2669	0.4343	0.2499	0.2127	0.6400	0.6749	0.3015	0.6968	0.8809
17	0.0706	0.9381	0.9503	0.1962	0.9858	0.0930	0.7267	0.5648	0.9845	0.8479
18	0.8425	0.7392	0.3501	0.3801	0.6990	0.9171	0.7220	0.4780	0.2085	0.3133
19	0.8231	0.6274	0.0786	0.5162	0.9337	0.2470	0.1332	0.0198	0.7103	0.7913
20	0.2868	0.4520	0.5488	0.3352	0.0588	0.6047	0.8662	0.1299	0.6969	0.2944
21	0.9063	0.1194	0.2156	0.7239	0.8540	0.1357	0.4514	0.0250	0.2737	0.5244
22	0.6594	0.2546	0.8567	0.0700	0.4355	0.3949	0.4965	0.2091	0.9288	0.5338
23	0.9698	0.3383	0.3892	0.4624	0.1078	0.7376	0.3623	0.6116	0.1182	0.1033
24	0.8053	0.1811	0.8920	0.3407	0.9640	0.6988	0.1055	0.0855	0.1504	0.4988
25	0.2179	0.5049	0.8055	0.4840	0.5455	0.3623	0.6134	0.9199	0.2009	0.5218
26	0.6098	0.6918	0.4044	0.9822	0.9184	0.5877	0.9103	0.5789	0.2773	0.9826
27	0.5361	0.2441	0.7406	0.4882	0.8118	0.5763	0.9624	0.3351	0.4076	0.7131
28	0.8469	0.5329	0.4787	0.5148	0.1348	0.4000	0.0405	0.3006	0.3301	0.8660
29	0.3340	0.0267	0.7893	0.3740	0.2005	0.3778	0.7914	0.8336	0.2229	0.0228
30	0.5355	0.2208	0.6430	0.0941	0.4173	0.2994	0.2918	0.8116	0.1496	0.5455
31	0.4453	0.8751	0.1142	0.8739	0.1349	0.4933	0.0452	0.0929	0.7734	0.8432
32	0.2352	0.4134	0.1758	0.9619	0.0400	0.7450	0.4431	0.0994	0.9161	0.8652
33	0.2830	0.8530	0.2892	0.3582	0.1367	0.1734	0.0463	0.2263	0.0605	0.1609
34	0.6501	0.7394	0.4484	0.1493	0.3050	0.9393	0.2156	0.6598	0.3997	0.6288
35	0.1005	0.1273	0.5380	0.7484	0.8514	0.4612	0.4360	0.9832	0.4450	0.3704
36	0.1056	0.7077	0.5064	0.2787	0.2421	0.7182	0.1922	0.6606	0.3401	0.2984
37	0.3235	0.3526	0.5865	0.0477	0.1864	0.3562	0.7448	0.9037	0.5878	0.7998
38	0.3106	0.8295	0.8361	0.0184	0.4062	0.2672	0.4854	0.8283	0.1978	0.4368
39	0.6274	0.6365	0.4537	0.0571	0.9529	0.2744	0.6390	0.3491	0.4552	0.5130
40	0.0168	0.5028	0.0410	0.6587	0.7962	0.5426	0.5161	0.4587	0.5123	0.3167
41	0.9225	0.9296	0.8231	0.3274	0.1799	0.7959	0.9142	0.7369	0.0626	0.2517

Table 17
Result of the rigorous MILP method for the fourth example (random data).

Number of clusters	Cluster	Objectives inside the clustering option			
		Option 1	Option 2	Option 3	Option 4
2	1	5	10	5,10	9
	2	1,2,3,4,6,7,8,9,10	1,2,3,4,5,6,7,8,9	1,2,3,4,6,7,8,9	1,2,3,4,5,6,7,8,10
	Clustering error	0.940	0.950	0.950	0.950
3	1	5	6	6	6
	2	6	8	5,8	9
	3	1,2,3,4,7,8,9,10	1,2,3,4,5,7,9,10	1,2,3,4,7,9,10	1,2,3,4,5,7,8,10
	Clustering error	0.900	0.920	0.920	0.920
4	1	4	1	4,9	5
	2	5	8	5	6
	3	6	10	6	9
	4	1,2,3,7,8,9,10	2,3,4,5,6,7,9	1,2,3,7,8,10	1,2,3,4,7,8,10
	Clustering error	0.850	0.890	0.890	0.890
5	1	5	5	5	1,2,3,7,8,9
	2	6	6	1,2,3,7,8,9	5
	3	4	1,2,3,7,8,9	6	6
	4	1,2,3,7,8,9	4	4	4
	5	10	10	10	10
	Clustering error	0.780	0.780	0.780	0.780
6	1	10	3,4,5,6,9	10	10
	2	8	10	3,4,5,6,9	8
	3	3,4,5,6,9	8	8	1
	4	1	1	1	3,4,5,6,9
	5	7	7	7	7
	6	2	2	2	2
	Clustering error	0.570	0.570	0.570	0.570
7	1	7	8	5,6,9	2
	2	3,8	4	10	10
	3	2	3,5,6,9	4	7
	4	1	2	7	1
	5	5,6,9	10	3,8	8
	6	10	7	2	4
	7	4	1	1	3,5,6,9
	Clustering error	0.480	0.480	0.480	0.480
8	1	10	10	10	10
	2	4	4	4	4
	3	2	2	2	2
	4	1	1	1	1
	5	7	7	7	3,5,9
	6	6	6	3,5,9	7
	7	8	3,5,9	6	6
	8	3,5,9	8	8	8
	Clustering error	0.260	0.260	0.260	0.260

Table 18
Comparison between the rigorous MILP and the statistical method for the fourth example (random data).

Number of clusters	Cluster	Objectives inside the cluster	
		MILP methodology	Dendrogram approach
2	1	5	2,4,5,7,9,10
	2	1,2,3,4,6,7,8,9,10	1,3,6,8
	Clustering error	0.940	0.980
3	1	5	2,4,5,7,9,10
	2	6	1.6
	3	1,2,3,4,7,8,9,10	3.8
	Clustering error	0.900	0.980
4	1	4	5,7,9,10
	2	5	2.4
	3	6	1.6
	4	1,2,3,7,8,9,10	3.8
	Clustering error	0.850	0.980
5	1	5	9,10
	2	6	5,7
	3	4	2,4
	4	1,2,3,7,8,9	1,6
	5	10	3,8
	Clustering error	0.780	0.980

Table 18 (*Continued*)

Number of clusters	Cluster	Objectives inside the cluster	
		MILP methodology	Dendrogram approach
6	1	10	9,10
	2	8	5,7
	3	3,4,5,6,9	2
	4	1	4
	5	7	1,6
	6	2	3,8
	Clustering error	0.570	0.920
7	1	7	9,10
	2	3,8	5,7
	3	2	2
	4	1	4
	5	5,6,9	1
	6	10	6
	7	4	3,8
	Clustering error	0.480	0.770
8	1	10	9,10
	2	4	5,7
	3	2	2
	4	1	4
	5	7	1
	6	6	6
	7	8	3
	8	3,5,9	8
	Clustering error	0.260	0.770

6. Conclusions

In this work, we have presented a rigorous MILP-based method for classifying objectives into clusters in MOO problems that is based on the concept of approximation error. Our method classifies similar objectives into clusters using a set of Pareto points generated via any standard algorithm for MOO. Our approach simplifies to a large extent multi-criteria decision-making from the viewpoints of calculation and analysis of the Pareto solutions, as it allows practitioners to conduct the analysis on a reduced number of objectives, each one extracted from a different cluster. We tested the capabilities of our method using several examples, where a comparison with a well-known statistical clustering method was carried out. Numerical results show that when the objectives are highly correlated, both methods (i.e., the MILP and the statistical method) lead to similar results, being the statistical approach the one with lower CPU times. In contrast, the statistical method lead to clusters with larger errors that differ from those produced by the rigorous MILP approach when the correlation between objectives is low. Our rigorous MILP-based approach is the only one that guarantees the identification of the minimum approximation error solution. In general, there will always be a trade-off between computational effort and quality of the results, so the application of one method or another will depend on the goal of the analysis. Note, however, that the MILP-based approach is the only one that provides, along with the clustering solution itself, an approximation error that quantifies the change in the dominance structure after removing objectives. Future work will concentrate on developing hybrid methods that will take advantage of the complementary strengths of both techniques.

References

Brockhoff, D. (2009). *Many-objective optimization and hypervolume-based search*. ETH Zurich

Brockhoff, D. (2010). *Java implementations of exact and greedy objective reduction algorithms [WWW Document]. Java implementations of exact and greedy objective reduction algorithms*. http://www.tik.ee.ethz.ch/sop/download/supplementary/objectiveReduction/. Accessed 03.20.13.

Brockhoff, D., Saxena, D. K., Deb, K., & Zitzler, E. (2007). On handling a large number of objectives a posteriori and during optimization. In J. Knowles, D. Corne, & K. Deb (Eds.), *Multiobjective problem solving from nature: From concepts to applications* (pp. 377–403). Springer.

Brockhoff, D., & Zitzler, E. (2006a). Are all objectives necessary? On dimensionality reduction in evolutionary multiobjective optimization. In T. P. Runarsson, H.-G. Beyer, E. Burke, J. J. Merelo-Guervós, L. D. Whitley, & X. Yao (Eds.), *Parallel problem solving from nature – PPSN IX* (pp. 533–542). Berlin/Heidelberg: Springer Berlin Heidelberg.

Brockhoff, D., & Zitzler, E. (2006b). *Dimensionality reduction in multiobjective optimization with (partial) dominance structure preservation: Generalized minimum objective subset problems (TIK Report No. 247)*. Computer Engineering and Networks Laboratory (TIK), ETH Zurich.

Brockhoff, D., & Zitzler, E. (2006c). *On objective conflicts and objective reduction in multiple criteria optimization (techreport)*. ETH Zurich: Computer Engineering and Networks Laboratory (TIK).

Brockhoff, D., & Zitzler, E. (2007a). Improving hypervolume-based multiobjective evolutionary algorithms by using objective reduction methods. In *Congress on evolutionary computation (CEC 2007)*. Singapore: IEEE Press.

Brockhoff, D., & Zitzler, E. (2007b). Dimensionality reduction in multiobjective optimization: The minimum objective subset problem. In K. H. Waldmann, & U. M. Stocker (Eds.), *Operations Research Proceedings 2006* (pp. 423–429). Saarbücken, Germany: Springer.

Brockhoff, D., & Zitzler, E. (2007c). *Offline and online objective reduction in evolutionary multiobjective optimization based on objective conflicts (TIK Report No. 269)*. ETH Zurich: Computer Engineering and Networks Laboratory (TIK).

Brockhoff, D., & Zitzler, E. (2009a). Objective reduction in evolutionary multiobjective optimization: Theory and applications. *Evolutionary Computation*, *17*, 135–166.

Brockhoff, D., & Zitzler, E. (2009b). Automated aggregation and omission of objectives to handle many-objective problems. In *Conference on multiple objective and goal programming (MOPGP 2008), lecture notes in economics and mathematical systems*. Springer.

Brooke, A. (1997). *GAMS Language Guide Release 2.25*. GAMS Development Corporation.

Brunet, R., Guillén-Gosálbez, G., & Jiménez, L. (2012). Cleaner design of single-product biotechnological facilities through the integration of process simulation, multiobjective optimization, life cycle assessment, and principal component analysis. *Industrial and Engineering Chemistry Research*, *51*, 410–424.

Copado-Méndez, P. J., Guillén-Gosálbez, G., & Jiménez, L. (2012). Rigorous computational methods for dimensionality reduction in multi-objective optimization. In *22nd European symposium on computer aided process engineering* (pp. 1292–1296). Elsevier.

Deb, K., & Saxena, D. K. (2005). On finding pareto-optimal solutions through dimensionality reduction for certain large-dimensional multi-objective optimization problems EMO for many objectives. *Kangal Report*, 3353–3360, 2005011.

Guillén-Gosálbez, G. (2011). A novel MILP-based objective reduction method for multi-objective optimization: Application to environmental problems. *Computers and Chemical Engineering*, *35*, 1469–1477.

Gutiérrez, E., Lozano, S., Moreira, M. T., & Feijoo, G. (2010). Assessing relationships among life-cycle environmental impacts with dimension reduction techniques. *Journal of Environmental Management*, *91*, 1002–1011.

Jacquemin, L., Pontalier, P.-Y., & Sablayrolles, C. (2012). *Life cycle assessment (LCA) applied to the process industry: A review*.

Jain, A. K., Murty, M. N., & Flynn, P. J. (1999). Data clustering: A review. *ACM Computing Surveys*, *31*, 316–323.

Kostin, A., Guillén-Gosálbez, G., Mele, F. D., & Jiménez, L. (2012). Identifying key life cycle assessment metrics in the multiobjective design of bioethanol supply chains using a rigorous mixed-integer linear programming approach. *Industrial and Engineering Chemistry Research*, *51*, 5282–5291.

Lehmann, E. L. (1998). *Nonparametrics: Statistical methods based on ranks, revised, Facsimile. ed.* Pearson Education.

MathWorks, n.d. Statistics Toolbox – Documentation [WWW Document]. MathWorks. http://www.mathworks.es/help/toolbox/stats/ Accessed 06.05.12.

Mucha, H.-J., Bartel, H.-G., & Morales-Merino, C. (2013). Visualisation of cluster analysis results. In A. Giusti, G. Ritter, & M. Vichi (Eds.), *Classification and data mining, studies in classification, data analysis, and knowledge organization* (pp. 261–270). Berlin/Heidelberg: Springer.

Pieragostini, C., Mussati, M. C., & Aguirre, P. (2012). On process optimization considering LCA methodology. *Journal of Environmental Management*, *96*, 43–54.

Pozo, C., Ruíz-Femenia, R., Caballero, J., Guillén-Gosálbez, G., & Jiménez, L. (2012). On the use of principal component analysis for reducing the number of environmental objectives in multi-objective optimization: Application to the design of chemical supply chains. *Chemical Engineering Science*, *69*, 146–158.

Sabio, N., Kostin, A., Guillén-Gosálbez, G., & Jiménez, L. (2012). Holistic minimization of the life cycle environmental impact of hydrogen infrastructures using multi-objective optimization and principal component analysis. *International Journal of Hydrogen Energy*, *37*, 5385–5405.

Salcedo, R., Antipova, E., Boer, D., Jiménez, L., & Guillén-Gosálbez, G. (2012). Multi-objective optimization of solar Rankine cycles coupled with reverse osmosis desalination considering economic and life cycle environmental concerns. *Desalination*, *286*, 358–371.

Seber, G. A. F. (2004). *Multivariate observations*. Hoboken, New Jersey: John Wiley & Sons. Inc.

Singh, H. K., Isaacs, A., & Ray, T. (2011). A Pareto corner search evolutionary algorithm and dimensionality reduction in many-objective optimization problems. *IEEE Transactions on Evolutionary Computation*, *15*, 539–556.

Späth, H. (1985). *Cluster dissection and analysis: Theory*. Horwood: FORTRAN programs, examples.

Taboada, H. A., & Coit, D. W. (2008). Multi-objective scheduling problems: Determination of pruned Pareto sets. *IIE Transactions (Institute of Industrial Engineers)*, *40*, 552–564.

Zhou, C., Zheng, J., Li, K., & Lv, H. (2009). Objective reduction based on the least square method for large-dimensional multi-objective optimization problem. In *5th International Conference on Natural Computation, ICNC* (pp. 350–354).