European Journal of Operational Research 210 (2011) 624–634

Contents lists available at ScienceDirect

European Journal of Operational Research

journal homepage: www.elsevier.com/locate/ejor

Decision Support

A clustering procedure for reducing the number of representative solutions in the Pareto Front of multiobjective optimization problems ☆

E. Zio [a,b,*], R. Bazzo [b]

[a] Ecole Centrale Paris and Supelec, Paris, France
[b] Politecnico di Milano, Milano, Italy

ARTICLE INFO

Article history:
Received 17 November 2009
Accepted 14 October 2010
Available online 24 November 2010

Keywords:
Multiobjective optimization
Subtractive clustering
Level Diagrams
Fuzzy preference assignment
Genetic algorithms
Redundancy allocation

ABSTRACT

In many multiobjective optimization problems, the Pareto Fronts and Sets contain a large number of solutions and this makes it difficult for the decision maker to identify the preferred ones. A possible way to alleviate this difficulty is to present to the decision maker a subset of a small number of solutions representatives of the Pareto Front characteristics.

In this paper, a two-steps procedure is presented, aimed at identifying a limited number of representative solutions to be presented to the decision maker. Pareto Front solutions are first clustered into “families”, which are then synthetically represented by a “head-of-the-family” solution. Level Diagrams are then used to represent, analyse and interpret the Pareto Front reduced to its head-of-the-family solutions. The procedure is applied to a reliability allocation case study of literature, in decision-making contexts both without or with explicit preferences by the decision maker on the objectives to be optimized.

1. Introduction

Multiobjective decision-making can be aided by methods which provide flexible ways of handling multiple objectives and decision maker (DM) preferences, rather than by quasi-prescriptive methods based on the aggregation of the multiple objectives into a single one. In this respect, the importance of determining and including DM preferences for the practice of multiobjective decision making and optimization is inarguable.

The particular context of reference for the present work are practical decision-making situations concerning high-consequence technologies, e.g. nuclear, oil and gas, transport etc. The starting point is the acknowledgment that technical analyses provide useful decision support in the sense that their outcomes inform the decision makers insofar as the technical side of the problem is relevant for the decision.

It is further understood that the actual decision outcome for a critical situation involving a potential for large consequences typically derives from a thorough process which combines (i) an analytic evaluation of the situation (i.e., the technical assessment) by rigorous, replicable methods evaluated under agreed protocols of an expert community and peer-reviewed to verify the assumptions underpinning the analysis, and (ii) a deliberative group exercise in which all involved stakeholders and decision makers collectively consider the decision issues, look into the arguments for their support, scrutinize the outcomes of the technical analysis and introduce all other values (e.g. social and political) not explicitly included in the technical analysis. This way of proceeding allows keeping the technical analysis manageable by complementation with deliberation for ensuring coverage of the non-modelled issues. In this way, the analytic evaluation (i.e., the technical assessment) supports the deliberation by providing numerical outputs of the relevant parameters, possibly to be compared with predefined numerical safety criteria for further guidance to the decision, and also all the argumentations behind the analysis itself, including the assumptions, hypotheses, parameters and their uncertainties.

The ultimate concern of the DM is to confidently fulfil his or her conflicting objectives, while satisfying the constraints posed by the problem itself.

In practice, the technical assessment amounts to the solution of a multiobjective optimization problem in terms of a discrete approximation of the Pareto Front and corresponding Pareto Set of solutions. The ultimate purpose of the technical assessment is to provide the DM with a clearly informed picture of the problem upon which he or she can confidently reason and deliberate. On the basis of the information provided by the technical assessment, the

☆ This work has been partially funded by the Foundation pour une Culture de Securité Industrielle of Toulouse, France, under the research contract AO2009-04.
* Corresponding author at: Ecole Centrale Paris and Supelec, Paris, France and Politecnico di Milano, Milano, Italy.

E-mail addresses: enrico.zio@ecp.fr, enrico.zio@supelec.fr, enrico.zio@polimi.it (E. Zio).

0377-2217/$ - see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.ejor.2010.10.021

DM is requested to select one or more feasible solutions according to criteria which depend on the decision situation. In the literature, it is well acknowledged that presenting the DM with too many alternatives increases the burden of his or her decision-making task.

Different approaches exist for introducing DM preferences in the optimization process; a common classification is based on when the DM is consulted: a priori, a posteriori, or interactively during the search. A priori methods use DM preferences to bias the search of optimal solutions towards a preferred region, for example by changing the definition of dominance (Molina et al., 2009; Zio et al., 2009), by weighting differently the objectives (Yang, 1996), by assigning reference values (goals) and priority levels to the objectives (Yang, 2000), by assuming a utility function describing the DM behaviour and interest in the alternative solutions (Malakooti, 1988). Interactive methods require the direct intervention of the DM in the optimal solution search, for example simply to stop an iterative trial-and-error search when satisfactory results are reached (Katagiri et al., 2008) or more effectively to drive the optimization by ranking and eliminating alternatives based on indicated preference strengths (Roy, 1968,1974,1986; Malakooti, 1988; De Boer et al., 1998) or by bounding DM utility functions by elicited preference information (Cho and Kim, 1997; Rios Insua and Martin, 1994), while accounting for the fact that the consequences of the alternative solutions may not be completely known, the problem definition may not be exact and the DM preferences may be only partially known and even partially inconsistent. A posteriori methods, on the other hand, apply DM preferences only after the optimal solutions of the Pareto Front are found.

The selection task by the DM can be difficult when the Pareto Front contains a large number of solutions. To make the task feasible, only a small number of solutions representatives of the Pareto Front should be offered for selection to the DM.

At the same time, it is important for the DM to be able to analyse the Pareto Front to evaluate the quality and adequateness of the results. For this reason, the number of solutions that the Pareto Front is reduced to has to be determined carefully: a number of solutions too small might not be sufficiently representative and informative, whereas a number too large might still be intractable for decision making purposes.

In this work, an a posteriori procedure is proposed for reducing the set of Pareto solutions on a Pareto Front to a small number of representative ones. Two situations are considered, depending on the presence or absence of explicit preferences of the DM on the objectives of the optimal decision. The procedure is made of two main steps. First, the set of optimal solutions constituting the Pareto Front and Set is partitioned in a number of clusters (here also called "families") of solutions sharing common features. The clustering is performed by considering the distance between solutions in the objective values space; to this purpose, different clustering algorithms are available, e.g., the *k*-means (Bandyopadhyay and Maulik, 2002), the *c*-means (Bezdek, 1974) and subtractive clustering (Chiu, 1994); in this paper, the latter algorithm is used for the technical reasons explained in Section 3. The second step consists in selecting for each family (or cluster) the representative solution (the "head of the family"): depending on the decision situations (i.e., the presence or the absence of preferences on the objectives), different selection criteria may apply to provide the DM with the best solutions according to his or her requirements.

The outcome of the procedure is a Pareto Front reduced to a number of representative solutions, thanks to the clustering technique, which are ranked either by a specific norm or by DM explicit preferences, if available. In both cases, one may actually proceed to identify the best solution, i.e. the highest ranked according to the specified norm or DM preferences. However, in various practical decision making situations it is important that the DM has a picture of the spectrum of solutions available on the Pareto Front, for a solid support to his or her decision or for considering alternative compromises and preferences on the objectives in light of the Pareto Front obtained. Within a deliberative process of decision making, supported by the quantitative analysis performed, knowledge of the information contained in the Pareto Front and Set allows critically discussing, questioning, revising the preferences assumed and defending, supporting the choices made, possibly even non-optimal for the given situation a posteriori of consideration of additional aspects of the situation involved in the decision.

The originality of the work mainly lies in the following aspects:

- the proposal of using the ideal solution (i.e., the solution which is optimal with respect to all the objectives simultaneously) for identifying the representatives of the clusters in decision situations in which the DM does not express preferences on the objectives of the optimization; such proposal stands on a definition of a 1-norm to measure the distance of the solutions from the ideal one and provides an algorithmic generalization of the empirical method introduced by the authors in (Zio and Bazzo, 2010a);
- the effective integration of a fuzzy scoring procedure introduced by the authors in (Zio and Bazzo, 2010b) for ranking the cluster representative solutions in situations in which the DM expresses preferences on the objectives of the optimization;
- the analysis of the clustered Pareto Front and Set within a Level Diagrams representation.

These developments render the proposed procedure of general applicability with respect to both the decision situations and the Pareto Front characteristics.

The procedure is applied to a reference case study regarding a redundancy allocation problem of literature with three objectives: system availability to be maximized, system cost and weight to be minimized (Taboada and Coit, 2007).

Level Diagrams (Blasco et al., 2008) are used to graphically represent, analyse and interpret the Pareto Front and Set considered in the analyses.

The remainder of the paper is organized as follows: Section 2 presents upfront the case study; Section 3 contains the analysis of the clustering algorithm, the methods considered for selecting the representative heads of the families and the results of their application to the case study; Section 4 provides a critical discussion of the results; Section 5 gives the conclusions that can be drawn from the findings of the work.

2. Case study: redundancy allocation in a multistate system

The case study is taken from (Taboada and Coit, 2007); it is a system reliability design problem with three conflicting objectives: system availability to be maximized; system cost and weight to be minimized. The system is made of $u = 5$ units (subsystems) connected in series; each unit can be provided with redundancy by selecting components from m_p types available in the market, $p = 1, \ldots, 5$. Each component is binary, i.e., at any time it can be in only two states: functioning at nominal capacity or failed, with zero capacity. The collective performance of these binary components leads to a multistate system behaviour. The types of components available are characterized by their availability, nominal capacity, cost and weight in arbitrary units (Table 1). Without loss of generality, component capacities are measured in terms of percentage of the maximum system demand. The different demand levels for a given period, i.e., the cumulative demand curve, are given in Table 2.

The three objective functions $J_i(\theta)$, $i = 1,\ldots,3$ driving the search to the optimal system design are:

Table 1
Characteristics of the components available on the market.

Subsystem p	Component Type h	Availability	Capacity (%)	Cost	Weight
1	1	0.980	120	0.590	35.4
	2	0.977	100	0.535	34.9
	3	0.982	85	0.470	34.1
	4	0.978	85	0.420	33.9
	5	0.983	48	0.400	34.2
	6	0.920	31	0.180	34.3
	7	0.984	26	0.220	32.6
2	1	0.995	100	0.205	26.5
	2	0.996	92	0.189	22.4
	3	0.997	53	0.091	20.3
	4	0.997	28	0.056	21.7
	5	0.998	21	0.042	25.2
3	1	0.971	100	7.525	42.1
	2	0.973	60	4.720	41.7
	3	0.971	40	3.590	40.8
	4	0.976	20	2.420	39.6
4	1	0.977	115	0.180	25.4
	2	0.978	100	0.160	23.9
	3	0.978	91	0.150	24.7
	4	0.983	72	0.121	24.6
	5	0.981	72	0.102	23.6
	6	0.971	72	0.096	26.2
	7	0.983	55	0.071	25.5
	8	0.983	25	0.049	22.6
	9	0.977	25	0.044	24.8
5	1	0.984	128	0.986	15.4
	2	0.983	100	0.825	15.3
	3	0.987	60	0.490	14.9
	4	0.981	51	0.475	15.0

Table 2
Cumulative demand curve.

Demand (%)	100	80	50	20
Duration (hours)	4203	788	1228	2536
Duration (%)	0.48	0.09	0.14	0.29

$$\text{Availability}: \quad \max_{\theta} J_1(\theta) = \max_{\theta_p, p=1,\ldots,5}\left[\prod_{p=1}^{u} A_p(\theta_p)\right] \tag{1}$$

$$\text{Cost}: \quad \min_{\theta} J_2(\theta) = \min_{\substack{\theta_{ph}, p=1,\ldots,5\\ h=1,\ldots,m_p}}\left[\sum_{p=1}^{u}\sum_{h=1}^{m_p} c_{ph}\theta_{ph}\right] \tag{2}$$

and

$$\text{Weight}: \quad \min_{\theta} J_3(\theta) = \min_{\substack{\theta_{ph}, p=1,\ldots,5\\ h=1,\ldots,m_p}}\left[\sum_{p=1}^{u}\sum_{h=1}^{m_p} w_{ph}\theta_{ph}\right] \tag{3}$$

subject to: $1 \leqslant \sum_{h=1}^{m_p}\theta_{ph} \leqslant n_{\max,p},\ \forall p = 1,\ldots,u,$

where A_p = availability of the pth subsystem, θ_{ph} = number of components of type h used in subsystem p, $\theta_p = [\theta_{p1}\ \theta_{p2}\ \ldots\ \theta_{pm_p}]$ is the vector of the indexes of the pth subsystem configuration, $\theta = [\theta_1\ \theta_2\ \ldots\ \theta_5]$ is the vector of the indexes of the system configuration, u = number of subsystems, m_p = number of components in the pth subsystem, $n_{\max,p}$ = user-defined maximum number of redundant components that can be placed in subsystem p, c_{ph}, w_{ph} = cost and weight of the hth type of component available on the market.

The optimization problem has been solved in (Taboada et al., 2007) using the MOMS-GA algorithm; the Universal Moment Generating Function (UMGF) approach (Ushakov, 1986; Levitin and Lisnianski, 2001; Levitin, 2005) was used to compute the multi-state system availability (1),

The resulting Pareto Set (Θ) is made of 118 solutions; the Pareto Front is showed in Fig. 1.

3. Clustering of the Pareto Front

3.1. Subtractive clustering algorithm

The first step of the procedure for reducing the number of solutions to present to the DM is to group the solutions of the Pareto Front in a number K of families of solutions sharing similar characteristics. In this work, subtractive clustering (Chiu, 1994) is used to identify the families of similar solutions (clusters) F^j, $j = 1,\ldots,K$, in the objective function space. The clustering is performed on the basis of the distance between solutions, i.e., in this case, between objective function values.

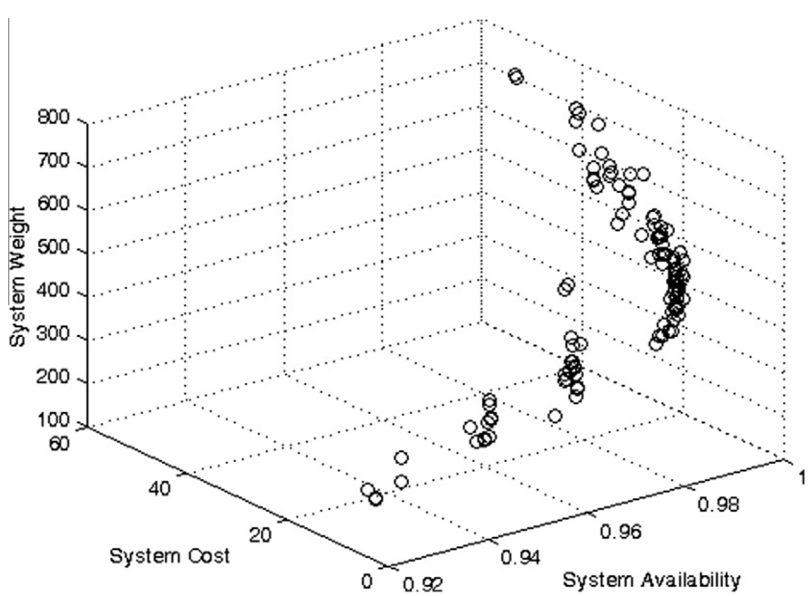

Fig. 1. Pareto Front, in the objective functions space, obtained by the MOMS-GA algorithm.

Subtractive clustering has been chosen over other methods like the k-means and fuzzy c-means, because it does not require the solution of a minimization problem of an objective function, nor any random initialization so that the results are independent on the initial cluster centers and membership function choice (Selim and Ismail, 1984; Spath, 1985; Tou and Gonzalez, 1974); also, differently from the k-means and the fuzzy c-means, the cluster centers $\underline{J}^j_{norm},\ j = 1,\dots,K$, found by subtractive clustering are vectors of objective functions values corresponding to existing solutions in the Pareto Front and Set and thus can be used directly as representative solutions on the reduced Pareto Front.

Let us consider a Pareto Set Θ made of n solutions; to the ith solution θ^i ($i = 1,\dots,n$) corresponds a vector of objective values

$$\underline{J}(\theta^i) = \left(J_1(\theta^i) J_2(\theta^i)\ \dots\ J_{Nobj}(\theta^i)\right) \tag{4}$$

where N_{obj} is the number of objective functions of the optimization problem. Since the objective functions are usually given in different units and scales, their values are normalized with respect to the minimum value in the Pareto Front:

$$\underline{J}_{norm}(\theta^i) = \left(J_{1,norm}(\theta^i)\ J_{2,norm}(\theta^i)\ \dots\ J_{Nobj,norm}(\theta^i)\right) \tag{5}$$

$$J_{s,norm}(\theta^i) = \frac{J_s(\theta^i) - J_{s,\min}}{J_{s,\max} - J_{s,\min}}, \quad s = 1,\dots,N_{obj} \tag{6}$$

where

$$J_{s,\min} = \min_i J_s(\theta^i) \quad \text{and } J_{s,\max} = \max_i J_s(\theta^i).$$

The normalization does not impact the successive clustering of the solutions performed by the subtractive clustering algorithm, which examines the optimal solutions along the Pareto Front to define the cluster centers according to the density of surrounding solutions.

The flowchart of the subtractive clustering algorithm is given in Fig. 2. Given the Pareto Front made of n normalized solutions ($\underline{J}_{norm}(\theta^i)$), the algorithm starts by calculating the following potential $P(\underline{J}_{norm}(\theta^i))$, $i = 1,2,\dots,n$:

$$P(\underline{J}_{norm}(\theta^i)) = \sum_{l=1}^{n} e^{-\alpha\left\|\underline{J}_{norm}(\theta^i) - \underline{J}_{norm}(\theta^l)\right\|^2}, \quad \alpha = \frac{4}{r_a^2} \tag{7}$$

where $r_a \in [0,1]$ is an input parameter called cluster radius, which indicates a cluster center's range of influence in each of the data

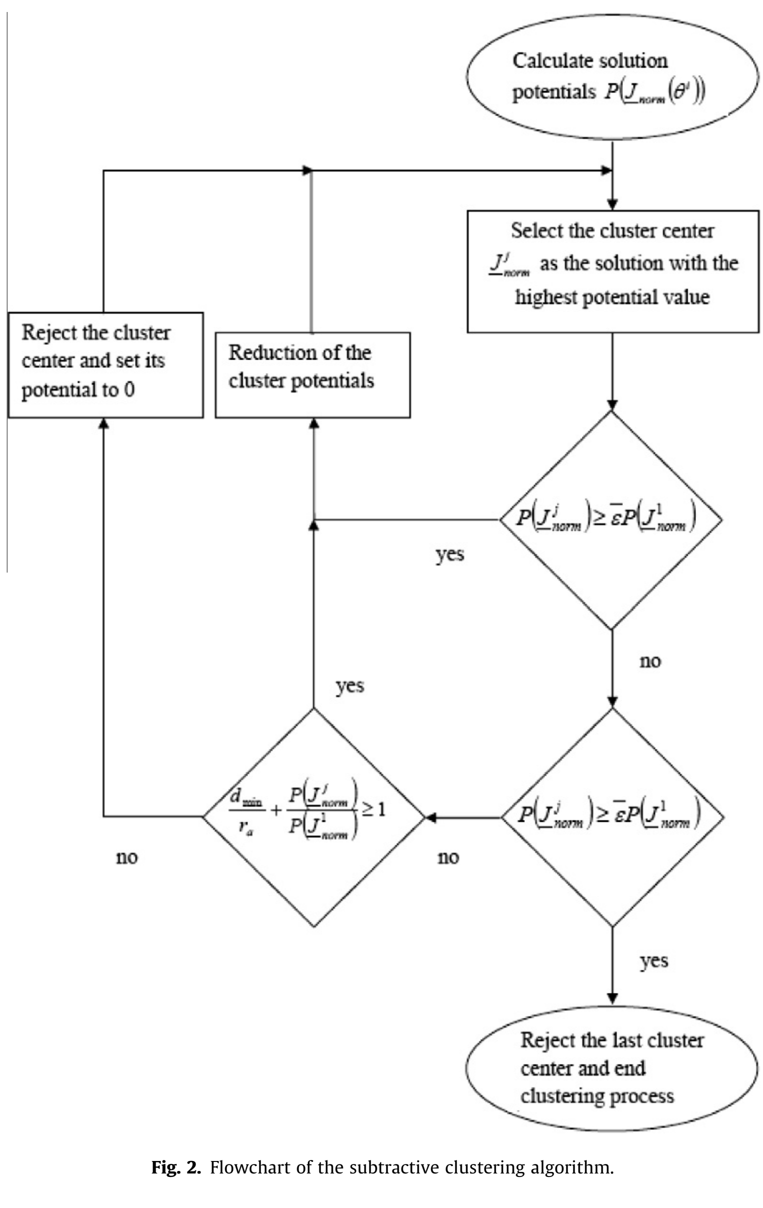

Fig. 2. Flowchart of the subtractive clustering algorithm.

dimensions, assuming the solution space is a unit hyperbox; a solution outside the cluster radius r_a of the cluster center has a small influence on the potential. The choice of the parameter r_a determines the number K of clusters that will be identified: a large value of r_a gives few and large clusters; usually this parameter is taken smaller than 0.5.

After the potentials of all solutions have been computed, the first cluster center $\underline{J}^1_{norm}$ is selected as the solution with the highest potential value $P(\underline{J}^1_{norm})$. All the other $n-1$ solutions potentials $P(\underline{J}_{norm}(\theta^i))$ are corrected subtracting the potential $P(\underline{J}^1_{norm})$ multiplied by a factor which considers the distance between the ith solution and the first cluster center:

$$P(\underline{J}_{norm}(\theta^i)) = P(\underline{J}_{norm}(\theta^i)) - P\left(\underline{J}^1_{norm}\right) e^{-\beta\left\|\underline{J}_{norm}(\theta^i)-\underline{J}^1_{norm}\right\|^2},$$
$$\beta = \frac{4}{r_b^2} \quad \text{and } r_b = qr_a \tag{8}$$

where q is an input parameter called squash factor, which indicates the neighborhood with a measurable reduction of potential expressed as a fraction of the cluster radius; a value of 2 of the squash factor indicates that the clusters found are far from each other.

From the $n-1$ remaining solutions, the one with the highest value of the reduced potential is chosen as the second cluster center $\underline{J}^2_{norm}$.

Generally, for the jth cluster center found $\underline{J}^j_{norm}$, with $j = 1,\ldots,K$, the potentials are reduced as follows:

$$P(\underline{J}_{norm}(\theta^i)) = P(\underline{J}_{norm}(\theta^i)) - P\left(\underline{J}^j_{norm}\right) e^{-\beta\left\|\underline{J}_{norm}(\theta^i)-\underline{J}^j_{norm}\right\|^2} \tag{9}$$

Two other input parameters are introduced: the accept ratio $\overline{\varepsilon}$ and the reject ratio $\underline{\varepsilon}$, which are respectively the fraction of the potential of the first cluster center above or below which another solution is accepted or rejected as a cluster center. Then, the process of finding new cluster centers and reducing the potential as in (9), is repeated according to the following criteria: if

$$P\left(\underline{J}^j_{norm}\right) \geqslant \overline{\varepsilon} P\left(\underline{J}^1_{norm}\right) \tag{10}$$

the cluster center $\underline{J}^j_{norm}$ is accepted and then the potential of the other solutions are further reduced as in (9). If

$$P\left(\underline{J}^j_{norm}\right) \leqslant \underline{\varepsilon} P\left(\underline{J}^1_{norm}\right) \tag{11}$$

the cluster center $\underline{J}^j_{norm}$ is rejected and the clustering process terminates. If neither (10) nor (11) are satisfied, then the acceptance criterion becomes:

$$\frac{d_{\min}}{r_a} + \frac{P\left(\underline{J}^j_{norm}\right)}{P\left(\underline{J}^1_{norm}\right)} \geqslant 1 \tag{12}$$

where $d_{\min}$ is defined as :

$$d_{\min} = \min_h \left\|\underline{J}^j_{norm} - \underline{J}^h_{norm}\right\|_2 \quad \text{with } h = 1,\ldots,j-1 \tag{13}$$

Finally, when all the cluster centers have been computed, a membership function matrix μ is found using the standard Gaussian distribution:

$$\mu_{j,i} = e^{-\alpha\left\|\underline{J}_{norm}(\theta^i)-\underline{J}^j_{norm}\right\|^2} \tag{14}$$

A solution $\underline{J}_{norm}(\theta^i)$ is then assigned to the jth cluster F^j, with $j = 1,\ldots,K$ if

$$\mu_{j,i} = \max_p \mu_{p,i}, \quad p = 1,\ldots,K. \tag{15}$$

Since the membership function (16) is computed as a negative power of the distance between the solution $\underline{J}_{norm}(\theta^i)$ and the cluster centers $\underline{J}^j_{norm}$, the smaller the distance, the higher the membership function.

The subtractive clustering algorithm allows avoiding the problem of objective function minimization; also, there is no random initialization so that the results are not dependent on the initial cluster centers or membership function choice; finally, the cluster centers $\underline{J}^j_{norm},\ j = 1,\ldots,K$, are objective values vectors corresponding to existing solutions in the Pareto Front and Set and thus could be used directly as representative solutions on the reduced Pareto Front.

3.2. Application of subtractive clustering to the case study of Section 2

The algorithm presented in Section 3.1 has been applied for clustering the Pareto Front solutions of the case study presented in Section 2.

The default input parameters are given in Table 3. They have been set after analysing the results of the clustering obtained with different values of the parameters.

The cluster radius r_a, which influences directly the number K of clusters obtained by the subtractive clustering algorithm, is an important parameter in the Pareto Front reduction procedure because it determines how many families (clusters), and consequently representative solutions, the Pareto Front is reduced to. To optimally set it, the so called global silhouette value is used to evaluate the quality of the clustering allocation (Rousseeuw, 1987; Rousseeuw et al., 1989). For any cluster partition of the Pareto Front, a global silhouette index, GS, is computed as follows:

$$GS = \frac{1}{K}\sum_{j=1}^{K} S_j \tag{16}$$

where S_j is the cluster silhouette of the jth cluster F^j indicating the heterogeneity and isolation properties of the cluster; it is computed as the average value of the silhouette widths $s(i)$ of the solutions in the cluster j, defined as:

$$s(i) = \frac{b(i) - a(i)}{\max\{a(i), b(i)\}} \tag{17}$$

where, $a(i)$ is the average distance from the ith solution of all the other solutions in the cluster, and $b(i)$ is the average distance from the ith solution of all the solutions in the nearest neighbor cluster, identified as the cluster with the minimum average distance of all its solutions from the ith solution.

The value of $s(i)$ is between +1 and -1; a value of +1, indicates solutions that are very distant from neighboring clusters; a value of 0, indicates solutions that are not distinctly in one cluster or another; a value of -1, indicates solutions that are probably assigned to the wrong cluster.

The maximum global silhouette can be evaluated to find the number of clusters (Rousseeuw, 1987).

The values of the global silhouette obtained for different values of r_a with default settings are given in Fig. 3: the maximum global silhouette value is reached for $r_a = 0.18$ with a resulting number of clusters K of 10.

The results of the subtractive clustering are showed in Fig. 4. The results are satisfactory in spite of the peculiar characteristics of the Pareto Front of the redundancy allocation problem, which is not "regular" and "smooth" and would pose significant

Table 3
Default input parameters for the subtractive clustering algorithm.

q	$\overline{\varepsilon}$	$\underline{\varepsilon}$
1.25	0.5	0.15

challenges to the minimization algorithms of k-means and fuzzy c-means because of the local minima present in the cost function of the clustering algorithm. This confirms the advantage of using subtractive clustering for reducing the number of Pareto Front solutions to be presented to the DM, independently on the Front characteristics and the number of the solutions therein.

3.3. Cluster representative solution selection

Each family of solutions F^j, $j = 1,\dots,K$, obtained by the subtractive clustering algorithm contains a number n^j of solutions of similar characteristics in the objective functions space and whose best representative solution, the head of the family $\underline{H}^j = \left(H_1^j \ \dots \ H_s^j \ \dots \ H_{N_{obj}}^j\right)$, needs to be found accounting for the DM specific requirements.

The DM would then be provided with only this reduced number K of solutions $\underline{H}^j$, $j = 1,\dots,K$, which best represent the Pareto Front under analysis and from which to select the final solution of the optimization problem.

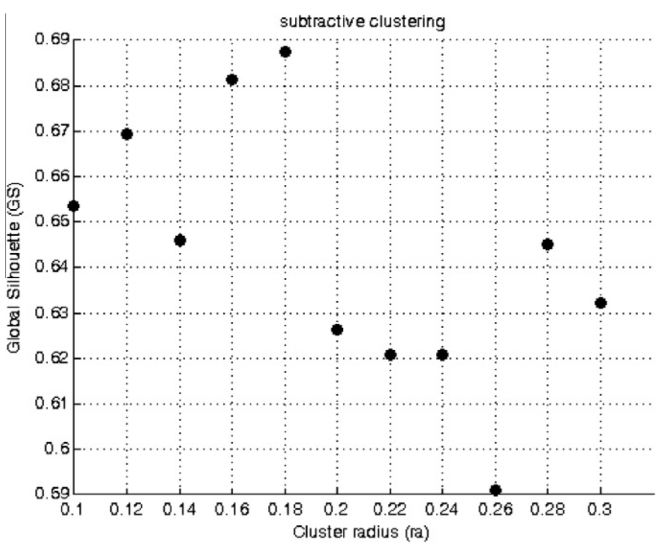

Fig. 3. GS for different cluster radius values.

The cluster representative solutions can be visualized using Level Diagrams (Blasco et al., 2008), which account for the distance of the Pareto Front and Set solutions from the ideal solution, optimal with respect to all the objectives simultaneously. Considering a multiobjective problem with l objectives to be minimized and m maximized (such that $N_{obj} = l + m$), n solutions in the Pareto Set and indicating by $\underline{J}(\theta^i) = (J_1(\theta^i) \dots J_s(\theta^i) \dots J_{Nobj}(\theta^i))$ the objective functions values vector corresponding to the solution θ^i, $i = 1,\dots,n$, each objective value $J_s(\theta^i)$, $s = 1,\dots,N_{obj}$, can be normalized with respect to its minimum and maximum values ($J_s^{\min}$ and $J_s^{\max}$) on the Pareto Front (Blasco et al., 2008):

$$\overline{J}_s(\theta^i) = \frac{J_s(\theta^i) - J_s^{\min}}{J_s^{\max} - J_s^{\min}}, \quad s = 1,\dots,l \tag{18}$$

and

$$\overline{J}_s(\theta^i) = \frac{J_s^{\max} - J_s(\theta^i)}{J_s^{\max} - J_s^{\min}}, \quad s = 1,\dots,m \tag{19}$$

so that now,

$$0 \leqslant \overline{J}_s(\theta^i) \leqslant 1, \quad s = 1,\dots,N_{obj} \tag{20}$$

where

- $\overline{J}_s(\theta^i) = 0$ means that the solution θ^i has the best value for the sth objective
- $\overline{J}_s(\theta^i) = 1$ means that the solution θ^i has the worst value for the sth objective.

To evaluate the distance to the ideal point

$$\theta^* : \overline{J}_s(\theta^i) = 0, \ \forall s = 1,\dots,N_{obj} \tag{21}$$

a suitable norm must be introduced (Blasco et al., 2008). Different norms can give different views on the characteristics of the Pareto Front and Set. In this paper, the norm considered is the following 1-norm:

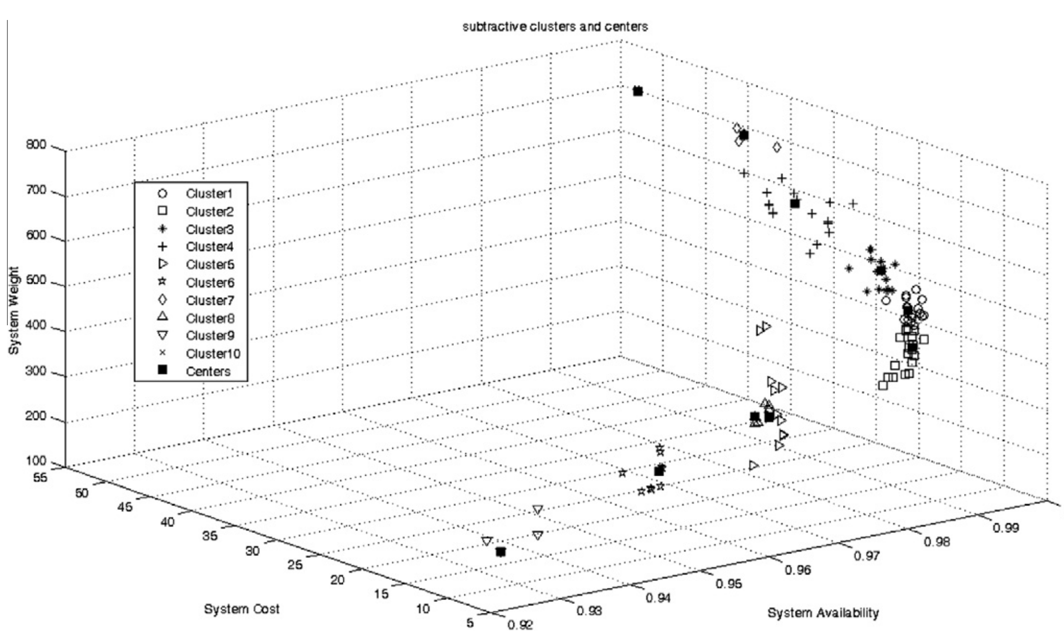

Fig. 4. Solutions clusters and centers obtained with the subtractive clustering algorithm in the objective function space of the redundancy allocation case study (Section 2).

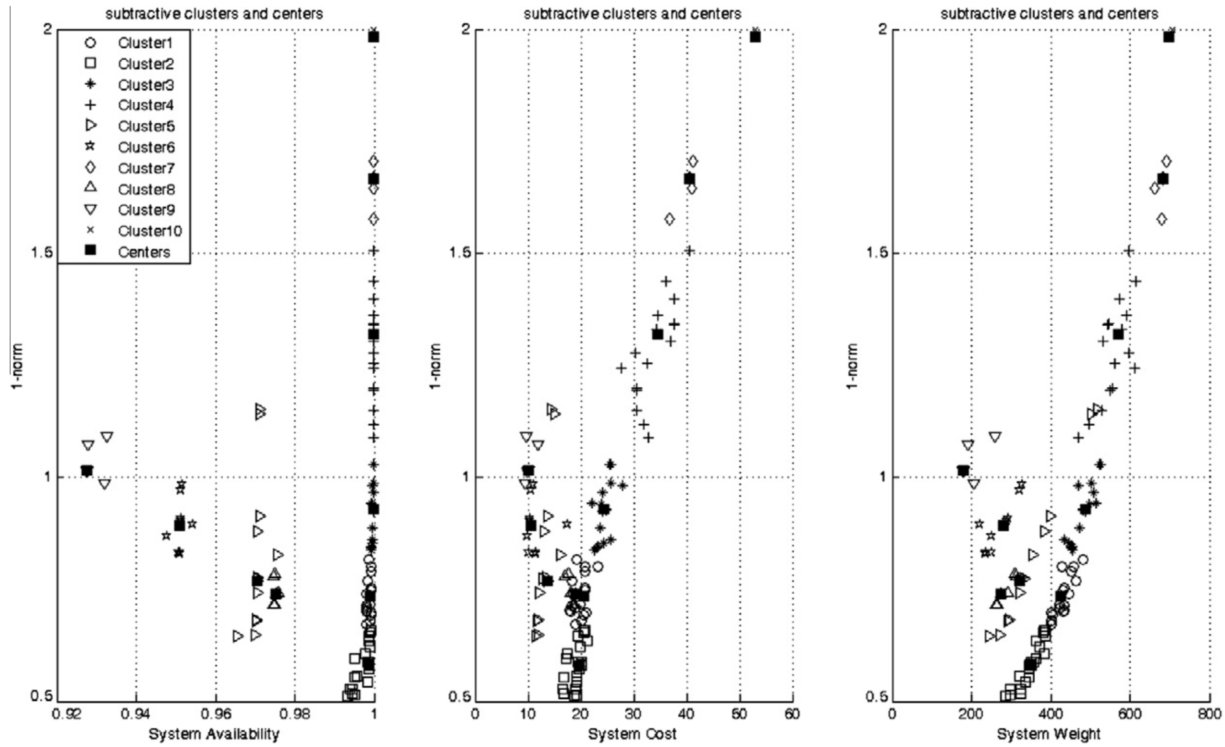

Fig. 5. Level Diagrams representation of the Pareto Front of the redundancy allocation case study.

$$\text{1-norm}: \left\| \underline{J}(\theta^i) \right\|_1 = \sum_{s=1}^{Nobj} \bar{J}_s(\theta^i), \quad \text{with } 0 \leqslant \|\bar{J}(\theta^i)\|_1 \leqslant s \tag{22}$$

The choice of this norm is due to the fact that it takes into account all the objectives, and thus can be used to analyse the overall performance of the solution (contrary to, for example, an infinite norm which would account only for the worst objective, i.e. the farthest from the ideal value). The plot of Level Diagrams is then done as follows: each objective J_s, $s = 1,\ldots,N_{obj}$, is plotted separately; the X axis corresponds to the objective in physical units of measurement, while the Y axis corresponds, for all the graphs, to the value $\|\underline{J}(\theta^i)\|_1$. This means that all the plots are synchronized with respect to the Y axis, i.e., all the information for a single solution of the Pareto Set will be plotted at the same level of the Y axis.

The clustered Pareto Fronts of the redundancy allocation case study of Section 2 is represented by Level Diagrams in Fig. 5.

When analysing the Pareto Front, the DM either:

1. analyses the Pareto Front to find the solution closest to the ideal one, i.e., that which optimizes all the objectives simultaneously;
2. applies his or her preferences on the objective functions values to identify the best solution according to these preferences.

The two decision situations, i.e., in presence or absence of preferences on the objectives values, may lead to the selection of different solutions and require different procedures of reduction of the solutions in the Pareto Front.

3.3.1. Selection of the families representative solutions in a decision situation with no preferences on the objective values

When preferences on the objective values are not given, the best solution is reasonably the one closest to the ideal solution, which optimizes all the objectives simultaneously.

By using the subtractive clustering algorithm, the cluster centers identified $\underline{J}^j_{norm}$, $j = 1,\ldots,K$, are solutions belonging to the Pareto Front, and can righteously be used as heads of the families $\underline{H}^j$. These solutions can be depicted on Level Diagrams according to their distances from the optimal solution. However, the cluster centers $\underline{J}^j_{norm}$ are identified by optimizing the distances among the solutions of the Pareto Front, with no consideration given to their relative distance from the ideal solution; in the family F^j, containing n^j solutions, there might be a solution $\underline{J}(\theta^i)_k$, $k = 1,\ldots,n^j$, closer to the ideal solution than the cluster center $\underline{J}^j_{norm}$.

For this reason, the head $\underline{H}^j$ of the jth generic family F^j is chosen as the solution with the lowest 1-norm value among the n^j solutions belonging to the cluster F^j, which implies, according to the Level Diagram definition, that this is the solution of the family closest to the optimal solution, ideal with respect to all the objectives:

$$\|\underline{H}^j\|_1 = \min \|\underline{J}(\theta^i)_k\|_1, \quad k = 1,\ldots,n^j \quad \text{and } j = 1,\ldots,K \tag{23}$$

The results of this procedure of selection of the family representatives applied to the case study considered is represented in Fig. 6, which shows that the Pareto Front shape is maintained after reduction, making it easier for the DM to analyse the Pareto Front which is now more intelligible and neat.

3.3.2. Selection of the families representative solutions in a decision situation with preferences on the objective values

If the DM preferences on the objective functions values are available, the solutions with minimum 1-norm value might not be the preferred ones. The solutions most relevant for the DM are those which are best with respect to the DM preferences.

In this work, the assignment of the DM preferences to the solutions is done by assigning objective values thresholds, which define different classes of merit. The objective values thresholds are given in a preference matrix P ($N_{obj} \times C$), where C is the number of objective functions thresholds used for the classification,

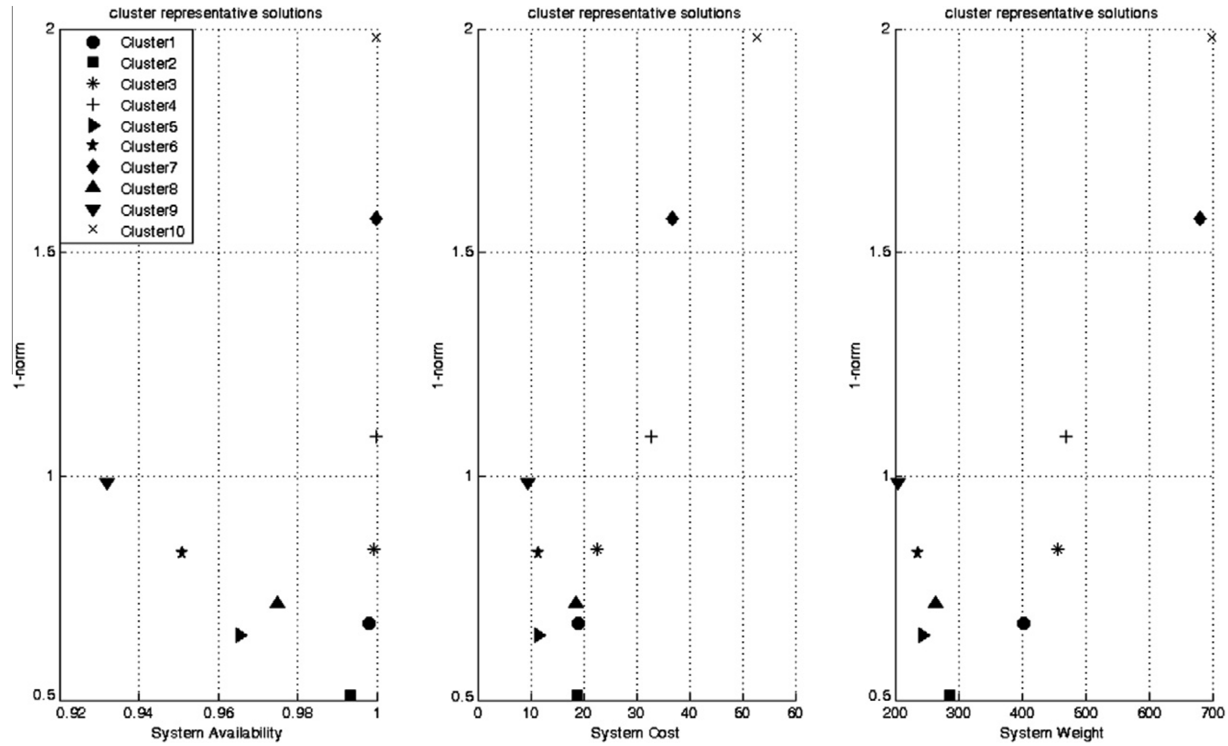

Fig. 6. Level Diagrams representation of the family representative solutions closest to the ideal solution optimizing all objectives, for the case study of reliability allocation problem (Section 2).

defining $C+1$ preference classes as in Fig. 7 (Blasco et al., 2008), where J_s^Z, $Z=1,\ldots,5$ are the thresholds values of the sth objective, l and m are the number of objectives to be minimized and maximized, respectively.

The fuzzy scoring procedure introduced by the authors in (Zio and Bazzo, 2010b) is then applied: each preference class is assigned a score $sv(r)$ (Blasco et al., 2008), with $r=1,\ldots,C+1$, such that

$$sv(C+1)=0; \quad sv(r)=N_{obj}\cdot sv(r+1)+1, \quad \text{for } r=C,\ldots,1 \tag{24}$$

and each objective value $J_s(\theta^i)$, with $i=1,\ldots,n$ and $s=1,\ldots,N_{obj}$, is assigned a membership function $\mu_{A_s^r}(J_s(\theta^i))$ which represents the degree with which $J_s(\theta^i)$ is compatible with the fact of belonging to the rth preference class, with $r=1,\ldots,C+1$.

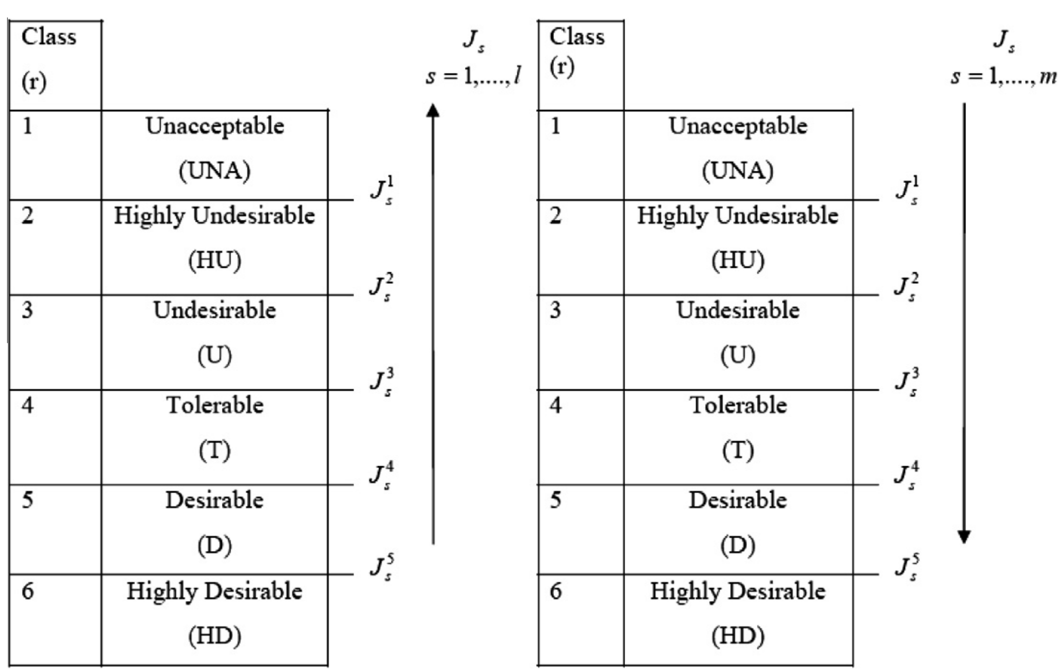

Fig. 7. Preferences thresholds.

A vector of $C+1=6$ membership functions is then defined for each objective J_s:

$$\underline{\mu}(J_s(\theta^i)) = \Big(\mu_{A_s^1}(J_s(\theta^i))\ \mu_{A_s^2}(J_s(\theta^i))\ \mu_{A_s^3}(J_s(\theta^i))\ \mu_{A_s^4}(J_s(\theta^i))\ \mu_{A_s^5}(J_s(\theta^i))\ \mu_{A_s^6}(J_s(\theta^i))\Big) \quad (25)$$

with $i=1,\ldots,n$, $s=1,\ldots,N_{obj}$.

The membership-weighted score of each individual objective is computed; given the scoring vector $\underline{sv} = (sv(1)\ sv(2)\ \ldots\ sv(C+1))$, whose components are defined in (26), and the membership functions vector $\underline{\mu}(J_s(\theta^i))$ in (25) for the ith solution and sth objective function, the score sv_s^i of the individual objective J_s is obtained by weighting the score $sv(r_s)$ of each class r_s the objective belongs to, by the respective membership function value $\mu_{A_s^{r_s}}(J_s(\theta^i))$, $r_s=1,\ldots,6$, and then summing the 6 resulting terms. This can be formulated in terms of the scalar product of the vectors $\underline{\mu}\left(J_s^i\right)$ and $\underline{sv}$ as follows:

$$sv_s^i = \frac{\left\langle \underline{\mu}(J_s(\theta^i)), \underline{sv}\right\rangle}{\sum_{r_s=1}^{6}\mu_{A_s^{r_s}}(J_s(\theta^i))}, \quad \text{with } i=1,\ldots,n \text{ and } s=1,\ldots,N_{obj} \quad (26)$$

where the denominator serves as the normalization factor.

Then, the score $S(\underline{J}(\theta^i))$ of the ith solution is the sum of the scores of the individual objectives

$$S(\underline{J}(\theta^i)) = \sum_{s=1}^{N_{obj}} sv_s^i, \quad \text{with } i=1,\ldots,n \quad (27)$$

and the lowest score is taken as the most preferred solution.

According to this fuzzy scoring procedure, the head $\underline{H}^j$ of the generic family F^j, $j=1,\ldots,K$, is chosen as the solution in F^j with lowest scores $S(\underline{J}(\theta^i))$:

$$S(\underline{H}^j) = \min S(\underline{J}(\theta^i)_k), \quad k=1,\ldots,n^j \quad \text{and } j=1,\ldots,K \quad (28)$$

Table 4
Preference threshold matrix P for the redundancy allocation case study.

	J_s^1	J_s^2	J_s^3	J_s^4	J_s^5
J_1 (availability)	0.960	0.965	0.970	0.980	0.985
J_2 (cost in arbitrary units)	35	30	25	22.5	20
J_3 (weight in arbitrary units)	350	325	300	275	250

For illustration purposes, let us introduce an arbitrary preference matrix P for the redundancy allocation case study (see Table 4). The resulting family representative solutions are visualized in Fig. 8 by Level Diagrams combined with a colouring methodology to represent the different scores $S(H^j)$ (Blasco et al., 2008): the Pareto Front shape is maintained after reduction allowing the DM to analyse the reduced and neat Pareto Front and Set, as in the previous case.

Note that the preference assignment is particularly critical. The DM sets preferences based on experience related to, but possibly not perfectly fitting with, the particular optimization problem, and/or valid only in a local region of the Pareto Front. A consequence might be that the most preferred solution is actually non-optimal for the problem at hand, in terms of its proximity to the ideal point. In this sense, the representation of the reduced Pareto Front combined with the colouring methodology, allows the DM to acknowledge the impact of his or her preferences on the Pareto Front and, in case, also to understand if and how preferences should be modified to reach the optimal region of the Front.

4. Discussion

The proposed procedure for reducing the solutions of the Pareto Front and Set for presentation to the DM is summarized in Fig. 9.

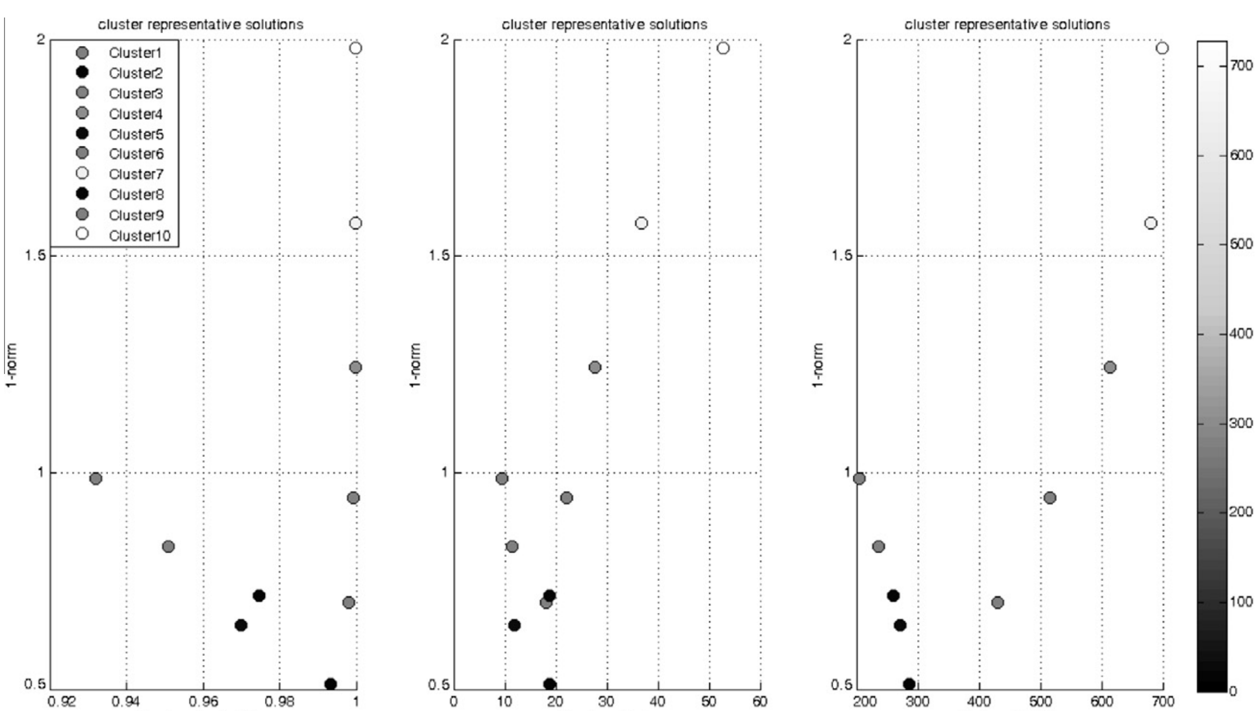

Fig. 8. Level Diagrams representation of the family representative solutions with lowest score $S(H^j)$, for the case study of reliability allocation (Section 2).

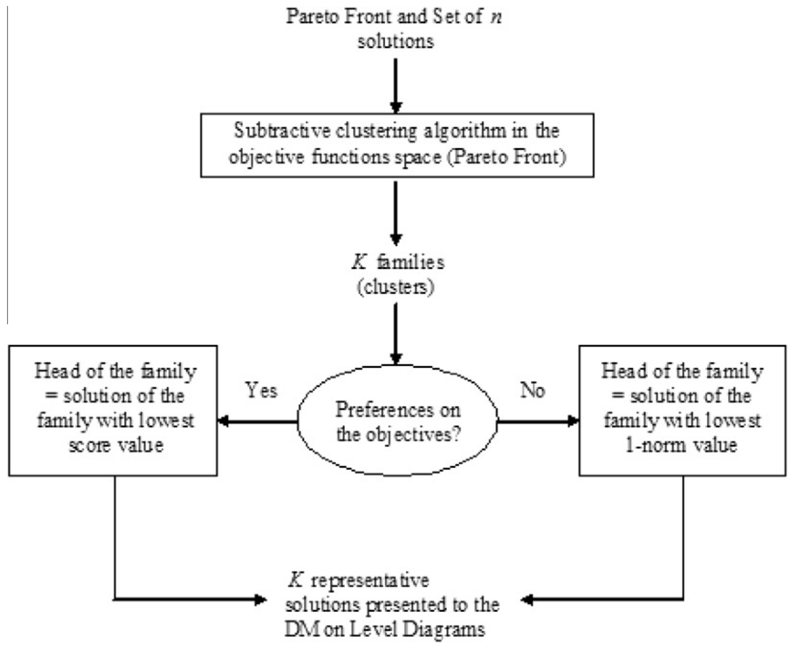

Fig. 9. Schematics of the procedure for reducing the Pareto Front and Set.

The combination of the clustering algorithm and the selection of the minimum 1-norm solutions as representative of the clustered families is a sound generalization of the empirical method introduced by the authors in (Zio and Bazzo, 2010a), where families of solutions were identified by looking for "vertical alignments" in the 1-norm Level Diagram of the availability objective (Fig. 5). Since the differences of the availability values of two successively ranked solutions, i.e., $J_1\ (\theta^{i+1}) - J_1(\theta^i)$, $i = 1,\ldots,n$, are very different in the two branches of the Front (Fig. 5), two distinct criteria were needed to decide whether two solutions can be considered vertically aligned in a cluster on the left and right of the minimum value of the norm. Given $i_{\min}$, the position in the objective J_1 of the solution with the lowest value of the 1-norm, the solution θ^{i+1} corresponding to $J_1(\theta^{i+1})$ belongs to a vertical cluster alignment if the previous solution θ^i is such that:

$$J_1(\theta^{i+1}) - J_1(\theta^i) \leqslant 10^{-3} \quad \text{for } 1 \leqslant i \leqslant i_{\min} \tag{29}$$

and

$$J_1(\theta^{i+1}) - J_1(\theta^i) \leqslant 10^{-5} \quad \text{for } i_{\min} < i \leqslant n_j \tag{30}$$

Then, the solutions belonging to a "vertical alignment" can be reduced to the minimum 1-norm representative solution. Such empirical method was shown effective in reducing the Pareto Front to a smaller number of solutions from 118 to 52, but was tailored on the characteristics of the particular Pareto Front, and thus of limited applicability.

5. Conclusions

Multiobjective decision-making is the process of choosing a possible course of action among the alternative solutions available, which are judged preferentially by one or more DMs with respect to several conflicting criteria. In fulfilling the conflicting goals, the DMs must account also for the constraints imposed by the system itself.

The issue can be formulated in terms of a multiobjective optimization problem whose solving produces a Pareto Set of non-dominated solutions among which the DMs have to select the preferred ones. The selection is difficult because the set of non-dominated solutions is usually large, and it is well known that presenting the DMs with many alternatives renders their decision-making task burdensome. For this reason, decision-making aid tools are developed to drive the search for the optimal solution, reducing the burden of the analysis.

In this work, a procedure for reducing the set of solutions for presentation to the DM is proposed, based on clustering. The non-dominated solutions of the Pareto Set are grouped in families by subtractive clustering, according to their relative distance in the objective functions space (Pareto Front); the parameters driving the subtractive clustering are chosen to maximize the quality of the resulting Pareto Front partition; a head of the family is then chosen as best representative solution of each family, with respect to the DM requirements; Level Diagrams are finally used to represent and analyse the reduced Pareto Front thereby obtained.

Two approaches to the selection of the heads of the families have been introduced to handle two different decision situations, depending on the presence or absence of DM preferences on the objectives. In the former situation, the representative solutions are sought by referring to their distance from the ideal solution (which optimizes all objectives simultaneously); in the latter situation, a fuzzy scoring procedure is used for ranking solution alternatives.

The procedure has been applied to a reference case study concerning a redundancy allocation problem. The results show that the reduction procedure makes it easier for the decision maker to select the final solution and allows him or her to discuss the outcomes of the optimization process on the basis of his or her assumed preferences. This is possible thanks to the clustering technique which is showed to maintain the Pareto Front shape and relevant characteristics.

The subtractive clustering algorithm, with parameters optimized to maximize the quality of the partition, can be applied to Pareto Fronts with different characteristics, and thus it is suitable for general application. Extension of the clustering algorithm to mathematical forms of distance other than the euclidean (e.g. Mahalanobis) might be worth investigating to provide more flexibility to the clusters in terms of shape and orientation; both the benefits and the costs of application of the thereby extended algorithm should be evaluated.

The overall proposed procedure is fit to be applied a posteriori of the optimality search, for reducing the set of solutions in the Pareto Front to a small number of representative ones; on the other hand, it can serve also as initial screening within a deliberative process of decision-making, possibly interactive and guided by techniques of preference elicitation and inclusion for direct ranking and elimination of solutions.

Acknowledgments

The authors are thankful to Professor David Coit of Rutgers University for providing the Pareto Front and Set data of the case study, and to the three anonymous reviewers for providing in-depth comments which have stimulated a thorough revision of the paper, for its improvement.

References

Bandyopadhyay, S., Maulik, U., 2002. An evolutionary technique based on *k*-means algorithm for optimal clustering in R^N. Information Science 146, 221–237.

Bezdek, J., 1974. Cluster validity with fuzzy sets. Journal of Cybernetics 3 (3), 58–71.

Blasco, X., Herrero, J.M., Sanchis, J., Martínez, M., 2008. A new graphical visualization of *n*-dimensional Pareto Front for decision-making in multiobjective optimization. Information Science 178, 3908–3924.

Chiu, S., 1994. Fuzzy model identification based on cluster estimation. Journal of Intelligent and Fuzzy Systems 2 (3), : 1240–1245.

Cho, K.I., Kim, S.H., 1997. An improved Interactive hybrid method for the linear multi-objective knapsack problem. Computers and Operations Research 24 (11), 991–1003.

De Boer, L., van der Wegen, L., Telgen, J., 1998. Outranking methods in support of supplier selection. European Journal of Purchasing and Supply Management 4, 109–118.

Katagiri, H., Sakawa, M., Kato, K., Nishizaki, I., 2008. Interactive multiobjective fuzzy random linear programming: Maximization of possibility and probability. European Journal of Operational Research 188, 530–539.

Levitin, G., 2005. Universal generating function in reliability analysis and optimization. Springer-Verlag, London.

Levitin, G., Lisnianski, A., 2001. A new approach to solving problems of multi-state system reliability optimization. Quality and Reliability Engineering International 17 (2), 93–104.

Malakooti, B., 1988. A decision support system and a heuristic interactive approach for solving discrete multiple criteria problems. IEEE Transactions on System, Man and Cybernetics 18, 273–284.

Molina, J., Santana, L.V., Hernandez-Diaz, A.G., Coello Coello, C.A., Caballero, R., 2009. *g*-Dominance: Reference point based dominance for multiobjective metaheuristics. European Journal of Operational Research 197, 658–692.

Rios Insua, D., Martin, J., 1994. Robustness issue under imprecise beliefs and preferences. Journal of Statistical Planning and Inference 40 (2–3), 383–389.

Rousseeuw, P.J., 1987. Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. Journal of Computational and Applied Mathematics 20, 53–65.

Rousseeuw, P., Trauwaert, E., Kaufman, L., 1989. Some silhouette-based graphics for clustering interpretation. Belgian Journal of Operations Research, Statistics and Computer Science 29 (3).

Roy, B., 1968a. Classement et choix en presence de points de vue multiples (la methode ELECTRE). RIRO 8, 57–75.

Roy, B, 1968. Criteres multiples et modelisation des preferences: l'apport des relations de surclassement revue d'Economie Politique.

Roy, B., 1974. Criteres multiples et modelisation des preferences: l'apport des relations de surclassement revue d'Economie Politique 84, 1–44.

Roy, B., Bouyssou, D., 1986. Comparison of two Decision-Aid Models Applied to a Nuclear Power Plant Siting Example. European Journal of Operational Research 25, 200–215.

Selim, S.Z., Ismail, M.A., 1984. *K*-means type algorithms: A generalized convergence theorem and characterization of local optimality. IEEE Transactions on Pattern Analysis and Machine Intelligence 6, 81–87.

Spath, H., 1985. Cluster Dissection and Analysis: Theory, FORTRAN Programs, Examples, New York: Halsted Press.

Taboada, H. and Coit, D, 2007. Recent Developed Evolutionary Algorithms for the Multi- Objective Optimization of Design Allocation Problems, in: Proceedings of the 5th International Conference on Quality & Reliability (ICQR5), Chiang Mai, Thailand.

Taboada, H., Espiritu, J., Coit, D., 2007. MOMS-GA: A multiobjective multi-state genetic algorithm for system reliability optimization design problems. IEEE Transactions on Reliability.

Tou, J.T., Gonzalez, R.C., 1974. Pattern Recognition Principles. Addison-Wesley, Reading.

Ushakov, I., 1986. A universal generating function. Soviet Journal of Computer and System Sciences 24 (5), 37–49.

Yang, J.B., 1996. Multiple Criteria Decision Making Methods and Applications. Hunan Publishing House, Changsha PR China.

Yang, J.B., 2000. Minimax reference point approach and its application for multiobjective optimisation. European Journal of Operational Research 126, 541–556.

Zio, E., Bazzo, R, 2010a. Level Diagrams for Decision Aiding in Multiobjective Optimization Problems, Proceedings of the European Safety and Reliability Conference, ESREL 2010, Rhodes, Greece, September 2010, CD-Rom.

Zio, E., Bazzo, R., 2010b. Optimization of the test intervals of a nuclear safety system by genetic algorithms, solution clustering and fuzzy preference assignment. Nuclear Engineering and Technology 42 (4).

Zio, E., Baraldi, P., Pedroni, N., 2009. Optimal power system generation scheduling by multi- objective genetic algorithms with preferences. Reliability Engineering and System Safety 94, 432–444.

Enrico Zio (BS in Nuclear Engng., Politecnico di Milano, 1991; MSc in Mechanical Engng., UCLA, 1995; PhD, in Nuclear Engng., Politecnico di Milano, 1995; PhD, in Nuclear Engng., MIT, 1998) is Director of the Chair in Complex Systems and the Energetic Challenge of Ecole Centrale Paris and Supelec, Director of the Graduate School of the Politecnico di Milano, full professor of Computational Methods for Safety and Risk Analysis, adjunct professor in Risk Analysis at the University of Stavanger, Norway, and invited lecturer and committee member at various Master and PhD Programs in Italy and abroad.

He has served as Vice-Chairman of the European Safety and Reliability Association, ESRA (2000–2005) and as Editor-in-Chief of the International journal Risk, Decision and Policy (2003–2004). He is currently the Chairman of the Italian Chapter of the IEEE Reliability Society (2001).

He is member of the editorial board of the International Scientific Journals Reliability Engineering and System Safety, Journal of Risk and Reliability, Journal of Science and Technology of Nuclear Installations, plus a number of others in the nuclear energy field.

He has functioned as Scientific Chairman of three International Conferences and as Associate General Chairman of two others, all in the field of Safety and Reliability.

His research topics are: analysis of the reliability, safety and security of complex systems under stationary and dynamic operation, particularly by Monte Carlo simulation methods; development of soft computing techniques (neural networks, fuzzy logic, genetic algorithms) for safety, reliability and maintenance applications, system monitoring, fault diagnosis and prognosis, and optimal design.

He is co-author of three international books and more than 150 papers on international journals, and serves as referee of more than 20 international journals.

Roberta Bazzo (BS in Energy Engineering, Politecnico di Milano, 2007, MS in Nuclear Engineering, Politecnico di Milano, 2009). Her interests include Pareto Front and Set analyses for decision making in RAMS and banking applications.