A new parsimonious AHP methodology: Assigning priorities to many objects by comparing pairwise few reference objects

A new parsimonious AHP methodology: assigning priorities to many objects by comparing pairwise few reference objects

Francesca Abastante[a], Salvatore Corrente[b], Salvatore Greco[b,c], Alessio Ishizaka[c], Isabella M. Lami[a]

[a] *Department of Regional and Urban Studies and Planning (DIST), Politecnico of Torino, Turin, Italy*
[b] *Department of Economics and Business, University of Catania, Italy*
[c] *CORL, Portsmouth Business School, University of Portsmouth, United Kingdom*

Abstract

We propose a development of the Analytic Hierarchy Process (AHP) permitting to use the methodology also for decision problems with a very large number of alternatives and several criteria. While the application of the original AHP method involves many pairwise comparisons between considered objects, that can be alternatives with respect to considered criteria or criteria between them, our parsimonious proposal is composed of five steps: (i) direct evaluation of the objects at hand; (ii) selection of some reference evaluations; (iii) application of the original AHP method to reference evaluations; (iv) check of the consistency of the pairwise comparisons of AHP and the compatibility between the rating and the prioritization with a subsequent discussion with the decision maker who can modify the rating or pairwise comparisons of reference points; (v) revision of the direct evaluation on the basis of the prioritization supplied by AHP on reference evaluations. Our approach permits to avoid the distortion of comparing more relevant objects (reference points) with less relevant objects. Moreover, our AHP approach avoids rank reversal problems, that is, changes of the order in the prioritizations due to adding or removing one or more objects from the set of considered objects. The new proposal has been tested and experimentally validated.

Keywords: Analytic Hierarchy Process, Parsimonious preference information, Pairwise comparison matrix.

1. Introduction

Our life is the result of our decisions. Most of our decisions are taken on the basis of our intuition and common sense. Most complex decisions require instead a more systematic approach and the adoption of proper methodologies of decision support. Since complex decision problems require to take into consideration a plurality of points of view, technically called criteria, concepts

Email addresses: francesca.abastante@polito.it (Francesca Abastante), salvatore.corrente@unict.it (Salvatore Corrente), salgreco@unict.it (Salvatore Greco), alessio.ishizaka@port.ac.uk (Alessio Ishizaka), isabella.lami@polito.it (Isabella M. Lami)

[Post-print] Please cite as: Abastante et al. 2019. A new parsimonious AHP methodology: assigning priorities to many objects by comparing pairwise few reference objects, Expert Systems with Applications, DOI: 10.1016/j.eswa.2019.02.036

and methodologies of Multiple Criteria Decision Aiding (MCDA, see e.g. Ishizaka and Nemery, 2013 and Greco et al., 2016) are very useful in this context. The main point of MCDA is the aggregation of alternatives' evaluations and the comparisons of feasible alternatives on the family of heterogeneous criteria. These comparisons require the measurement of the attractiveness of each alternative with respect to each criterion by assignment of reliable priorities. A similar prioritization is required also to assign a weight to the criteria in the family. The priorities' assignment must represent perceptions and preferences of the Decision Maker (DM) that, with this aim, can be requested to supply two different types of judgments (Blumenthal, 1977):

- absolute judgments representing the magnitude of a stimulus,
- comparative judgments representing the relations between two stimuli.

Absolute judgments seem to have a great disadvantage with respect to comparative judgments. Indeed, they involve a relation between a single stimulus and other analogous stimuli stored in the memory implying a great effort for the already limited capability of our brain (see e.g. Miller, 1956). This would suggest to adopt methodologies assigning priorities on the basis of comparative judgments that, instead, are based on comparison of immediate impressions both present to the DM. On the basis of this considerations, the Analytical Hierarchy Process (AHP) (Saaty, 1977, 1980) has been proposed to build the prioritization of each alternative based on the pairwise comparisons of each object with all the others. The great success of AHP, being one of the most adopted MCDA methods, confirms that DMs and experts appreciate to deal with comparative judgments. However, there is also a considerable limitation for AHP as well as for all other methods based on pairwise judgments such as MACBETH (Bana e Costa and Vansnick, 1994). In fact these methods require from the DMs involved in the decision problem to supply a huge quantity of information. For example, in a not so big decision problem presenting 21 alternatives and 9 criteria, it would be necessary to ask 210 pairwise comparisons for each one of the non-numerical criteria, that are 1,890 pairwise comparisons in total! This is a huge amount of preference information required. To give an idea of the effort asked to the DM, let us suppose that for each pairwise comparison he needs one minute in average. Accordingly, he would need 31 hours and half to supply all these preferences. Realistically, this cannot be asked to the DM. This bottleneck is well known in the literature on AHP: for example Saaty and Ozdemir (2003) demonstrate that the number of elements to be considered in AHP should be no more than seven. Thus, what to do? Should we renounce to use AHP, as well as other methods based on comparative judgment, to deal with decision problems presenting more than seven alternatives? As the reader can imagine, this is a fundamental problem for the application of a very well-known and appreciated method such as AHP. Indeed, real world decision problems very often have more than seven alternatives and, even more, the case in which the alternatives are seven or less is really a small subset of the family of decision problems we encounter in real life. Consequently, several approaches have been developed to reduce the number of pairwise comparisons. Let us remember some of them.

A first class of approaches use clustering technique to produce a granulation of the feasible evaluations, so that the pairwise comparisons are requested on a family of sets of more or less homogeneous evaluations considered as a whole. Already in the original publication, Saaty (1977) proposes to group similar objects into clusters of no more than 7 elements. The DM would compare the alternatives within the clusters and then the clusters themselves. Ishizaka (2012) developed a variant where each cluster has a common object with neighbouring clusters. This permits to join clusters without compare them in a separate matrix. In both cases, the number of pairwise comparisons

still increases exponentially.
A second approach to reduce the pairwise comparisons requested to the DM is based on the idea of a partial fulfillment of the pairwise comparison matrix. The DM supplies some specific pairwise comparisons of the matrix, typically those he is more confident in, and the missing values are then estimated from the given evaluations. Several techniques have been developed for this purpose (Benítez et al., 2014; Bozóki et al., 2010; Chen et al., 2015; Csató and Rónyai, 2016; Fedrizzi and Giove, 2007; Gomez-Ruiz et al., 2010; Harker, 1987). This approach presents two main questions: Which evaluations do I need to provide? How many evaluations do I need to provide to have a satisfactory vector of priorities? Carmoner et al. (1997) found that 50% of the evaluations need to be provided, which is only a reduction of half of the pairwise comparisons required.
Another approach is based on the comparison of some benchmarking. In this case, instead of comparing each objects to all the others, they are compared only with a set of fixed objects as in AHPSort (Ishizaka et al., 2012) or in the Best-Worst method (Rezaei, 2015). The number of required evaluations is greatly reduced but it is still proportional to the number of objects. For example, in the case of the Best-Worst method, each object is compared with the best object and with the worst object so that in case of n objects, the DM has to supply $2n-3$ pairwise comparisons corresponding to the sum of

- $n-2$ comparisons of each object different from the best and the worst with the best object,
- $n-2$ comparisons of each object different from the best and the worst with the worst object,
- one comparison between the best and the worst object.

In this paper we present in detail and we test experimentally a new methodology permitting to use AHP also if the number of objects is relatively great. The methodology we propose permits to use AHP, nevertheless reducing the number of preference comparisons asked to the DM. The new proposal, that has to be repeated for each criterion, is composed of five steps: in the first step, the DM has to provide a direct rating of the objects; in the second step the DM, in accordance with the analyst, has to select some reference evaluations; in the third step, the DM has to pairwise compare the reference evaluations by using the original version of the AHP method in order to get their priorities; in the fourth step, the consistency of the pairwise comparisons of AHP and the compatibility between the rating and the prioritization obtained for the reference evaluations are checked and possibly discussed with the DM who can modify the rating or pairwise comparisons of reference points if he is willing to; in the fifth step, the priorities of all the other evaluations are obtained by interpolation according to the priority values obtained for the reference evaluations. The new proposal reduces in a considerable way the enormous cognitive effort asked to the DM in comparing alternatives and criteria in problems of big dimensions. However, it maintains the basic idea of AHP of relying on pairwise comparisons to prioritize the elements to be evaluated. Thus, we think that this procedure can be used to deal with any decision problem in which the application of original AHP is prevented by a number of alternatives larger than seven. Observe also that our methodology is strongly based on the concept of reference point that is so important in the real world decisions as explained by many models and results in behavioural economics, the most important of which related to the Prospect Theory of Kahneman (1979) (see also, for example, Barberis and Xiong, 2009; Camerer et al., 1997; Genesove and Mayer, 2001; Tversky and Kahneman, 1992). In this sense, by taking into account some reference points selected by the DM and asking him to compare them we obtain two important results:

- we take advantage of the natural tendency to focus on reference points to get more aware and reliable pairwise comparisons,
- we avoid the distortions of comparing non reference point objects with reference point ones.

Our approach has also the advantage to avoid rank reversal problems. Let us remember that in AHP the addition or deletion of one or more alternatives can modify the final rank (Belton and Gear, 1983). The study of the causes, the interpretation and the possible remedies for this problem have been the subject of a long discussion in the specialized literature (see Maleki and Zahir, 2013 for a comprehensive review and Krejčí and Stoklasa, 2018 for a recent paper dealing with this topic). Some salient points that need to be taken into account are the following. In case of multiple criteria, even fully consistent pairwise comparisons can originate a rank reversal and, in fact, this is the case considered in the seminal paper of Belton and Gear (1983) and in many other contributions on the subject. Moreover, in case of multiple criteria, there is no normalization that can prevent rank reversal. Instead, if there is only one criterion and the pairwise comparisons are fully consistent, there cannot be any rank reversal (Harker and Vargas, 1990; Saaty, 1990b). Anyway, Stam and Silva (2003) pointed out that rank reversal can be caused solely by inconsistency of pairwise comparisons even in the case of a single criterion. In general, simulation results (Raharjo and Endah, 2006) proved that the greater the number of alternatives and the more inconsistent the pairwise comparisons, the greater the probability of rank reversal.

Our approach permits to handle rank reversal with respect to any single criterion and, a fortiori, also in case of more than one criterion, because, once the reference points have been selected and the related pairwise comparison matrix has been obtained, the prioritization of all feasible performances remains fixed independently of the set of considered alternatives. The argument that in some situations the rank reversal can be justified (Leskinen and Kangas, 2005; Saaty and Vargas, 1984) is taken into consideration in our approach with respect to the reference points. Indeed, if the addition or the deletion of one or more reference points change the ranking of the other reference points, this can be discussed with the DM. In this cases, due to the generally small number of reference points, let us say never more than 9, the DM has the possibility to take into consideration all of them together following the general rule that our brain is able to control a number of element equal to "the magical number seven plus or minus two" (Miller, 1956). Therefore, the DM can reflect and ponder his preference information and he can verify the reasonableness of the prioritization of the reference points. In this way, in an interactive way, some pairwise comparisons can be modified until he is convinced and he accepts the values assigned by AHP to the reference points. This in-deep analysis cannot be conducted considering the whole set of alternatives when they are numerous. Therefore, we can say that regarding rank reversal our approach has two types of advantages with respect to the standard AHP:

- ex-ante advantage, due to the smaller probability of rank reversal in case of smaller number of elements to be considered that, in this case, are well-known to the DM, since they are reference points selected with his cooperation;
- ex-post advantage, due to the possibility of analyzing in detail the prioritization supplied by AHP because, again, the compared elements are few and well-known to the DM.

One can see that the methodology we are proposing requires that the inconsistency in absolute and comparative judgments provided by the DM has to be discussed with him. This can be seen as a week point since it requires a heavier involvement of the DM. However, according to Roy (2010) we believe that multiple criteria decision aiding must be based on models that are, at least partially, co-constructed through interaction with the DM and that the co-constructed model must be a tool for looking deeper into the subject, exploring, interpreting, debating and even arguing. In this perspective, we believe that the discussion of inconsistency in supplied preference information with the DM increases the reliability of the recommendation obtained through the decision process and, consequently, it has to be seen as strong point of our methodology.

In order to validate the new proposal, we conducted an experiment with around 100 students of the University of Catania (Italy) that were requested to evaluate the area of ten geometric figures. The students were split in two groups of similar size. All of them solved the same problem but one group applied the original AHP method, while the other used the methodology we are proposing. The results showed that the new proposal gives better results than those obtained using the original methodology.

The paper is organized as follow: Section 2 presents the basic concepts of AHP and its application to multiple criteria decision aiding; Section 3 presents the new methodology; Section 4 presents an experiment that shows the merits of the new approach; Section 5 presents a didactic example that illustrates how the parsimonious AHP we are proposing can be applied in multiple criteria decision aiding problem; last Section collects conclusions.

2. Methodological framework

2.1. Introduction to MCDA

Multiple Criteria Decision Aiding (MCDA; see Ishizaka and Nemery, 2013 for a general introduction and Greco et al., 2016 for an updated and comprehensive collection of state of the art surveys), considers a set $A = \{a, b, c, \ldots\}$ of alternatives evaluated with respect to a coherent family of points of view, technically called criteria, $G = \{g_1, g_2, \ldots, g_n\}$, with $g_j : A \rightarrow \mathbb{R}$, $j = 1, \ldots, n$, such that, without loss of generality, we can suppose that for all $a, b \in A$, $g_j(a) \geq g_j(b)$ means that a is at least as good as b with respect to g_j. Four main different decision problems can be considered with the methodologies developed within MCDA: ranking, choice, sorting and description (Roy, 1996). Since the only objective information stemming from the evaluations of the alternatives on the considered criteria, which is the dominance relation, is quite poor, different aggregation methods have been proposed in literature based on: value functions (Keeney and Raiffa, 1976), binary relations (Brans and Vincke, 1985; Roy, 1996) or decision rules (Greco et al., 2001). Value functions assign to each alternative a real value being representative of the goodness of the considered alternative with respect to the problem at hand. Binary relations are the basis of the outranking methods for which aSb means that a is at least as good as b. Decision rules connect the evaluations of the alternatives on the considered criteria with the recommendations on the problem at hand. For example, "If the maximum speed of a car is at least 140 km/h and its consumption is at least 17 km/l, then the car is considered at least good".

2.2. The Analytic Hierarchy Process

AHP (Saaty, 1977, 1990a) is an MCDA method based on ratio scales for measuring performances on considered criteria and the importance of these criteria (see Ishizaka and Labib, 2011 for a review of the main developments of AHP and Duleba and Moslem, 2019; Nazari et al., 2018 for some recent contributions related to AHP). The problem at hand is structured in AHP in a hierarchical way where the overall goal is set at the top of the hierarchy and the alternatives, being the object of the decision, are placed at the bottom of the hierarchy. The criteria on which the alternatives need to be evaluated are in the middle between the overall goal and the alternatives themselves. Given n criteria, the DM is supported to provide a value for the evaluations $g_j(a)$ ($g_j \in G$ and $a \in A$) and for the weights $w_1, \ldots, w_n$ in a weighted sum aggregation. The fundamental idea is that it is more convenient to perform pairwise judgments rather than giving direct evaluations of performances $g_j(a)$ and weights $w_1, \ldots, w_n$ as experimentally seen in Millet (1997) and Por and Budescu (2017). Using AHP, for each criterion $g_j \in G$, the DM is asked to compare each couple of alternatives $\{a_r, a_s\}$ indicating the preferred alternative and expressing the degree of preference with a verbal judgment on a nine point scale (Saaty, 1990a) defined as

- 1: indifferent,
- 3: moderately preferred,
- 5: strongly preferred,
- 7: very strongly preferred,
- 9: extremely preferred

with 2, 4, 6 and 8 intermediate values between the two adjacent judgments. Denoting by $a_{rs}^{(j)}$ the pairwise comparison between the priorities $e_r^{(j)}$ and $e_s^{(j)}$, it is possible to build a positive square reciprocal matrix $M^{(j)}$ of order $|A|$.

$$M^{(j)} = \begin{pmatrix} 1 & a_{12}^{(j)} & \cdots & a_{1|A|}^{(j)} \\ {}^{1}/_{a_{12}^{(j)}} & 1 & \cdots & a_{2|A|}^{(j)} \\ \vdots & \vdots & \ddots & \vdots \\ {}^{1}/_{a_{1|A|}^{(j)}} & {}^{1}/_{a_{2|A|}^{(j)}} & \cdots & 1 \end{pmatrix}$$

The basic idea is that if the comparative judgments are perfectly consistent, then $a_{rs}^{(j)} = \frac{e_r^{(j)}}{e_s^{(j)}}$ so that , the condition $a_{rs}^{(j)} a_{sk}^{(j)} = a_{rk}^{(j)}$ should be satisfied for all r, s, k. In this case to compute the $e_r^{(j)}$ for $r = 1, \ldots, |A|$, is quite straightforward. Indeed, for example, putting $e_s^{(j)} = 1$ for some $s = 1, \ldots, |A|$, one has

$$e_r^{(j)} = a_{rs}^{(j)}, r = 1, \ldots, |A|.$$

However, in general the comparative judgments of matrix $M^{(j)}$ are not consistent and, therefore, several procedures have been proposed to determine the evaluations $e_r^{(j)}$ and among them the most well-known is based on the computation of the right eigenvector of the $M^{(j)}$ matrix (Saaty, 1977), the row arithmetic mean vector, that is

$$e_r^{(j)} = \frac{\sum_{s=1}^{|A|} a_{rs}^{(j)}}{|A|}$$

and the row geometric mean vector (Crawford and Williams, 1985), that is

$$e_r^{(j)} = \sqrt[|A|]{\prod_{s=1}^{|A|} a_{rs}^{(j)}}.$$

Recently, a new approach has been proposed to compute priorities $e_r^{(j)}$ by taking into account all possible subsets of independent comparison judgments $a_{rs}^{(j)}$ (Tsyganok, 2010; Siraj et al., 2012), such as, for example, the set of comparisons judgments from row r in matrix $M^{(j)}$ (that is, $\{a_{r1}^{(j)}, \ldots, a_{r|A|}^{(j)}\}$) or the set of comparisons judgments from column s in matrix $M^{(j)}$ (that is, $\{a_{1s}^{(j)}, \ldots, a_{|A|s}^{(j)}\}$). Each one of these possible subsets of independent comparisons corresponds to a spanning tree whose nodes are $r, r = 1, \ldots, |A|$. To each spanning tree corresponds a prioritization. For example, to the spanning tree related to row r in matrix $M^{(j)}$ corresponds the prioritization

$$e_s^{(j)} = \frac{1}{a_{rs}^{(j)}}, s = 1, \ldots, |A|$$

while, to the spanning tree related to column s in matrix $M^{(j)}$ corresponds the prioritization

$$e_r^{(j)} = a_{rs}^{(j)}, r = 1, \ldots, |A|.$$

Finally, an overall prioritization is obtained as an average of the priorities corresponding to all the spanning trees. More precisely, the arithmetic mean has been proposed in Tsyganok (2010); Siraj et al. (2012) and the geometric mean in Lundy et al. (2017).
If the entries of the matrix would be perfectly consistent, the right eigenvector, the row arithmetic mean and the row geometric mean, as well as all possible subsets of independent comparison judgments corresponding to the spanning trees should give the same evaluations $e_r^{(j)}, r = 1, \ldots, |A|$. As this is not the case in real world applications, the consistency of the judgments supplied by the DM can be tested computing the consistency ratio $CR = CI/RI$ where $CI = (\lambda_{max} - |A|)/(|A| - 1)$ is the consistency index, λ_{max} is the principal eigenvalue of A and RI is the ratio index. The ratio index is the average of the consistency indices of 500 randomly filled matrices. Saaty (1977) considers that a consistency ratio exceeding 10% may indicate a set of judgments too inconsistent to be reliable and therefore it is recommended to revise the evaluations. Let us underline that even if in the following we are considering a CR threshold equal to 10%, in literature there is a lively discussion on the proper value of this threshold and on the fact that a threshold greater than 0.1 does not imply necessarily an intolerable inconsistency in the provided judgments. For example, in Ishizaka and Siraj (2018) the authors studied the helpfulness of the AHP method with respect to users preferences in relation to the CR of the pairwsise comparison table he provided concluding that AHP was helpful even for users with a high level of inconsistency. We think that the CR threshold has to be suggested by the analyst depending on the complexity of the problem. For example, if a big number of criteria, alternatives or reference alternatives is taken into account, he can admit a CR threshold greater than the 10%. However, we would like to underline that the procedure we are proposing is independent on this point and a discussion on the way the CR threshold has to be fixed is out of the scope of the paper.

The same pairwise procedure described above is applied on the weights to calculate their importance $w_1, \ldots, w_n$. Once the local priorities and the weights have been obtained, the global priority $U(a)$ of the alternative $a \in A$ is obtained as

$$U(a) = \sum_{j=1}^{n} e_a^{(j)} w_j.$$

3. Reduction of pairwise comparisons in AHP: the new proposal

As observed in the previous section, the application of the AHP method involves $\binom{n}{2} + n\binom{|A|}{2}$ pairwise comparisons ($\binom{n}{2}$ pairwise comparisons between importance of criteria and $\binom{|A|}{2}$ pairwise comparisons between alternatives on each considered criterion). The amount of information asked to the DM can be quite huge even for small problems. For example, considering a problem of reasonable dimension composed of 5 criteria and 10 alternatives, the DM has to provide $\binom{10}{2}+10\binom{5}{2}$ pairwise comparisons, which are 145 pairwise comparisons. Consequently, it is quite difficult for the DM providing all these preference comparisons because of his limited capacity of processing information (Miller, 1956). If the criterion has an objective numerical evaluation (for example the maximum speed in a decision about cars), Corrente et al. (2016) proposed to apply AHP to prioritize a small set of reference levels of the considered criterion and to build the subjective measure of attractiveness of all the evaluations on the considered criterion by linear interpolation. Here we extend that approach to the case in which the considered criterion has not an objective numerical evaluation, so that the above mentioned interpolation cannot be applied and the basic AHP should be applied asking to the DM the pairwise comparisons of each alternative with all the others. Our proposal aims to reduce the cognitive effort involved in the application of the basic AHP and can be summarized in the following procedure where, for the sake of clarity we consider prioritization of pairwise comparisons of alternatives with respect to some criterion (of course, an analogous procedure can be applied to assign priorities to criteria in order to assess their weights):

Step 1 For each criterion g_j, we ask the DM to give a rating to alternatives in A on a common scale (for example 0-100). We shall denote by $r_j(a)$ the rating provided by the DM to the alternative a with respect to criterion g_j;

Step 2 For each criterion g_j, the DM, in accordance with the analyst, has to fix t_j reference evaluations $(\gamma_{j1}, \ldots, \gamma_{jt_j})$ on the considered common scale ordered from the lowest to the greatest one;

Step 3 Following Corrente et al. (2016), the DM is therefore asked to apply the AHP to the set composed of the reference evaluations defined on Step 2 obtaining the normalized evaluations $u(\gamma_{js})$, for all $j = 1, \ldots, n$ and for all $s = 1, \ldots, t_j$;

Step 4 The following checks are conducted and discussed with the DM:

- the consistency of the pairwise comparisons is computed through the consistency ratio (CR),
- the normalized evaluations $u(\gamma_{js})$, for all $j = 1, \ldots, n$ and for all $s = 1, \ldots, t_j$, are compared with the corresponding ratings $r_j(\gamma_{js})$ controlling that the monotonicity is satisfied, that is verifying that $r_j(\gamma_{js_1}) > r_j(\gamma_{js_2})$ iff $u(\gamma_{js_1}) \geq u(\gamma_{js_2})$.
- the DM has the possibility to modify the rating or the pairwise comparisons in order to get tolerable inconsistent pairwise comparisons (or consistent ones in the extreme case) that satisfy monotonicity and that are finally accepted by the DM.

Step 5 The rating of the evaluations provided by the DM which are not reference evaluations are obtained by interpolating the normalized evaluations got in the previous step. For each $r_j(a) \in [\gamma_{js}, \gamma_{js+1}]$, the following value is computed:

$$u(r_j(a)) = u(\gamma_{js}) + \frac{u(\gamma_{js+1}) - u(\gamma_{js})}{\gamma_{js+1} - \gamma_{js}}(r_j(a) - \gamma_{js}). \tag{1}$$

While in the original AHP method, the DM was asked to provide the pairwise comparison of all pairs of alternatives on all considered criteria, in this case he is asked to provide, at first, the rating of the alternatives on the considered criteria and applying the AHP on a small subsets of reference evaluations defined for each criterion. It is to observe that if the considered criterion has an objective numerical evaluation, the above procedure can be applied substituting this numerical evaluation to the direct rating supplied by the DM in Step 1. In fact, this is the procedure proposed in Corrente et al. (2016) where criteria have objective numerical evaluations.

Let us compare our methodology with the other approaches proposed to reduce the preference information requested to the DM in AHP.

The approaches based on clustering techniques ask the DM to compare set of alternatives with similar performances. These comparisons can be not so easy for the DM that is accustomed to compare single alternatives rather than groups of alternatives. Of course, this can have some consequence on the reliability of the obtained information and the resulting prioritization. Our approach does not suffer of this problem, because it continues to be based on comparison of performances related to single alternatives. Even more, the compared alternatives correspond to some reference points that are selected together with the DM and that are, consequently, very significant for him. In such a way, the corresponding pairwise comparison can be expected to be very reliable.

The approaches based on the partial completion of the pairwise comparison matrix present some similarity to our methodology because both of them consider a small set of pairwise comparisons. However, there are also some relevant differences. Indeed, our approach "corrects" the rating given to the performances of all the alternatives by means of the priorities obtained comparing only few reference evaluations meaningful for the DM. In the standard AHP approach, no rating is taken into consideration because it could not be reliable, and because no reference point is considered. Indeed, consideration of alternatives different from those ones on which the decision has to be taken can originate rank reversal (see e.g. Belton and Gear, 1983; Saaty and Vargas, 1984; Wang and Elhag, 2006) and in any case they increase the number of pairwise comparisons asked to the DM. Therefore, our methodology transforms in advantages for AHP two issues that are considered problematic in the theory supporting AHP:

- the direct rating, which is natural for the DM, but untrustworthy, can be amended through the interpolation with respect to the priorities assigned to the reference points;
- introducing performances not related to the alternatives of the decision problem at hand, permits to get more reliable evaluations since the attention of the DM is focused on reference points that are particularly meaningful for him.

Finally, with respect to the approaches based on some benchmarking, let us compare our approach with the most representative method of this group, which is the Best-Worst method (Rezaei, 2015). The Best-Worst method asks the DM to compare each performance with two reference performances being the best and the worst ones. This permits to ask much less pairwise comparison than the standard AHP method obtaining highly reliable priority values. We believe that our approach can improve these good properties of the Best-Worst method. Indeed, the Best-Worst method obtains its good results considering only the two extreme reference points leaving uncovered the central part of the spectrum of feasible performances. At the contrary, with our approach, the reference points are well distributed covering the whole interval of feasible performances. Moreover, the preference information requested by our approach is more parsimonious in most typical decision

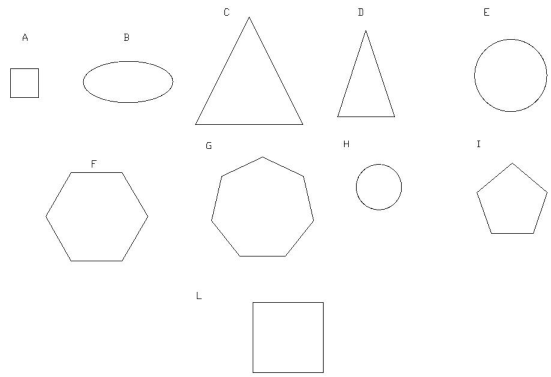

Figure 1: Geometrical figures which area have to be estimated.

problems than in the Best-Worst method. In fact, with n alternatives and r reference points, our method asks $n+\frac{r(r-1)}{2}$ pieces of preference information to the DM, that is

- n direct ratings for the alternatives of the problem at hand and
- $\frac{r(r-1)}{2}$ pairwise comparisons of the reference points.

Therefore we have to compare the $n+\frac{r(r-1)}{2}$ pieces of preference information required by our approach with the $2n-3$ pairwise comparisons asked by the Best-Worst method. Consequently, if the number of alternatives n is greater than $\frac{r(r-1)}{2}+3$, our approach requires less preference information than the Best-Worst method. This means that with 3 reference points, just with 7 alternatives our approach requires less preference information than the Best-Worst method. Considering 4 reference points, our approach is more advantageous from 10 alternatives and above. Table 1 presents the minimum number of alternatives given the number of reference points for which our approach requires less preference information than the Best-Worst method.

Table 1: Minimum number of alternatives given the number of reference points for which parsimonious AHP requires less preference information than the Best-Worst method

Number of reference points	Minimum number of alternatives
3	7
4	10
5	14
6	19
7	25
8	32
9	40

3.1. An illustrative example

In this section, we shall apply the new AHP procedure to the same problem taken into account in the experiment that we shall present in the next section. Let us suppose that the DM has to apply AHP to estimate the area of the ten geometrical figures shown in Figure 1.

To apply the AHP methodology, the DM is asked to perform $\binom{10}{2}=45$ pairwise comparisons between the areas of the considered figures. We shall present how our new proposal can be applied in this case, describing in detail the five steps presented in Section 2.3.

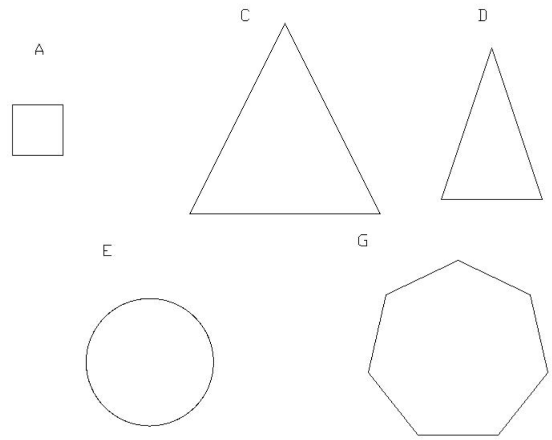

Figure 2: Reference geometric figures selected to correct the direct estimation.

Step 1 The DM is asked to provide a numerical evaluation of the area of the 9 figures B-L shown in Figure 1 knowing that the area of Figure A is equal to 1. The estimated values for the area of geometrical figures supplied by the DM are the following:

	A	**B**	**C**	**D**	**E**	**F**	**G**	**H**	**I**	**L**
Estimated Area : $\mathbf{r}(\cdot)$	1	2.75	8.5	3	4	6.5	5	2	3.75	4

Step 2 Let us suppose that the DM, in accordance to the analyst, chooses the five geometrical figures (A,C,D,E,G) shown in Figure 2 as reference alternatives. Using the previous notation and considering the estimate of the area of the five considered figures, we have $\gamma_1 = 1$, $\gamma_2 = 3$, $\gamma_3 = 4$, $\gamma_4 = 5$ and $\gamma_5 = 8.5$[2].

Step 3 We apply the AHP methodology to the five geometrical figures shown in Figure 2. For this reason, at first the geometrical figures have to be pairwise compared by the DM using the classical 1-9 AHP scale. The obtained pairwise comparison matrix is shown in Table 2.

Table 2: First pairwise comparison matrix

	A	**C**	**D**	**E**	**G**
A	1	$1/9$	$1/5$	$1/5$	$1/9$
C	9	1	6	4	$1/5$
D	5	$1/6$	1	$1/4$	$1/7$
E	5	$1/4$	4	1	$1/8$
G	9	5	7	8	1

Then, after observing that $CR = 0.1587$, the following priorities of the 5 geometrical figures are obtained:

[2]Of course this second step could be performed choosing a different number of geometrical figures as well as different geometrical figures.

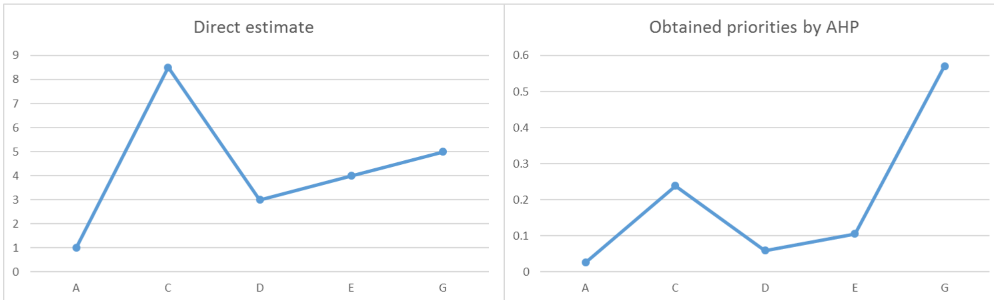

Figure 3: Direct estimate of the areas provided by the DM and their priorities obtained by applying the AHP method.

	A	**C**	**D**	**E**	**G**
Priority	0.0261	0.0590	0.1056	0.2389	0.5704

Step 4 At this point, the analyst discusses the collected information and the obtained results with the DM. First of all, the analyst observes that $CR > 0.1$, which means that the pairwise comparisons are not consistent enough. Moreover, the analyst shows that the priorities obtained through AHP on the pairwise comparisons table supplied by the DM are not concordant with estimates of the same DM. Indeed, looking at Table 3, one can observe that he estimated the areas of C and G in 8.5 and 5, respectively, while the priorities corresponding to the two figures are 0.0590 and 0.5704, being exactly in the opposite order of the estimated areas.

The analyst then looked at the provided pairwise comparisons and let the DM observe that even if he stated that the area of C is greater than the area of G, he introduced the value 5 in correspondence of the pair (G,C), meaning that G is 5 times greater than C.

Table 3: Estimated areas and priorities obtained using the first pairwise comparisons provided by the DM.

Figure	Estimated Area	Priorities obtained by AHP
A	1	0.0261
C	8.5	0.0590
D	3	0.1056
E	4	0.2389
G	5	0.5704

The previous observations made by the analyst became evident to the DM plotting the estimation of the areas he provided together with their priorities obtained by applying AHP to the pairwise comparison matrix as shown in Figure 3.

In consequence of these observations, the DM thought again to the provided information. He is still sure about the estimated areas but he decided to modify the two mentioned pairwise comparisons. The new pairwise comparison matrix provided by the DM is therefore shown in Table 4.

After applying the AHP to the new pairwise comparison matrix, one gets CR=0.0868 and the new priorities shown in Table 5.

The analyst observes that the pairwise comparisons are now consistent enough since the $CR < 0.1$ and that there is not any contradiction between the order of the figures w.r.t. the

Table 4: Second pairwise comparison matrix

	A	**C**	**D**	**E**	**G**
A	1	$^1/_9$	$^1/_5$	$^1/_5$	$^1/_9$
C	9	1	6	4	2
D	5	$^1/_6$	1	1	$^1/_7$
E	5	$^1/_4$	1	1	$^1/_8$
G	9	$^1/_2$	7	8	1

Table 5: Estimated areas and priorities obtained using the second pairwise comparison matrix provided by the DM.

Figure	**Estimated Area**	**Priorities obtained by AHP**
A	1	0.0279
C	8.5	0.4201
D	3	0.0797
E	4	0.0849
G	5	0.3871

estimated areas and the order of the same figures w.r.t. the priorities obtained by AHP.

As shown in Figure 4, the information provided by the DM is far from being linear. Once more, this underlines the importance of using the AHP to get the priorities of the considered areas.

Assigning a value of 1 to the area of figure A, using the priorities of AHP the areas of the five reference geometrical figures can therefore be estimated as follows:

	A	**C**	**D**	**E**	**G**
Area (u)	1	15.0364	2.8538	3.0404	13.8564
	$u(\gamma_1)$	$u(\gamma_5)$	$u(\gamma_2)$	$u(\gamma_3)$	$u(\gamma_4)$

Step 5 The area of all non-reference geometrical figures (B, F, H, I, L) can be now estimated by linear interpolation with respect to the values of the areas obtained in the previous step. For example, since the estimated area of F was 6.5 and it belongs to the interval of reference evaluations $[5, 8.5]$ (being the estimated areas for reference geometrical figures G and C), we get that the area of F is equal to

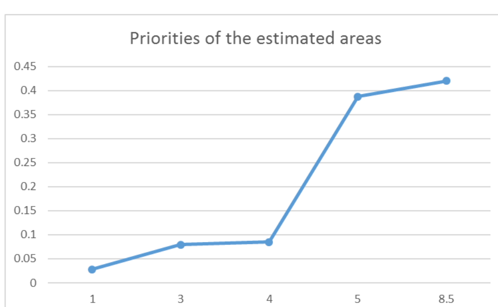

Figure 4: Priorities of the reference areas provided by the DM and obtained by AHP after the correction in the preference information.

$$
\begin{aligned}
Area(r(F)) &= u(\gamma_4) + \frac{u(\gamma_5) - u(\gamma_4)}{\gamma_5 - \gamma_4}(r(F) - \gamma_4) = \\
&= 13.8564 + \frac{15.0364 - 13.8564}{8.5 - 5}(6.5 - 5) = 14.3621.
\end{aligned}
$$

Therefore, the areas of the ten geometrical figures obtained by the application of the newly proposed AHP procedure are the following:

	A	**B**	**C**	**D**	**E**
Final Area	1	2.6221	15.0364	2.8538	3.0404
	F	**G**	**H**	**I**	**L**
	14.3521	13.8564	1.9269	2.9937	3.0404

4. An experiment on the reliability of the proposed approach

In this section, we shall present the results of the performed experiment. Let us point out that analogous experiments, but with other aims, have been performed and described in Ishizaka and Nguyen (2013) and Meesariganda and Ishizaka (2017).

We selected a group of 98 undergraduate students from the Department of Economics and Business of the University of Catania (Italy). This group was split in two subgroups of 46 and 52 students. Both groups were presented with a sheet on which ten geometrical figures (triangles, rectangles, circles, etc.) of different area were pictured (see Figure 1). Anyway, the components of the two groups were asked to fulfill two different assignments:

1. Students belonging to the first group were asked to:
   - give an estimate of the area of the geometric figures B-L, knowing that the area of the geometric figure A was equal to 1,
   - fill in a pairwise comparison matrix related to the five geometric figures shown in Figure 2, giving an evaluation of the ratio between their area using the usual 9-point qualitative scale used by AHP (in this case “equally large”, “weakly larger”, “strongly larger”, “very strongly larger”, “absolutely larger” corresponding to 1, 3, 5, 7 and 9, respectively, with 2, 4, 6, 8 being related to intermediate evaluations).
2. Students belonging to the second group were asked to fill in the pairwise comparison matrix relative to all ten figures as in the original AHP method. This assignment involved $\binom{10}{2} = 45$ pairwise comparisons of the different areas.

For the students of the first group, we computed the area of the ten geometric figures according to the methodology described in section 3 and presented as an example in section 3.1. For the students of the second group, we computed their area using the traditional AHP approach, that is, by computing the normalized right principal eigenvectors of the comparison matrix. Finally, denoting by $Area^{real}$ and $Area^{estimated}$ the vectors containing the real area and the estimated area of the ten geometric figures shown in Figure 1, we computed the mean squared error (MSE) of these two vectors, that is,

$$MSE = \frac{1}{10}\sum_{i=1}^{10}\left(Area_i^{real} - Area_i^{estimated}\right)^2.$$

With respect to AHP we computed the MSE and its standard deviation for:

- all questionnaires (AHP_0),
- all questionnaires for which the corresponding CR was lower than 0.1 (AHP_1).

With respect to the newly proposed AHP procedure, to show the goodness of the methodology as well as to underline the necessity to consider both reliable pairwise comparisons as well as correct estimates, we computed the MSE and its standard deviation for:

- all questionnaires (NewAHP_0),
- all questionnaires presenting a CR lower than 0.1 (NewAHP_1),
- all questionnaires for which the priorities obtained for the reference evaluations were concordant with the estimated area (NewAHP_2),
- all questionnaires presenting a CR lower than 0.1 and for which the priorities obtained for the reference evaluations were concordant with the estimated area (NewAHP_3).

All these data are shown in Table 6.

Table 6: Results of the experiment for the classical AHP method and for the new proposal.

	AHP_0	**AHP_1**	**NewAHP_0**
Average MSE	12.2819	11.716	11.33
StD MSE	7.6621	6.8388	10.47

	NewAHP_1	**NewAHP_2**	**NewAHP_3**
Average MSE	7.5538	10.48	7.15
StD MSE	4.4705	11.264	3.469

The data in Table 6 show that the results obtained for AHP_1 (the classical AHP procedure but considering only those questionnaires presenting $CR > 0.1$), are more or less similar in average to NewAHP_0 (the results obtained from all questionnaires with the new methodology), even if the results obtained with the classical procedure AHP_1 are more stable as shown by the smaller standard deviation of AHP_1. However, if we clean the data related to the newly proposed methodology by removing those pairwise comparisons matrices with $CR > 0.1$, then the performance becomes really better (both in average and in stability), since the MSE is equal to 7.5538 and the standard deviation is 4.4705. These data still improve if we consider questionnaires for which the pairwise comparisons are reliable enough (presenting therefore $CR \leq 0.1$) and for which there is not any inversion between the order of the reference alternatives w.r.t. their estimated area and the order of the same reference figures w.r.t. the obtained priorities. Indeed, it seems that NewAHP_3 can be considered as the best one since the corresponding MSE is the lowest, being 7.15. Let us observe also that instead there is not a great improvement removing the questionnaires for which the priorities obtained for the reference evaluations were not concordant with the estimated areas (NewAHP_2). In this case, not only the MSE has only a negligible reduction (10.48 vs 11.33), but, even worse, the standard deviation is increased (11.264 vs 10.47), which points out that the obtained results are basically more unstable. This suggests that both the consistency ratio and the inversions of the direct estimates and the priorities obtained from the AHP approach on the reference values have to be taken into consideration in the discussion with the DM.

5. An application of the proposed method to an MCDA problem

In this section we shall describe how the proposed methodology can be applied to an MCDA problem. Let us assume that the dean of an high school has to decide which worthy student, among 20 candidates, deserves a scholarship. Students are evaluated on three main subjects that are Mathematics (M), Physics (P) and Literature (L). For handling such a problem, we suggested to apply the new proposal. In the following, we shall detail all the steps of the considered procedure introduced in section 3:

Step 1 The dean, helped by the professors of the considered subjects, evaluated students on a [18,30] scale as shown in Table 7, on the left.

Table 7: Students' evaluations (left), ratings (center) and comprehensive evaluation (right)

	M	**P**	**L**		**M**	**P**	**L**		$U(\cdot)$
$\mathbf{S_1}$	28	29	19	$\mathbf{S_1}$	0.4388	0.5443	0.0527	$\mathbf{S_1}$	0.4265 (5th)
$\mathbf{S_2}$	29	26	19	$\mathbf{S_2}$	0.5508	0.1990	0.0527	$\mathbf{S_2}$	0.4287 (4th)
$\mathbf{S_3}$	21	25	30	$\mathbf{S_3}$	0.0692	0.1736	0.6815	$\mathbf{S_3}$	0.1479 (13rd)
$\mathbf{S_4}$	30	20	30	$\mathbf{S_4}$	0.6627	0.0708	0.6815	$\mathbf{S_4}$	0.5356 (1st)
$\mathbf{S_5}$	30	24	28	$\mathbf{S_5}$	0.6627	0.1482	0.4365	$\mathbf{S_5}$	0.5301 (2nd)
$\mathbf{S_6}$	19	23	29	$\mathbf{S_6}$	0.0532	0.1227	0.5590	$\mathbf{S_6}$	0.1146 (17th)
$\mathbf{S_7}$	28	30	26	$\mathbf{S_7}$	0.4388	0.6594	0.1916	$\mathbf{S_7}$	0.4642 (3rd)
$\mathbf{S_8}$	18	29	30	$\mathbf{S_8}$	0.0451	0.5443	0.6815	$\mathbf{S_8}$	0.2119 (11th)
$\mathbf{S_9}$	26	27	27	$\mathbf{S_9}$	0.2148	0.3141	0.3141	$\mathbf{S_9}$	0.2455 (9th)
$\mathbf{S_{10}}$	23	26	20	$\mathbf{S_{10}}$	0.1117	0.1990	0.0635	$\mathbf{S_{10}}$	0.1263 (16th)
$\mathbf{S_{11}}$	27	18	21	$\mathbf{S_{11}}$	0.3268	0.0442	0.0742	$\mathbf{S_{11}}$	0.2422 (10th)
$\mathbf{S_{12}}$	18	19	28	$\mathbf{S_{12}}$	0.0451	0.0575	0.4365	$\mathbf{S_{12}}$	0.0836 (19th)
$\mathbf{S_{13}}$	27	22	30	$\mathbf{S_{13}}$	0.3268	0.0973	0.6815	$\mathbf{S_{13}}$	0.3093 (7th)
$\mathbf{S_{14}}$	18	23	22	$\mathbf{S_{14}}$	0.0451	0.1227	0.0849	$\mathbf{S_{14}}$	0.0657 (20th)
$\mathbf{S_{15}}$	27	28	20	$\mathbf{S_{15}}$	0.3268	0.4292	0.0635	$\mathbf{S_{15}}$	0.3250 (6th)
$\mathbf{S_{16}}$	24	23	26	$\mathbf{S_{16}}$	0.1461	0.1227	0.1916	$\mathbf{S_{16}}$	0.1451 (14th)
$\mathbf{S_{17}}$	27	27	21	$\mathbf{S_{17}}$	0.3268	0.3141	0.0742	$\mathbf{S_{17}}$	0.3010 (8th)
$\mathbf{S_{18}}$	26	26	20	$\mathbf{S_{18}}$	0.2148	0.1990	0.0635	$\mathbf{S_{18}}$	0.1976 (12th)
$\mathbf{S_{19}}$	19	24	30	$\mathbf{S_{19}}$	0.0532	0.1482	0.6815	$\mathbf{S_{19}}$	0.1313 (15th)
$\mathbf{S_{20}}$	22	25	20	$\mathbf{S_{20}}$	0.0773	0.1736	0.0635	$\mathbf{S_{20}}$	0.0970 (18th)

Step 2 The dean, in accordance with the analyst, chooses to fix 4 reference levels for all criteria that are $\gamma_1 = 18$, $\gamma_2 = 22$, $\gamma_3 = 26$ and $\gamma_4 = 30$;

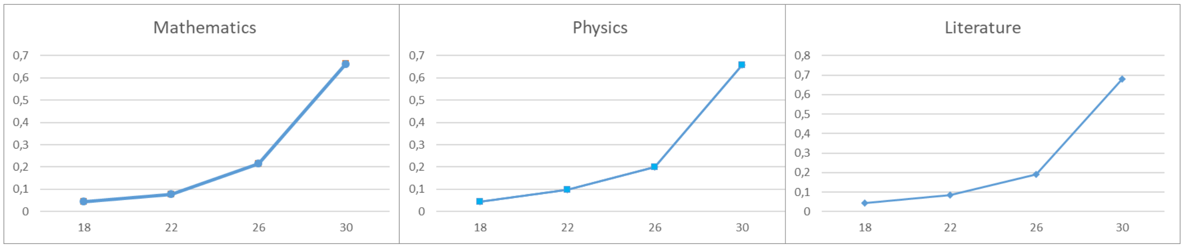

Figure 5: Priorities of the reference levels for the considered subjects.

Step 3 The dean provides 4 pairwise comparison tables. In the first one (Table 8(a)) he expresses his preferences on the three considered subjects. In the other three (Tables 8(b)-8(d)), he compares the reference levels for each subject.

Table 8: Pairwise comparison tables on criteria as well as on the considered subjects

(a) Subjects (CR=0.0462)

	M	**P**	**L**
M	1	4	6
P	$1/4$	1	3
L	$1/6$	$1/3$	1

(b) Mathematics (CR=0.0611)

	18	**22**	**26**	**30**
18	1	$1/2$	$1/7$	$1/9$
22	2	1	$1/3$	$1/8$
26	7	3	1	$1/5$
30	9	8	5	1

(c) Physics (CR=0.0511)

	18	**22**	**26**	**30**
18	1	$1/3$	$1/6$	$1/9$
22	3	1	$1/2$	$1/9$
26	6	2	1	$1/4$
30	9	9	4	1

(d) Literature (CR=0.0872)

	18	**22**	**26**	**30**
18	1	$1/3$	$1/6$	$1/9$
22	3	1	$1/3$	$1/8$
26	6	3	1	$1/6$
30	9	8	6	1

The analyst is able to obtain the priorities of the criteria and the priorities of the reference levels for all considered criteria.

$$\begin{array}{lcccc}
\textit{Subjects} & \mathbf{M} & \mathbf{P} & \mathbf{L} & \\
\textbf{Priority} & 0.691 & 0.2176 & 0.0914 & \\
 & & & & \\
\textit{Mathematics} & \gamma_1 = \mathbf{18} & \gamma_2 = \mathbf{22} & \gamma_3 = \mathbf{26} & \gamma_4 = \mathbf{30} \\
\textbf{Priority} & 0.0451 & 0.0773 & 0.2148 & 0.6627 \\
 & & & & \\
\textit{Physics} & \gamma_1 = \mathbf{18} & \gamma_2 = \mathbf{22} & \gamma_3 = \mathbf{26} & \gamma_4 = \mathbf{30} \\
\textbf{Priority} & 0.0442 & 0.0973 & 0.1990 & 0.6594 \\
 & & & & \\
\textit{Literature} & \gamma_1 = \mathbf{18} & \gamma_2 = \mathbf{22} & \gamma_3 = \mathbf{26} & \gamma_4 = \mathbf{30} \\
\textbf{Priority} & 0.042 & 0.0849 & 0.1916 & 0.6815
\end{array} \tag{2}$$

As shown in Figure 5, the preference provided by the DM on the reference evaluations is not linear. This underlines, once again, the usefulness of applying AHP to find the priorities of these levels on the considered subjects.

Step 4 The analyst computes the CR for all tables observing that it is always lower than 10%. At the same time, he observes that the monotonicity of the priorities with respect to the reference evaluations is respected. Consequently, there is not the necessity to review the information provided by the dean.

Step 5 The rating of the students on the three subjects that are not reference evaluations are obtained by linear interpolation and shown in Table 7. For example, considering that the evaluation of $\mathbf{S_{16}}$ on M is 24, this value belongs to the interval of reference evaluations $[22, 26]$ and that the priorities assigned to these reference evaluations are 0.0773 and 0.2148, respectively (see (2)), we get

$$\begin{aligned} u_M(24) &= u_M(22) + \frac{u_M(26) - u_M(22)}{26 - 22} \cdot (24 - 22) = \\ &= 0.0773 + \frac{0.2148 - 0.0773}{4} \cdot 2 = 0.1461. \end{aligned}$$

For each student, multiplying his/her ratings (Table 7, middle) for the priorities of the three subjects (2) we obtain the final evaluation shown again in Table 7, on the right. It is evident from the last column in the mentioned table that, on one hand, $\mathbf{S_3}$ and $\mathbf{S_4}$ are the best students, while, on the other hand, $\mathbf{S_{12}}$ and $\mathbf{S_{14}}$ are the worst ones.

Once more we would like to underline the usefulness of using our approach compared with the classical AHP. Indeed, if the dean had decided to apply the AHP method, he had been asked to provide $3 \cdot \binom{20}{2} + \binom{3}{2} = 573$ pairwise comparisons ($\binom{20}{2}$ for each criterion and $\binom{3}{2}$ for the criteria comparison), while, in our approach, he was asked to provide $3 \cdot \binom{4}{2} + \binom{3}{2} = 21$ pairwise comparisons only. Assuming that the dean needed 30 secs for each pairwise comparison and 20 secs for each criterion evaluation, to apply AHP he had needed 17,190 secs (approximately 4.775 hours) while in applying the proposed method he has taken 1830 secs, that is, 30 minutes and half, implying a reduction of the 89.35% of the time.

6. Conclusions

In this paper, we proposed a novel prioritization method using pairwise comparison matrices. The aim of the proposed methodology is to ask a parsimonious preference information to the Decision Maker (DM), maintaining the soundness of the obtained priorities. More precisely, taking into account that comparative judgments are more reliable than absolute judgments but only for a limited number of objects to be prioritized, our methodology proposes to correct absolute judgments by taking into account priorities related to a small number of reference objects. It permits to use the Analytic Hierarchy Process (AHP), being one of the most adopted Multiple Criteria Decision Aiding (MCDA) methods, even if the considered decision problem presents a high number of alternatives or criteria. Observe that even if we focused on AHP, the relevance of our proposal goes beyond this specific method and it is related to all the MCDA domain. Indeed, the idea of correcting absolute judgments through prioritization obtained from comparative judgments of a small number of reference objects, can be applied to other methods using pairwise comparison matrices such as MACBETH, but also in applying any other MCDA method that can benefit from prioritizations obtained from comparative judgments. This can be the case, for example, of the elicitation of weights to be applied in other methods through AHP, as successfully proposed for PROMETHEE methods (Brans and Vincke, 1985) in Macharis et al. (2004). The new proposal is composed of five different steps: (i) direct evaluation of the objects at hand; (ii) definition of some reference evaluations in accordance with the analyst; (iii) pairwise comparison of the reference objects by using the AHP method; (iv) control and discussion with the DM of the consistency of the supplied pairwise comparisons and the monotonicity of the prioritization with respect to the rating of the reference points; the DM can modify the rating or pairwise comparisons of reference points to get consistency and monotonicity until the priorities of the reference points are satisfying for him; (v) interpolation of the values obtained by AHP in the previous step in order to get the normalized evaluations of all objects. Our new parsimonious AHP has two main advantages with respect to the standard AHP: 1) it avoids the distortion of comparing more relevant objects (reference points) with less relevant objects; 2) it avoids rank reversal problems. The new proposal has been tested on a group of undergraduate students belonging to the University of Catania and the experiment confirmed the goodness of the proposed methodology. We believe that our proposal of a parsimonious version of AHP can be considered a relevant contribution for the basic theory and application of the method. Indeed AHP is one of the most adopted MCDA methods. Anyway its adoption is limited to problems with small number of alternatives or criteria due to the impossibility to ask to the DM many pairwise comparative judgments. Our approach gives an answer to this bottleneck and therefore substantially strengthens the potential of AHP. In this sense, the many real life decision problems with many alternatives and criteria can now be effectively handled with AHP. Among these decision problems, we retain particularly interesting the following: assignments of rotation schedule (Michalos et al., 2011, 2010), designing of assembly line alternatives (Michalos et al., 2012) and assembly line reconfiguration (Michalos et al., 2016). In all these problems, several hundreds of different alternatives are evaluated with respect to conflicting criteria to choose the best alternative. Generally, the evaluations of the alternatives on the criteria at hand are aggregated by a weighted sum and, consequently, to make comparable the evaluations, different normalization techniques are applied. We think that the methodologies and techniques with which these problems are handled could largely benefit from the integration with AHP that would avoid the use of an aseptic normalization technique in favor of a more advanced decision model taking adequately into consideration the preferences of the DM. However the application of the standard AHP method would be impossible in these cases since the DM cannot provide thousands of pairwise comparisons, while it is definitely reasonable with our methodology where the DM is asked only to compare the reference evaluations that are the most representative ones in his opinion and that are very limited in number. Another class of significant decision problems in which the contribution of our methodology could be relevant is constituted by the evaluation procedures based on composite indices such as those ones underlying ranking of Universities, Countries, Cities and so on (for a recent survey see Greco et al., 2018). Also in these cases, it is not possible to pairwise compare hundreds and even thousands of alternatives with respect to several criteria and, nevertheless it seems appealing to take advantage of the reliability of comparative judgments ensured by AHP. Finally, a class of decision problems where there is no possibility to apply the standard AHP and that, instead, could be adequately approached with our methodology is given by the repetitive standard evaluation procedures such as credit scoring analysis (see e.g. Mareschal and Brans, 1991; Zopounidis, 1987). These procedures are able to produce an assessment for any possible alternative that could be submitted to evaluation. In this case, beyond all the considerations about the number of pairwise comparisons to be asked to the DM, it is not possible even imagine pairwise comparisons between alternatives that are not known a-priori, as it is the case of customers that have not yet applied for a credit. In such a situation, the only possibility to collect and to process comparative judgments is to ask the DM, for example the bank manager, pairwise comparisons of reference evaluations. Again, using linear interpolation for priorities of reference points we can assign a priority to the evaluation of any alternative on all considered criteria. In view of all these considerations, we expect the adoption of our methodology in diversified real world decision problems and we are confident in the good results that can be obtained as it was the case for the application described by Abastante et al. (2018). We expect also that the approach proposed in this paper could stimulate further methodological advances on the direction of a more parsimonious request of preference information in AHP. More in detail, we believe that our approach could be extended taking into account group decisions (see e.g. Dyer and Forman, 1992; Escobar et al., 2015; Gass and Rapcsák, 1998), fuzzy evaluations (see e.g. Gu and Zhu, 2006) and dynamic contexts (Benítez et al., 2012).

Acknowledgments

Salvatore Corrente and Salvatore Greco gratefully acknowledge the funding by the research project "Data analytics for entrepreneurial ecosystems, sustainable development and wellbeing indices"' of the Department of Economics and Business of the University of Catania. Salvatore Greco has also benefited of the fund "Chance" of the University of Catania

References

Abastante, F., Corrente, S., Greco, S., Ishizaka, A., Lami, I., 2018. Choice architecture for architecture choices: Evaluating social housing initiatives putting together a parsimonious AHP methodology and the Choquet integral. Land Use Policy 78, 748–762.

Barberis, N., Xiong, W., 2009. What drives the disposition effect? An analysis of a long-standing preference-based explanation. The Journal of Finance 64, 751–784.

Belton, V., Gear, T., 1983. On a short-coming of Saaty's method of analytic hierarchies. Omega 11, 228–230.

Benítez, J., Delgado-Galván, X., Izquierdo, J., Pérez-García, R., 2012. An approach to AHP decision in a dynamic context. Decision Support Systems 53, 499–506.

Benítez, J., Izquierdo, J., Pérez-García, R., Ramos-Martínez, E., 2014. A simple formula to find the closest consistent matrix to a reciprocal matrix. Applied Mathematical Modelling 38, 3968–3974.

Blumenthal, A., 1977. The process of cognition. Prentice Hall/Pearson Education.

Bozóki, S., Fülöp, J., Rónyai, L., 2010. On optimal completion of incomplete pairwise comparison matrices. Mathematical and Computer Modelling 52, 318–333.

Brans, J., Vincke, P., 1985. A preference ranking organisation method: The PROMETHEE method for MCDM. Management Science 31(6), 647–656.

Camerer, C., Babcock, L., Loewenstein, G., Thaler, R., 1997. Labor supply of New York City cabdrivers: One day at a time. The Quarterly Journal of Economics 112, 407–441.

Carmoner, F., Kara, A., Zanakis, S., 1997. A Monte Carlo investigation of incomplete pairwise comparison matrices in AHP. European journal of operational research 102, 538–553.

Chen, K., Kou, G., Tarn, J., Song, Y., 2015. Bridging the gap between missing and inconsistent values in eliciting preference from pairwise comparison matrices. Annals of Operations Research 235, 155–175.

Corrente, S., Greco, S., Ishizaka, A., 2016. Combining Analytical Hierarchy Process and Choquet integral within Non Additive Robust Ordinal Regression. Omega 61, 2–18.

Bana e Costa, C., Vansnick, J.C., 1994. MACBETH-An interactive path towards the construction of cardinal value functions. International Transactions in Operational Research 1, 489–500.

Crawford, G., Williams, C., 1985. A note on the analysis of subjective judgment matrices. Journal of mathematical psychology 29, 387–405.

Csató, L., Rónyai, L., 2016. Incomplete pairwise comparison matrices and weighting methods. Fundamenta Informaticae 144, 309–320.

Duleba, S., Moslem, S., 2019. Examining Pareto optimality in analytic hierarchy process on real Data: An application in public transport service development. Expert Systems with Applications 116, 21–30.

Dyer, R., Forman, E., 1992. Group decision support with the analytic hierarchy process. Decision support systems 8, 99–124.

Escobar, M., Aguarón, J., Moreno-Jiménez, J., 2015. Some extensions of the precise consistency consensus matrix. Decision Support Systems 74, 67–77.

Fedrizzi, M., Giove, S., 2007. Incomplete pairwise comparison and consistency optimization. European Journal of Operational Research 183, 303–313.

Gass, S., Rapcsák, T., 1998. A note on synthesizing group decisions. Decision Support Systems 22, 59–63.

Genesove, D., Mayer, C., 2001. Loss aversion and seller behavior: Evidence from the housing market. The Quarterly Journal of Economics 116, 1233–1260.

Gomez-Ruiz, J., Karanik, M., Peláez, J., 2010. Estimation of missing judgments in AHP pairwise matrices using a neural network-based model. Applied Mathematics and Computation 216, 2959–2975.

Greco, S., Ehrgott, M., Figueira, J., 2016. Multiple Criteria Decision Analysis: State of the Art Surveys. Springer, Berlin.

Greco, S., Ishizaka, A., Tasiou, M., Torrisi, G., 2018. On the methodological framework of composite indices: A review of the issues of weighting, aggregation, and robustness. Social Indicators Research DOI: 10.1007/s11205-017-1832-9.

Greco, S., Matarazzo, B., Słowiński, R., 2001. Rough sets theory for multicriteria decision analysis. European Journal of Operational Research 129, 1–47.

Gu, X., Zhu, Q., 2006. Fuzzy multi-attribute decision-making method based on eigenvector of fuzzy attribute evaluation space. Decision Support Systems 41, 400–410.

Harker, P., 1987. Shortening the comparison process in the AHP. Mathematical Modelling 8, 139–141.

Harker, P., Vargas, L., 1990. Reply to “remarks on the analytic hierarchy process” by J.S. Dyer. Management Science 36, 269–273.

Ishizaka, A., 2012. Clusters and pivots for evaluating a large numberof alternatives in ahp. Pesquisa Operacional 32, 87–102.

Ishizaka, A., Labib, A., 2011. Review of the main developments in the analytic hierarchy process. Expert systems with applications 38, 14336–14345.

Ishizaka, A., Nemery, P., 2013. Multicriteria Decision Aid: Methods and software. Wiley, Chichester.

Ishizaka, A., Nguyen, N., 2013. Calibrated fuzzy AHP for current bank account selection. Expert Systems with Applications 40, 3775–3783.

Ishizaka, A., Pearman, C., Nemery, P., 2012. AHPSort: an AHP-based method for sorting problems. International Journal of Production Research 50, 4767–4784.

Ishizaka, A., Siraj, S., 2018. Are multi-criteria decision-making tools useful? An experimental comparative study of three methods. European Journal of Operational Research 264, 462–471.

Kahneman, D., 1979. Prospect theory: An analysis of decisions under risk. Econometrica 47, 263–291.

Keeney, R., Raiffa, H., 1976. Decisions with multiple objectives: Preferences and value tradeoffs. J. Wiley, New York.

Krejčí, J., Stoklasa, J., 2018. Aggregation in the analytic hierarchy process: Why weighted geometric mean should be used instead of weighted arithmetic mean. Expert Systems with Applications 114, 97–106.

Leskinen, P., Kangas, J., 2005. Rank reversals in multi-criteria decision analysis with statistical modelling of ratio-scale pairwise comparisons. Journal of the Operational Research Society 56, 855–861.

Lundy, M., Siraj, S., Greco, S., 2017. The mathematical equivalence of the “spanning tree” and row geometric mean preference vectors and its implications for preference analysis. European Journal of Operational Research 257, 197–208.

Macharis, C., Springael, J., De Brucker, K., Verbeke, A., 2004. PROMETHEE and AHP: The design of operational synergies in multicriteria analysis.: Strengthening PROMETHEE with ideas of AHP. European Journal of Operational Research 153, 307–317.

Maleki, H., Zahir, S., 2013. A comprehensive literature review of the rank reversal phenomenon in the analytic hierarchy process. Journal of Multi-Criteria Decision Analysis 20, 141–155.

Mareschal, B., Brans, J.P., 1991. Bankadviser: An industrial evaluation system. European Journal of Operational Research 54, 318–324.

Meesariganda, B., Ishizaka, A., 2017. Mapping verbal AHP scale to numerical scale for cloud computing strategy selection. Applied Soft Computing 53, 111–118.

Michalos, G., Makris, S., Mourtzis, D., 2011. A web based tool for dynamic job rotation scheduling using multiple criteria. CIRP Annals-Manufacturing Technology 60, 453–456.

Michalos, G., Makris, S., Mourtzis, D., 2012. An intelligent search algorithm-based method to derive assembly line design alternatives. International Journal of Computer Integrated Manufacturing 25, 211–229.

Michalos, G., Makris, S., Rentzos, L., Chryssolouris, G., 2010. Dynamic job rotation for workload balancing in human based assembly systems. CIRP Journal of Manufacturing Science and Technology 2, 153–160.

Michalos, G., Sipsas, P., Makris, S., Chryssolouris, G., 2016. Decision making logic for flexible assembly lines reconfiguration. Robotics and Computer-Integrated Manufacturing 37, 233–250.

Miller, G., 1956. The magical number seven, plus or minus two: some limits on our capacity for processing information. Psychological review 63, 81–97.

Millet, I., 1997. The effectiveness of alternative preference elicitation methods in the Analytic Hierarchy Process. Journal of Multi-Criteria Decision Analysis 6, 41–51.

Nazari, S., Fallah, M., Kazemipoor, H., Salehipour, A., 2018. A fuzzy inference-fuzzy analytic hierarchy process-based clinical decision support system for diagnosis of heart diseases. Expert Systems with Applications 95, 261–271.

Por, H., Budescu, D., 2017. Eliciting Subjective Probabilities through Pair-wise Comparisons. Journal of Behavioral Decision Making 30, 181–196.

Raharjo, H., Endah, D., 2006. Evaluating relationship of consistency ratio and number of alternatives on rank reversal in the AHP. Quality Engineering 18, 39–46.

Rezaei, J., 2015. Best-worst multi-criteria decision-making method. Omega 53, 49–57.

Roy, B., 1996. Multicriteria Methodology for Decision Aiding. Nonconvex Optimization and its Applications, Kluwer Academic Publishers, Dordrecht.

Roy, B., 2010. Two conceptions of decision aiding. International Journal of Multicriteria Decision Making 1, 74–79.

Saaty, T., 1977. A scaling method for priorities in hierarchical structures. Journal of Mathematical Psychology 15, 234–281.

Saaty, T., 1980. The Analytic Hierarchy Process. New York, McGraw-Hill.

Saaty, T., 1990a. How to make a decision: the Analytic Hierarchy Process. European Journal of Operational Research 48, 9–26.

Saaty, T., Ozdemir, M., 2003. Why the magic number seven plus or minus two. Mathematical and Computer Modelling 38, 233–244.

Saaty, T., Vargas, L., 1984. The legitimacy of rank reversal. Omega 12, 513–516.

Saaty, T.L., 1990b. An exposition of the AHP in reply to the paper "remarks on the analytic hierarchy process". Management science 36, 259–268.

Siraj, S., Mikhailov, L., Keane, J., 2012. Enumerating all spanning trees for pairwise comparisons. Computers & Operations Research 39, 191–199.

Stam, A., Silva, A., 2003. On multiplicative priority rating methods for the AHP. European Journal of Operational Research 145, 92–108.

Tsyganok, V., 2010. Investigation of the aggregation effectiveness of expert estimates obtained by the pairwise comparison method. Mathematical and Computer Modelling 52, 538–544.

Tversky, A., Kahneman, D., 1992. Advances in prospect theory: Cumulative representation of uncertainty. Journal of Risk and uncertainty 5, 297–323.

Wang, Y.M., Elhag, T., 2006. An approach to avoiding rank reversal in AHP. Decision Support Systems 42, 1474–1480.

Zopounidis, C., 1987. A multicriteria decision-making methodology for the evaluation of the risk of failure and an application. Found. Control Eng. 12, 45–64.