Università di Pisa

Dipartimento di Informatica

Technical Report : TR-99-06

Evaluation of DEA validity as a MCDA/M tool: some problems and issues

Fernanda Farinaccio, Anna Ostanello

March 29, 1999

ADDR: Corso Italia 40,56125 Pisa,Italy. TEL: +39-50-887111. FAX: +39-50-887226

Evaluation of DEA validity as a MCDA/M tool: some problems and issues

Fernanda Farinaccio,[1] *Anna Ostanello*[2]

1 . Introduction

Data Envelopment Analisys (DEA) and Multicriteria Decision Aid/Making (MCDA/M) are two important research areas in Operational Research and Management Science, which have been receiving considerable attention in OR/MS literature over the past two decades (see [13], for a large DEA bibliography, [12], [14], for a survey of MCDA/M).

DEA is a set of models, based on Linear Programming (LP), which is used to evaluate the *relative* efficiency of *Decision Making Units* (DMUs) performing similar *transformation processes* of several *inputs* , into several *outputs* . The principal aim is to provide a "measure of efficiency/inefficiency", so as to distinguish between *efficient* and *not efficient* DMUs. The measure of inefficiency obtained by DEA gives an assessment of the improvement needed to reach efficiency (i.e. good relative performance).

There is an analogy between the problem tackled by DEA and the *ranking problem* in Multicriteria Analysis ([10]). The DEA and the MCDA/M ranking problems, indeed, are apparently similar, by interpreting the DMUs as a set of candidates to evaluate, and the input/ouput factors as criteria to be minimised/maximised. This analogy has led some authors to begin to studying the interactions between DEA and other approaches of MCDA/M (see [2], [5], [15]). Some questions regarding the possibility of considering DEA as an MCDA/M tool have been raised and criticisms have also been made [3]. Some of the given answers, however, are "problematic" more than "resolutive", and important conceptual and formal problems may be added to the initial ones. The multiplicity of "criteria" (i.e. input and output factors) involved in DEA models does not necessarily mean that such kind of representation might answer some questions which concern the preferences on a given candidate set, as other MCDA/M tools can do validly (both conceptually and formally) (see [6]).

A few methodological considerations and research proposals regarding these problematical aspects shall be presented in the paper. We introduce a multi-judge MCDA/M model, with two decision levels (strategic and operational), which allows us to show how

[1] ADDR: Dipartimento di Informatica - Corso Italia, 40 - 56125 Pisa (Italy) - E_MAIL: farinacc@di.unipi.it
[2] ADDR: DSPEA - Politecnico di Torino, Corso Duca degli Abruzzi, 24 - 10129 Torino -(Italy)

DEA models can be used to answer the evaluation problems at the operational level. The importance of clearly defining the *Decision Maker* figure (DM) is underlined, so as to distinguish situations in which an "individual" or a group rather than a multiplicity of judges is involved in the evaluation and ranking of a candidate set.

2. The DEA evaluation approach

The problem formulation of a DEA model can be viewed as a MCDA/M problem formulation, if one considers the following assumptions.

- Let T_0 be a set of observed DMUs, considered as "real" candidates; then T_0 is **the candidate set A**:

$$A=\{a_1,...a_j,...,a_n\}=\{DMU_1,...,\ DMU_j,...,\ DMU_n\}=T_0\ .$$

- The input/output factors, X_i , $i=1,...m$, Y_r, $r=1,...,s$, are considered as **the criteria set** $G=[g_1,\mathrm{K}\ ,g_p]$, such that $\forall a_j \in A$:

- $$\left.\begin{array}{lcll} g_i(a_j) & = & X_{ij}, & i=1,\mathrm{K}\ ,m \\ g_{m+r}(a_j) & = & Y_{rj}, & r=1,\mathrm{K}\ ,s \end{array}\right\}.$$

The criteria scales are non increasing with preference, for the input factors, and non decreasing for the output ones.

- The set T , defined as a convexification of the discrete set T_0 , in the criteria space, is a population of "potential" DMUs, obtained by applying a set of production postulates.

- *Relative evaluation* of T_0 is the **problem** tackled by DEA; it may be assumed as equivalent to the *ranking problem* ([10]).

The *efficiency notion* assumed in DEA is intended in terms of Pareto efficiency ([4]). Thus, the *efficient* DMUs are elements of the non dominated subset of T , in the criteria space, with respect to the Pareto preference structure defined as follows.

Definition (Pareto preference)

Let P, I, R be the relations of Preference, Indifference and Incomparability (see [10]). Then, the triplet $\{P,I,R\}$ of relations on T represents a Pareto preference structure, if and only if, $\forall (a_j,a_k) \in A \times A$:

$(X_j,Y_j)P(X_k,Y_k) \Leftrightarrow X_j \le X_k,\ \ Y_j \ge Y_k,\ \ (X_j,Y_j) \ne (X_k,Y_k)$;

$(X_j,Y_j)I(X_k,Y_k) \Leftrightarrow (X_j,Y_j) = (X_k,Y_k)$;

$(X_j,Y_j)R(X_k,Y_k) \Leftrightarrow (X_j,Y_j)\neg P(X_k,Y_k) \wedge (X_k,Y_k)\neg P(X_j,Y_j) \wedge (X_k,Y_k)\neg I(X_j,Y_j)$.

◆

These relations are defined independently from any information on the relative importance of criteria. The Pareto preference (or dominance relation) is a strict **partial** order ([16]) which may include incomparability between some pairs of candidates.

The Pareto preference of DMU_j over DMU_k translates a sort of unanimity of points of view in favor of DMU_j . Thus, whatever *the aggregation procedure* might be, the global preference relation S, which is obtained on T_0 , implies that:

$$(X_j,Y_j)P(X_k,Y_k) \Rightarrow (X_j,Y_j)S(X_k,Y_k).$$

Let $E(P,T)$ denote the set of non dominated points in T with respect to P; then, the following inclusions are true:

$$E(S,T) \cap T_0 \subseteq E(P,T) \cap T_0 \subseteq E(P,T_0).$$

Since DEA considers the set T of "potential" DMUs, it is more selective than the relation P on T_0 . In fact, DEA models identifies as non efficient, not only the DMUs dominated by other DMUs in T_0 , but also those DMUs which are *convex-dominated* .

Let f be a *value function* , defined on T , such that $f(X_0,Y_0)$ is the "distance" of the DMU_0 from the *efficient frontier* of T ; then $f(X_0,Y_0)$ is obtained through the resolution of one of the basic DEA models. It is possible to prove that f is isotone with respect to the Pareto preference on T , i.e.:

i) $(X_j,Y_j)P(X_k,Y_k) \Rightarrow f(X_j,Y_j) < f(X_k,Y_k)$

ii) $f(X_j,Y_j) = f(X_k,Y_k) \Rightarrow (X_j,Y_j)R(X_k,Y_k) \vee (X_j,Y_j)I(X_k,Y_k)$;

iii) $f(X_j,Y_j) < f(X_k,Y_k) \Rightarrow \exists h \text{ such that } f(X_h,Y_h) = f(X_j,Y_j) \text{ e } (X_h,Y_h)P(X_k,Y_k)$

If one intends to give a "score" to each DMU, then this "score" must be non increasing with the value function f , which constitutes an estimate of the improvement needed to reach efficiency.

The function f implicity deals the candidate incomparability in arbitrary manner; so that if f is used to rank the DMUs in T_0 , then some problems may appears. Indeed, the DMU evaluations (and, consequently, the ranking) depend on the assumptions made on the production set T (convexification of T_0), and on the way by which one DMU is projected

on the frontier (envelopment metric). DEA tecnique ranks all Pareto efficient DMUs before all other candidates. This result appears as disturbing, since it is well-known that some non efficient candidates are likely to be *brillant second* . (See [3] for the example of Figure 1).

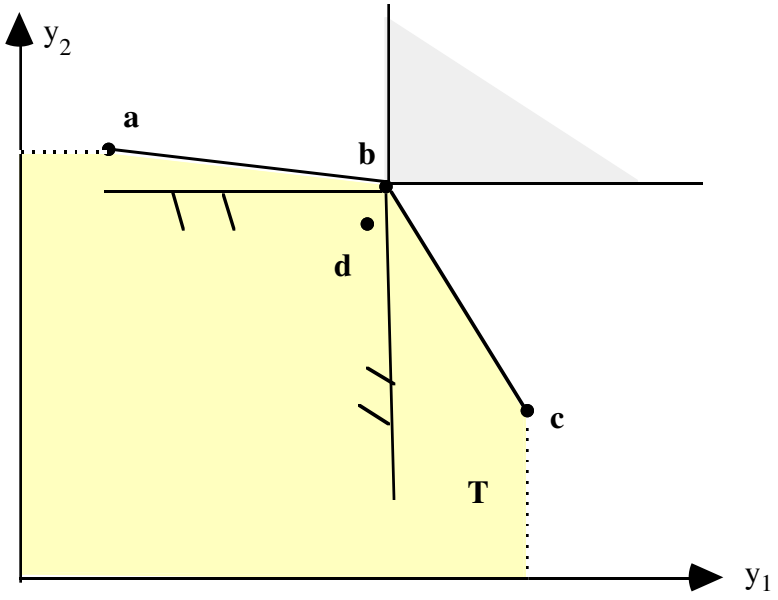

Figure 1

The DMUs *a, b, c* , are efficient and DEA assigns them the highest score. The DMU *d*, which is dominated by *b* , is not comparable with respect to *c* and *a* , but it receives a lower score; *d* could however be an excellent second in the ranking of the DMUs.

3. A multi-judge/multi-criteria model with two decision levels

We propose a new multi-judge/multi-criteria model, with two decision levels, which allows us to set the DEA models within the paradigm of MCDA/M (see [10]). With this purpose, we focus on the ratio characterization of the CCR model (see [4]).

The CCR ratio construction incorporates a process of assigning "weights" to the criteria, in an apparently similar way as it is done with the *single synthesizing criterion approach* of MCDA (see [10]) (This includes the methods of MultiAttribute Utility Theory (MAUT) (see [7])). The "weight" denomination, however, does not have the same meaning in DEA and in MAUT or MCDA approaches. It would be more correct to call the "weights" as "model parameters", to avoid a methodological confusion or some errors. One example of this is shown by the fact that a set of "weights" is determined *for each DMU* to show it in the

best possible light, instead of a unique definition of "weights" common to all DMUs, *as required by MAUT and by all MCDA procedures* .

The *efficiency measure* of a DMU_0 is formally defined as a ratio between a "weighted" sum of outputs (*virtual* output) and a "weighted" sum of inputs (*virtual* input). This ratio has to be maximized and constitutes the objective function of the formal model (1) for efficiency evaluation of one specific DMU_0:

$$\max_{\mu,\omega} \; h_0 = \frac{\sum_{r=1}^{s} \mu_r y_{r0}}{\sum_{i=1}^{m} v_i x_{i0}} \tag{1}$$

s.t.

$$\frac{\sum_{r=1}^{s} \mu_r y_{rj}}{\sum_{i=1}^{m} v_i x_{ij}} \le 1, \quad \forall j = 1,\ldots,n$$

$$\mu_r, v_i > 0, \quad i = 1,\ldots,m; \quad r = 1,\ldots,s.$$

The fractional linear program (1) can be transformed into an equivalent Linear Programming problem ([4]). It is possible, therefore, to measure the efficiency of a DMU set, by solving n LP problems.

The model (1) seeks the best *strategy* for DMU_0 global evaluation, that is the "best set of weights" favourable to DMU_0, considering the performances of the other units in T. Such a problem formulation for evaluating the efficiency of DMU_0 may be intended as an aggregation procedure through a unique synthesizing criterion approach. In fact, starting from the optimal solution (v^{0*}, μ^{0*}) of model (1), it is possible to define a *value function* , $h_0 : T \to R$, which assigns the following "efficiency score" to each DMU:

$$h_0(x,y) = \frac{\sum_{r=1}^{s} \mu_r^{0*} y_r}{\sum_{i=1}^{m} v_i^{0*} x_i}. \tag{2}$$

The value function, $h_0(\cdot)$, determines a total preorder P^0 on T , according to the decision strategy which is the most favourable to DMU_0. The total preorder P^0 is defined on T_0 , as follows:

$$DMU_j P^0 DMU_k \Leftrightarrow h_0(X_j, Y_j) > h_0(X_k, Y_k); \tag{3}$$

$$DMU_j I^0 DMU_k \Leftrightarrow h_0(X_j, Y_j) = h_0(X_k, Y_k).$$

The solution of model (1), for each observed DMU of T_0, supplies n value functions in the form (2), and, consequently, a total preference preorder P^j, for each observed DMU_j:

$$h_j(x,y) = \frac{\sum_{r=1}^{s} \mu_r^{j*} y_r}{\sum_{i=1}^{m} v_i^{j*} x_i}, \qquad j = 1,\ldots,n. \tag{4}$$

These may be considered as "marginal" evaluations, assigned by some particular observers (or judges). Let us, therefore, assume a decision situation in which two decision levels are recognizable: a **strategic** decision level, where a unique decision maker (individual or a group) is identified, and an **operational** decision level where a "judge" is defined for each DMU of T_0 (Figure 2).

The **operational** decision level is characterized by the following elements:

- object (or candidate) set: $T_0=\{DMU_j, j=1,...,n\}$;
- criteria set: $G = [g_1,\ldots,g_p]$;
- the evaluator of any DMU_0 of T_0 is a "sponsor" judge G_0. This judge is a "party-man", which seeks a strategy s (that is a set of "weights") that permits the best possible efficiency DMU_0's score to be achieved. In these terms, the evaluation may be intended as a self-evaluation process, by which each DMU assesses its own efficiency measure.

The judge G_0 uses the value function $h_0(\cdot)$ to comparatively evaluate both DMU_0 and the other DMUs, so as to provide a total preorder P^0 on T_0, defined as in (3).

The **strategic** decision level may be characterized by the following elements:

- object (or candidate) set: T_0 ;
- criteria set: the total preorders P^j, $j = 1,\ldots,n$, each provided by one of the n "sponsor" judges;
- the evaluator of the observed DMUs in T_0 is the Decision Maker (DM). DM is one subject whose evaluation process reflects, generally, some economical/financial, managerial/political goals to be pursued.

The problem that remains "open" is that of defining an aggregation procedure for the strategic level, which might be considered "objective" as most as possible.

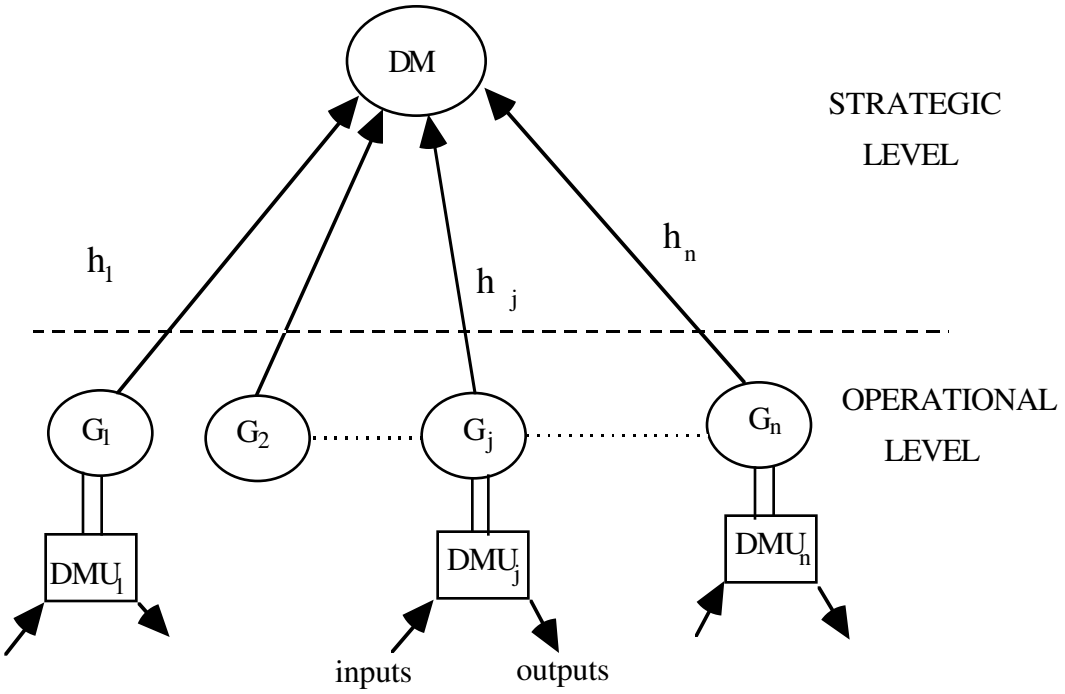

Figure 2.

4. Aggregation problem at strategic level

The decision situation at a strategic level may be then interpreted as a multicriteria evaluation problem, in which the n total preorders are the criteria.

The DM's problem tackled is a typical aggregation problem ([10]) that cannot be solved by DEA, as shown in section 3.

Each preorder P^j associates to T_0 one evaluation function $h_j(\cdot)$, $j = 1,\ldots,n$. It is, then, necessary to select one suitable strategy, regarding the aggregation.

Two strategy are possible:

choice 1) to aggregate the evaluations associated to the n preorders, by a unique synthesizing criterion approach;

choice 2) to aggregate the n preorders according to an outranking relation or another ordinal rule of synthesis.

Different aggregation procedures can be applied for each strategy, all requiring a suitable definition of the criteria "weight" ([8], [9], [11], [16]).

5. Restricting weight flexibility in DEA

Most of the recent DEA literature deals with extensions of the basic DEA models, in order to incorporate value judgments ([1]).

One way of modelling such value judgments in DEA leads to define some constraints on the "weight" set. This situation is interpretable in the multi-judge framework as follows: the DM gives information on his "weight" set, so as to put common constraints to all judges G_j, $j=1,\ldots,n$.

The limit situation occurs when the DM is able to made explicit its own "weights". In this case, all judges G_j should apply the same set of weight in the self-evaluation process.

The aggregation approach proposed in section 4 allows a more "dialectic" evaluation process in organizational contexts characterized by a multi-actiorial system. The strategic DM may express his preference in a more clear way and take some advantages from an interaction with the "sponsor" judges, for istance by a learning process. He can also control the local processes of DMUs by assigning different "weights" to judges, selecting an ordinal or cardinal aggregation procedure, or according to an interactive approach, seeking an agreement with the n judges.

References

[1] Allen, R., A. Athanassopoulos, R.G. Dyson, and E. Thanassoulis (1997) "Weights Restrictions and Value Judgments in Data Envelopment Analysis: Evolution, Development and Future Directions", *Annals of Operations Research* , 73, 13-34.

[2] Belton, V. and S.P., Vickers (1993) "Demystifying DEA - A Visual Interactive Approach Based on Multiple Criteria Analysis", *Journal of the Operational Research Society*, 44:9, 883-96.

[3] Bouyssou, D. (1997) "DEA as a Tool for MCDM: some remarks", in N.Meskens and M. Roubens (eds), *Extended-Abstract Proceedings of the International Conference on Methods and Applications of MCDM*, MONS, (14-16 May 1997), 57-60. Mons: University.

[4] Charnes, A., W.W. Cooper, and E. Rhodes (1978), "Measuring the Efficiency od Decision Maling Units", *European Journal of Operational Research* 2:6, 429-44.

[5] Doyle, J. and R. Green (1993), "Data Envelopment Analysis and Multiple Criteria Decision Making", *OMEGA* , 21:6, 713-15.

[6] Farinaccio, F. (1998), "Data Envelopment Analysis come strumento di Analisi Multicriteria", Tesi di Dottorato, Università di Pisa.

[7] Keeny, R. and H. Raiffa (1976), *Decisions with Multiple Objectives; Preferences and Value Trade-offs* . Chichester: John Wiley and Sons.

[8] Larichev, O.I. and H.M. Moshkovich (1995), "ZAPROS-LM - A method and system for ordering multiattribute alternatives", *European Journal of Operational Research* , 82, 503-521.

[9] Marcotorchino, J.F. and P. Michaud (1979), *Optimisation en analyse ordinale des données* . Paris: Masson.

[10] Roy, B. (1985), *Méthodologie Multicritère d'Aide à la Décision* . Paris: ECONOMICA.

[11] Roy, B. and D. Bouyssou (1993), *Aide Multicritère à la Décision: Méthodes et Cas*. Paris: ECONOMICA.

[12] Roy, B. and D. Vanderpooten (1996), "The European School of MCDA: Emergence, Basic Features and Current Works", *Journal of Multi-Criteria Decision Analysis*, 5, 22-38.

[13] Seiford, L.M. (1997), "A Bibliography for Data Envelopment Analysis (1978-1996)", *Annals of Operations Research* , 73, 393-438.

[14] Stewart, T.J. (1992), "A Critical Survey on the Status of Multiple Criteria Decision Making Theory and Practice", *OMEGA*, 20:5/6, 569-86.

[15] Stewart, T.J. (1996), "Relationships between Data Envelopment Analysis and Multicriteria Decision Analysis", *Journal of the Operational Research Society* , 47, 654-65.

[16] Vincke, P. (1992), *Multicriteria Decision Aid* . Chichester: John Wiley and Sons.