

Instituto de Engenharia de Sistemas e Computadores de Coimbra
Institute of Systems Engineering and Computers
INESC - Coimbra

Reinhard Madlener

Carlos Henggeler Antunes

Luis C. Dias

Multi-Criteria versus Data Envelopment
Analysis for Assessing the Performance of Biogas Plants

No. 6

2006

(Also published as CEPE Working Paper No. 49, ETH Zurich)

ISSN: 1645-2631

Instituto de Engenharia de Sistemas e Computadores de Coimbra

INESC - Coimbra

Rua Antero de Quental, 199; 3000-033 Coimbra; Portugal

www.inescc.pt

Multi-criteria versus data envelopment analysis for assessing the performance of biogas plants¹

Reinhard Madlener^a, Carlos Henggeler Antunes^b, and Luis C. Dias^c

^a Centre for Energy Policy and Economics (CEPE), Department of Management, Technology, and Economics, ETH Zurich, Zurichbergstrasse 18 (ZUE E), 8032 Zurich, Switzerland.
Email: rmadlener@ethz.ch

^b Department of Electrical Engineering and Computers, University of Coimbra and INESC Coimbra, Rua Antero de Quental, 199, Coimbra, Portugal
Email: cantunes@inescc.pt

^c Faculty of Economics, University of Coimbra and INESC Coimbra, Rua Antero de Quental, 199, Coimbra, Portugal
Email: ldias@inescc.pt

July 2006

ABSTRACT. This paper compares multi-criteria decision aiding (MCDA) and data envelopment analysis (DEA) approaches for assessing renewable energy plants, in order to determine their performance in terms of economic, environmental, and social criteria and indicators. The case is for a dataset of 41 agricultural biogas plants in Austria using anaerobic digestion. The results indicate that MCDA constitutes an insightful approach, to be used alternatively or in a complementary way to DEA, namely in situations requiring a meaningful expression of managerial preferences regarding the relative importance of evaluation aspects to be considered in performance assessment.

KEYWORDS. Multi-criteria decision analysis; DEA; Renewable energy; Biogas

1. INTRODUCTION

Over the last two decades, a growing environmental awareness has changed the focus of energy planning processes from an almost exclusive concern with cost minimization of supply-side options to the need of explicitly including multiple and conflicting aspects, such as cost and environmental issues, in decision support models. It is now widely recognized that the largest source of atmospheric pollution is fossil fuel combustion, on which current energy production and use patterns rely heavily. Therefore, most crucial environmental problems derive from energy demand to sustain human needs and economic growth. Despite of this, their effective potential is far from being exploited, and Renewable Energy Sources (RES) are becoming increasingly important as supply-side options to satisfy energy needs, taking into account their dispersed generation capabilities, low levels or absence of pollutant emissions, and waste valuation potential. However, some drawbacks can also be associated with RES, such as their intermittent nature, as in the case of wind turbines, and various types of environmental impacts.

¹ Paper presented at the 19th Mini EURO Conference on Operational Research Models and Methods in the Energy Sector (ORMMES'06), Coimbra, Portugal, 6-8 September 2006.

The Kyoto Protocol and the EU Renewables Directive are examples of ambitious political goals, which fostered the development of generation technologies based on RES. In this paper we address the case of the effective promotion of agricultural biogas plants in Austria, which use mainly energy crops (silage) for digestion, through guaranteed feed-in tariffs for electricity sold to the grid.

The assessment of the global performance of different entities (potential solutions, courses of action) can no longer be based on a single-dimensional axis of evaluation, such as cost or benefit. Multiple, incommensurate and often conflicting axes of evaluation of distinct nature are inherently at stake. Therefore, economic, technical, societal, and environmental aspects must be explicitly taken into account in models for decision support, rather than aggregated in a single indicator (generally of economic nature).

DEA is generally used to evaluate the efficiency of Decision Making Units (DMUs), which are comparable organizational entities performing similar tasks in a homogeneous operating environment. The introduction of preference managerial information is often relevant when assessing the units' performance. In fact, a manager is not indifferent as to whether a unit is efficient, using a less important combination of inputs and/or outputs and neglecting inputs/outputs of the utmost importance.

Uncertainty is an intrinsic characteristic of real-world problems arising from multiple sources of distinct nature. Uncertainty emerges from the ever-increasing complexity of interactions within social, economic and technical systems, characterized by a fast pace of technological evolution, changes in market structures, and new societal concerns. It is generally impracticable that decision aid models could capture all the relevant inter-related phenomena at stake, get through all the necessary information, and also account for the changes and/or hesitations associated with the expression of the stakeholders' preferences. Besides structural uncertainty associated with the global knowledge about the system being modeled, input data may also suffer from imprecision, incompleteness, or be subject to changes. In this context, it is important to provide managers and decision makers with robust conclusions. The concept of robust solution is generally linked to a certain degree of "immunity" to data uncertainty, to an adaptive capability (or flexibility) regarding an uncertain future or ill-specified preferences, guaranteeing an acceptable performance even under changing conditions (drifting from "nominal data").

This paper uses both DEA and MCDA approaches for assessing the efficiency of 41 agricultural biogas plants, with the purpose of obtaining complementary insights about these evaluations as well as the underlying methodologies. This has led us to face the issue of how can MCDA methods be used in the context of efficiency evaluation, trying to keep the spirit behind DEA, while being able to use MCDA's capabilities of incorporating the preferences of a decision maker.

The paper is organized as follows. Section 2 introduces and compares the two analytical frameworks studied. Section 3 describes the case study and how DEA and MCDA have been applied. The main results obtained are reported in section 4. In section 5 a discussion of the findings is made and some conclusions are drawn.

2. COMPARISON OF ANALYTICAL FRAMEWORKS

2.1 DEA

The attainment of high levels of performance is a key issue for the success of every organization. Therefore, an adequate management framework is necessary for evaluating the current performance, identifying benchmarks to use in seeking improvements, and understanding why some units in a particular organization are operating (in)efficiently.

DEA is a non-parametric performance measurement technique, based on linear programming, for assessing the relative efficiency of DMUs. DMUs are homogeneous entities (such as sales outlets, electricity distribution companies, bank branches, schools, university departments, etc.) with some decision autonomy, operating a production process that converts a set of inputs into a set of outputs. DEA models use these inputs and outputs to compute an efficiency score for a given DMU when this DMU is compared with all the other DMUs. The relative efficiency of a DMU is defined as a ratio between the sum of its weighted output levels to the sum of its weighted input levels. In contrast to other parametric econometric approaches, such as stochastic frontier analysis, DEA does not assume any specific functional form, thus avoiding problems of model misspecification.

In DEA, a DMU is considered efficient if there is no other DMU, or a linear combination of inputs and outputs of several DMUs, that can improve one input or output, without worsening the value of at least another one. The frontier is defined by the observed values of the (relatively) efficient DMUs. If a DMU does not belong to this envelopment surface (the convex hull of the efficient DMUs) and lies in its interior, then that DMU is operating inefficiently. DEA models usually return an efficient projection point of operation on the frontier for each inefficient DMU, thus identifying the DMUs that can be used as performance benchmarks for the DMUs that are operating inefficiently.

Three basic DEA models are generally distinguished (see Charnes et al., 1994, for a presentation and comparative analysis of these models):

- 1) *CCR model* – This model was presented in the seminal work of Charnes, Cooper and Rhodes (1978). The CCR model is based on the radial minimization (maximization) of all inputs (outputs) and assumes an environment of Constant Returns to Scale (CRS);
- 2) *BCC model* – The Banker, Charnes and Cooper (1984) model is the Variable Returns to Scale (VRS) version of the CCR model. The difference between the two types of envelopment surfaces, CRS and VRS, is the presence of a convexity constraint;
- 3) *Additive model* – The additive model originates in the work of Charnes et al. (1985). This model maximizes the L_1 distance (also known as “city-block distance”) of the DMU under analysis to the observed efficient frontier and assumes VRS.

DEA models have been extensively used to assess the performance of DMUs in a broad range of real-world problems. However, some important issues regarding the application of DEA with real-world data remain. Firstly, the complete weight flexibility assumed by DEA models often leads to efficiency results that are difficult to justify. The freedom of each DMU to choose the weights of inputs and outputs that

show it under the best possible light can lead to the assignment of very low weights. In practice, this means that certain inputs or outputs are effectively ignored (a disturbing effect of the free specialization allowed in DEA models, which is not generally acceptable in practice).

Moreover, the inputs and outputs can be weighted in a manner that contradicts the views and/or preferences of the organization and their stakeholders, or even in a quite counterintuitive manner by valuing secondary inputs or outputs more than priority ones (Joro and Viitala, 2004). In fact, the inputs and outputs are not generally equally relevant and some preference information must be included in the analysis. Also, whenever the number of inputs and outputs grows the trend is that more DMUs become efficient, thus impoverishing the discriminating power of the DEA models.

One of the techniques generally used to circumvent these issues is the introduction of additional restrictions on the variation allowed for the weights. These restrictions (absolute lower and upper bounds on weights or the ordering of input or output weights) intend to capture value judgments elicited from the managers on the perceived importance of inputs and outputs. However, as pointed out by some authors (Podinovski, 2004), the resulting efficiency score of weight-restricted models cannot be interpreted as a realistic improvement factor (because the efficient radial target of an inefficient DMU is no longer technologically feasible).

In real-world problems, in which inputs or outputs are less tangible, the application of DEA models is also problematic. In fact, market costs and prices may not be readily available, which introduces an additional degree of uncertainty to the results.

Another issue associated with DEA is the need of having in the DMUs set units with comparable production levels. A very large unit is at once deemed efficient because there are no other units with a similar production level. DEA models also lead to an amplification of the scale effect because small units are often appointed as scale efficient.

2.2 MCDA

These considerations led us to envisage the use of MCDA models to perform efficiency evaluation. However, instead of attempting to assign an efficiency measure to each DMU we believe that, in most real-world situations, assigning the DMUs to ordered efficiency categories is sufficient for analysis and provides more confidence on the results, in the sense of robustness to changes either in data or managers' preferences, than a single numerical figure. Moreover, a more detailed analysis within each efficiency category is always possible whenever it is found useful to improve the discrimination of the evaluation model.

In assessments of the performance of DMUs in which technical, economic and environmental aspects are at stake for, it is often important to use known standards (or theoretical maxima) and efficiency profiles. Also, there are situations in which DMUs must be appraised for efficiency on an "as they come" basis, i.e., they are not included in a given set of DMUs. This required capability of evaluating each DMU in absolute terms, and not just in comparison with other peers, as well as the need to include evaluation aspects expressed in different units and even measured qualitatively (that is, allowing independence towards scales), can be achieved using the ELECTRE TRI method (Yu, 1992).

The ELECTRE TRI method belongs to the ELECTRE family of multi-criteria methods developed by Bernard Roy and his co-workers (Roy, 1991; 1996). ELECTRE methods are based on the construction and exploitation of a so-called outranking relation between courses of action (DMUs in our context). Roy distinguishes three main problem types: choice, ranking and sorting (or classification). ELECTRE TRI is devoted to the sorting problem, which consists in assigning each alternative to one of a set of pre-defined ordered categories according to a set of evaluation criteria. The categories (C^h) are defined by specifying their boundaries (b^h) by means of reference actions, in terms of the performance they achieve in each criterion (see fig. 1).

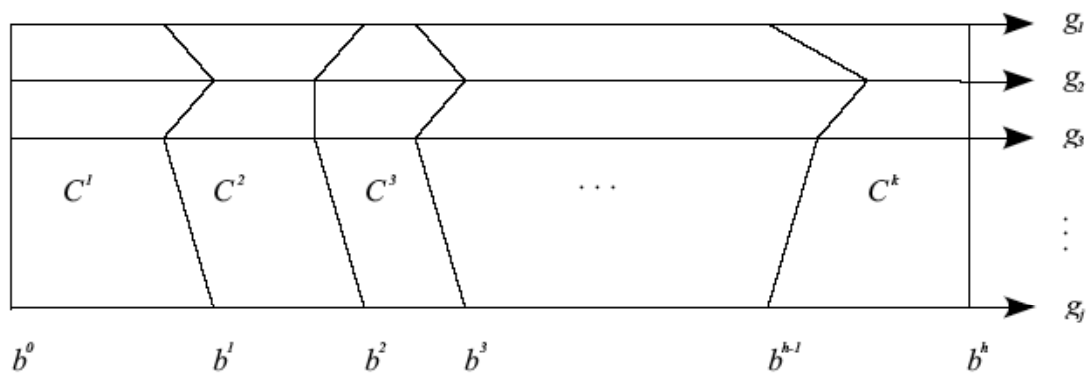


Figure 1. Definition of categories C^h with reference actions b^h .

The assignment of each alternative a to a category C^h is done by comparing its value in each criterion to the performances of the reference actions. The procedure assigns each action to the highest category such that its lower bound b^{hw} is outranked by a . The outranking relation is verified by comparing a credibility index, computed using the differences in performance and the criterion weights, with a cutting level λ ($\lambda \in [0.5, 1]$) which defines the “majority requirement”, hence the exigency of the classification. For further details about ELECTRE TRI see Yu (1992) and Mousseau et al. (1999), among others.

In the present application of dealing with an efficiency evaluation problem by means of a multi-criteria sorting model the software package IRIS 2.0 (Dias and Mousseau, 2003) has been used. IRIS implements a methodology developed by Dias et al. (2002) that is based on the ELECTRE TRI method, but which accepts uncertainty in the input parameters. The main characteristics of this software package are:

- Acceptance of imprecision regarding the criterion weights and the cutting level through the definition of intervals for each parameter, or the definition of linear constraints.
- Acceptance of classification examples, with the input of the better and worse category that each action can be assigned to. This is translated into constraints on the parameters that guarantee that those example results are reproduced.
- Inference of a combination of parameters that limits the violation of the constraints in the case of inconsistency, by minimizing the maximum deviation. It

is also possible to find the constraint subsets, which must be removed to restore consistency.

- Inference of a central combination of parameters through the maximization of the minimum slack associated with the constraints, when the constraints are consistent. For each alternative, it is shown which category represents this central combination, and the other possible classifications that respect the imposed constraints.

Multi-criteria methods usually require a set of parameters that embody the preferences of the decision makers. The ELECTRE TRI method requires the specification of the reference profiles associated with the categories (b^0, \dots, b^h), the criterion weights, and the cutting level (λ). Also, a set of indifference (q_j), preference (p_j) and veto (v_j) thresholds for each criterion and reference profile can be defined. Indifference and preference thresholds characterize the acceptance of imprecision in the judgment by considering as indifferent two actions when their performances in each criterion j differ less than a specified amount q_j . Moreover, the transition from indifference to preference is made gradual, changing linearly from q_j to p_j . The veto thresholds are aimed at capturing situations in which very bad scores in any criterion should prevent an alternative of being classified in the best category or if these bad scores should force it to be classified in the worst category, independently of having very good scores in all other criteria. This enables, as it is often required in practice, to introduce a certain level of non-compensation into the evaluation model.

A relevant issue in this context is the meaning of the weights in ELECTRE methods. In this type of methods, weights are perceived as true coefficients of importance assigned to the criteria, and not just as technical devices for translating the performances in the several criteria into a common value measure. Therefore, they are scale independent (that is, they are not linked to the scales in which each criterion is measured), thus making them easier to be specified by managers. These parameters bear the preference information and insights into the sorting process. In principle, they must be elicited from managers and stakeholders (preferably via an analyst with expertise on the methodological component). It must be noticed that this method imposes a non-negligible burden associated with the specification of all the parameters required. However, some of these parameter data can be preset according to the experience of the analyst, in general associated with previous case studies. For instance, indifference and preference thresholds can be fixed as percentages (say 1% and 10%, respectively) of the value ranges in each category.

The IRIS software allows for the consideration of uncertainty in the weights (as well as in the cutting level). This feature contributes to reducing the data requirements and increasing the confidence in the results.

3. CASE STUDY

In Austria, an effective promotion of renewable energy technologies has been pursued in recent years, driven by the need to achieve ambitious energy and climate policy goals. Examples of this trend are the goals contained in the Kyoto Protocol (-13% greenhouse gas emissions by 2008/12, relative to 1990 levels) or the EU Renewables Directive (renewable electricity share of Austria to be raised to 78.1% by 2008, compared to e.g. 70% in 1997). In particular, the last few years witnessed a remarkable boom in the construction of agricultural biogas plants due to the

introduction of feed-in tariffs between 10.3-16.5 Cents€/kWh_{el}, guaranteed for a period of 13 years, for ‘biogas’ electricity fed into the grid (Green Electricity Act, 2002). As a consequence, the number of plants rose from 110 at the end of 2003 to more than 200 by the end of year 2004 (Madlener et al., 2006). These plants use mainly energy crops (silage) for digestion.

However, up to now the promotion of energy crop digestion was hardly linked to any cost effectiveness or energy efficiency or other performance criteria. As a result many different technologies and specific applications occurred on the market, some of which were either not very productive, energy-efficient, or reliable.

Due to the attractive feed-in tariffs granted, anaerobic digestion of energy crops currently mainly aims at the generation of electricity. As a consequence, the heat energy produced in co-generation units remains largely wasted. Also, many plants use electricity for cooling purposes, in order to prevent adverse effects from self-heating of crop digesters. Therefore, in many cases up to two thirds of the available technical energy potential remains unused (Braun et al., 2005; Walla, 2005).

A monitoring and benchmarking project was initiated in 2004, which includes a detailed investigation of 41 Austrian energy crop digestion plants (cf. research database entry #8289 in www.rdb.ethz.ch). The project also aims at creating and establishing an evaluation system for the transparent assessment and benchmarking of the productivity of biogas plants by means of energetic, business economic, ecological and socio-economic criteria, characterizing the overall production cycle of biogas. Since anaerobic digestion has the potential of reducing greenhouse gas emissions (Braschkat et al., 2003), an important objective of the project is to evaluate the environmental impacts through the overall “crops to energy” process. Finally, positive and negative socio-economic impacts have been accounted for to a limited extent by means of a questionnaire survey among plant operators (subjective valuation, supplemented by measurable data).

4. RESULTS

4.1 Description of the data and parameters used

The DMUs are a representative set of energy crop digestion plants in Austria, aimed at covering the whole spectrum of existing plant types and operating conditions. Samples were taken from the substrate, digester, fermentation residues and biogas plant types. Also cooling, safe transport and appropriate storage were scrutinized. Installations are geographically well distributed over the country. They range from small-scale installations (down to 18 kW_{el}) in agricultural regions to larger scale plants (up to 1.7 MW_{el}). Single substrate (energy crops) installations as well as co-digestion plants (agricultural by-products and industrial bio-wastes) have been considered in the analysis.

The main groups of evaluation aspects at stake for assessing the efficiency of energy crop digestion plants are: (1) substrate provision, storage and pre-treatment; (2) biogas production (digestion); (3) net utilization of heat and electricity; (4) digestate handling and disposal; and (5) methane emissions.

In a first series of model specification, the following criteria have been considered for evaluating the efficiency of the energy crop digestion plants (for the sake of comparison between the DEA and the MCDA approaches): (1) labor (i.e. time) spent

for plant operation; (2) amount of substrate (organic dry substance, ODS) used; (3) amount of biogas produced or net electricity produced; and (4) methane emissions to the atmosphere (an undesirable output that contributes to the greenhouse gas problem, measured in CO₂ equivalents). For further details on data collection see Braun et al. (2005).

4.2 DEA

Figure 2 depicts the outcome of the DEA for two different model specifications (CCR and BCC output-oriented). In each of the model specifications reported, we have used substrate and labor as factor inputs and the amount of net electricity and external heat as (desirable) outputs.² Methane emissions have been considered as well in these first models. Some descriptive statistics are displayed in Table 1.

Table 1: Descriptive statistics ($N = 41$)

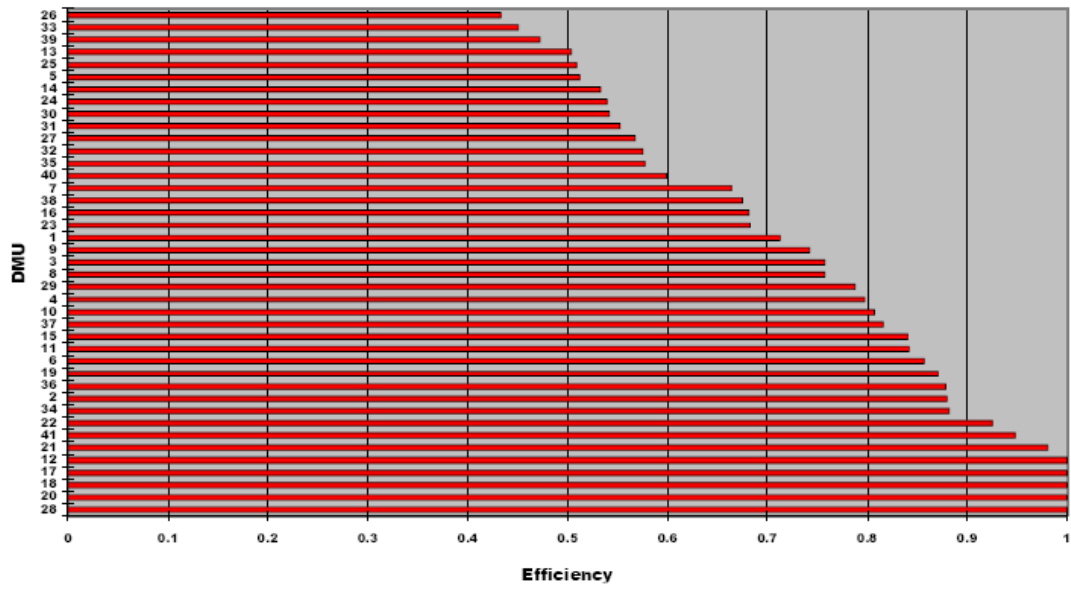
	Mean	SD	Min	Max
<i>Inputs:</i>				
i_1 – LABOR	1 581.29	1 934.47	50.42	10 950.00
i_2 – ODS	1 450.13	1 184.93	119.94	5 029.62
<i>Outputs:</i>				
o_1 – ELECPROD_NET	1 871 946.00	1 813 348.00	123 600.80	7 760 000.00
o_2 – HEATUSE_EXT	735 319.80	1 099 279.00	0.00	6 000 000.00
o_3 – METHANE	1 345.04	1 302.59	83.73	5 354.67

As can be seen in fig. 2, DMUs 12, (for BBC also 15), 17, 18, 20 and 28 form the efficiency frontier. For the case of the CCR model, 14 DMUs are below an (arbitrarily chosen) efficiency score of 0.6 and three below 0.5, while for the BCC model only 10 DMUs are below 0.6 and none below 0.5.

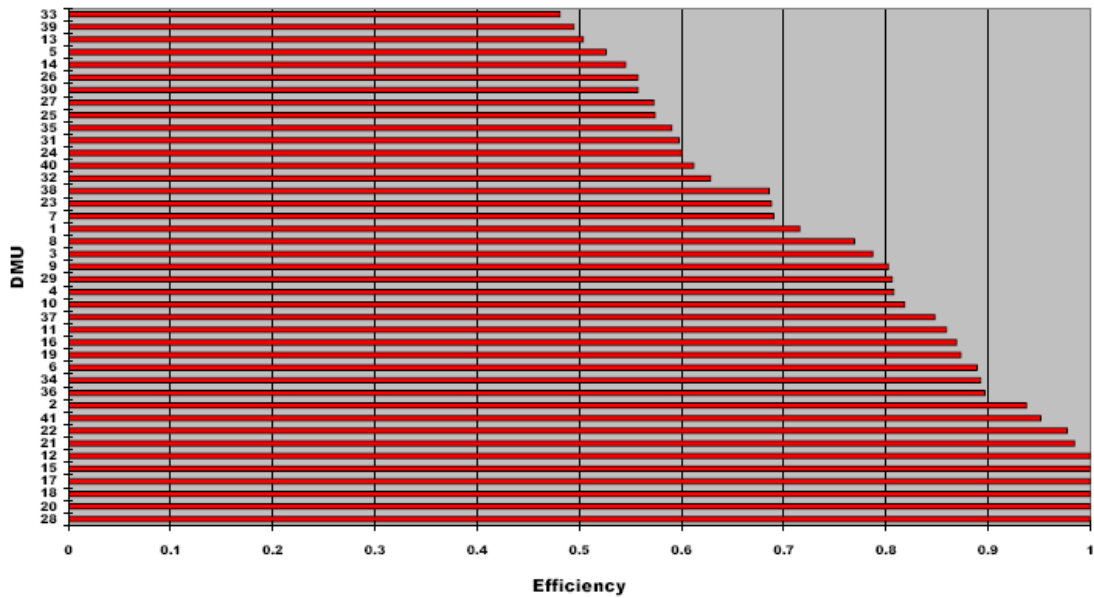
We also studied the sensitivity of the results with respect to the inputs and outputs considered in the DEA model. Interestingly, DMUs 12, 17, 20 and 28 are always on top of the rankings, determining the efficiency frontier, irrespective of whether labor is included as an input or not and whether the output is electricity only, heat and electricity, or biogas only.

For the CCR model, the ranking of the best is more sensitive. In particular, the efficiency frontier is often determined by one or two DMUs only (12 and 28 ranking on top or at least showing a very high score).

² ‘Net electricity’ and ‘external heat’ refer to the amount of electricity and heat delivered by the biogas plant for external consumption (i.e. net of what the biogas plant consumes itself), including farm operations not directly related to the biogas plant.

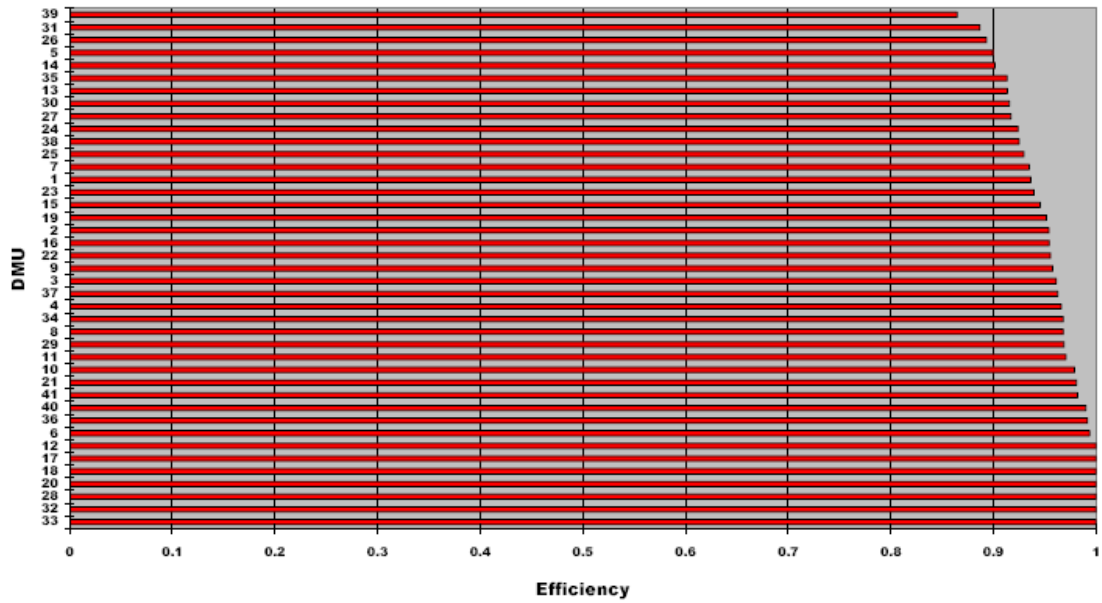


(a) CCR-O

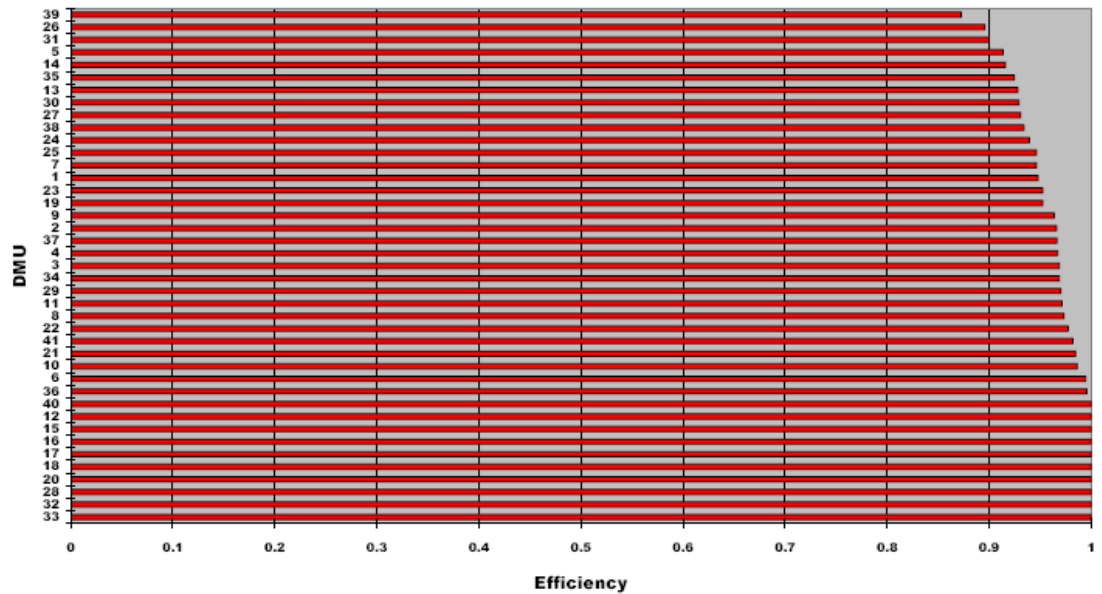


(b) BCC-O

Figure 2. DEA results (inputs: substrate used, labor spent; outputs: net electricity produced, external heat used).



(a) CCR-O



(b) BCC-O

Figure 3. DEA results (inputs: substrate used, labor spent, methane emissions; outputs: net electricity produced, external heat used).

CCR and BCC models were also built keeping the same inputs and outputs but adding methane emissions, which are considered an undesirable output, as another “input” to be minimized. These results are depicted in fig. 3. The main consequence of incorporating this new factor into the analysis is that DMUs 32 and 33 join the set of efficient solutions in both models. Note that DMU 33 was one of the worst-performing units, with scores lower than 0.5, before adding methane emissions, and it becomes efficient due to a very good performance on this new evaluation dimension. On the other end of the spectrum, DMUs 5, 14, 26, 31, and 39 turn out to be the least efficient units.

4.3 MCDA

Four efficiency categories are defined to classify the DMUs according to their efficiency: C^1 = “Poor”, C^2 = “Fair”, C^3 = “Good”, C^4 = “Very good”. The aim is to assign each plant to one of these ordered categories according to the multiple evaluation criteria.

To define the different categories, it is necessary to set the category bounds b^0, \dots, b^4 according to n_{crit} criteria/indicators which we denote as the evaluation functions $g_j(\cdot)$ ($j=1, \dots, n_{\text{crit}}$). The decision maker must set these bounds, taking into account that according to an indicator $g_i(\cdot)$, a DMU $_k$ with $g_i(\text{DMU}_k) \in [g_i(b^{j-1}), g_i(b^j)[$ should be placed into C^j . It then becomes clear that the DEA’s inputs and outputs cannot be taken as sorting criteria without some sort of adaptation. For instance, in which efficiency category should a DMU with electricity production (indicator o_1) of 900 000 be sorted? Clearly, this question regarding efficiency cannot be answered without knowing at least the dimension of the plant.

One of the modeling options is to consider as sorting indicators the ratio of inputs and outputs regarding some surrogate for the dimension of the plant, e.g., the amount of organic dry substrate (ODS) used. The indicators would then be Labor/ODS ($g_1=i_1/i_2$), Electricity/ODS ($g_2=o_1/i_2$), and Heat/ODS ($g_3=o_2/i_2$). Another option, which we have used in the experiments described below, is to use all ratios between the outputs and the inputs in the DEA model. That is, the multiple efficiency indicators are ratios combining an output to maximize or minimize (in the numerator) with an input. Thus, the following indicators have been considered: Electricity/Labor $g_1=o_1/i_1$ (max), Electricity/ODS $g_2=o_1/i_2$ (max), Heat/Labor $g_3=o_2/i_1$ (max), Heat/ODS $g_4=o_2/i_2$ (max), Methane/Labor $g_5=o_3/i_1$ (min), and Methane/ODS $g_6=o_3/i_2$ (min). The corresponding input data for the MCDA model is displayed in table 2. Although this approach leads to a high number of indicators as the number of criteria increases, it mimics the spirit of DEA: to allow each DMU to be evaluated according to multiple indicators and to choose the most favorable indicators (within the constraints that the decision maker may impose, as we will illustrate).

After the performance data of the 41 energy crop digestion plants for the different criteria have been introduced, the reference profiles are required, which define the limits between each category (see table 3). The profiles were defined such that approximately 1/4 DMUs would be placed in each category according to each indicator individually. Hence, for each indicator g_j and for each category C_k , there are approximately 10 plants that would be sorted into that category according to that indicator. The cutting level is specified within the interval [0.51, 0.67], that is it may vary from a simple majority to a 2/3-majority requirement.

Table 2. Efficiency indicators for the DMUs (most efficient DMUs marked boldface)

DMU	g_1 (max) Elec./Labor	g_2 (max) Elec./ODS	g_3 (max) Heat/Labor	g_4 (max) Heat/ODS	g_5 (min) Meth./Labor	g_6 (min) Meth./ODS
1	2 206.991	1 082.766	195.339	95.835	1.606	0.788
2	941.380	1 556.940	517.732	856.273	0.680	1.125
3	1 897.554	1 182.872	1 079.137	672.698	1.348	0.841
4	4 767.123	1 059.745	399.384	88.784	3.499	0.778
5	570.919	899.147	254.237	400.401	0.430	0.677
6	3 547.623	1 212.243	1 656.495	566.034	2.491	0.851
7	1 667.605	1 037.066	309.368	192.393	1.213	0.754
8	2 712.902	1 106.421	319.744	130.403	1.922	0.784
9	668.526	1 333.570	342.466	683.148	0.479	0.955
10	1 510.743	1 299.587	67.042	57.672	1.057	0.909
11	2 922.523	1 239.889	0.000	0.000	2.078	0.882
12	584.381	1 874.742	13.699	43.946	0.428	1.374
13	877.005	817.373	141.554	131.929	0.650	0.606
14	1 259.058	836.539	724.638	481.461	0.946	0.629
15	1 018.011	1 449.703	1 111.607	1 582.990	0.741	1.055
16	391.082	1 169.178	525.862	1 572.117	0.284	0.851
17	2 451.587	1 030.552	2 935.537	1 233.986	1.661	0.698
18	12 814.425	1 240.897	1 361.426	131.835	9.329	0.903
19	3 358.980	1 249.741	131.297	48.850	2.462	0.916
20	2 948.888	1 542.860	2 280.068	1 192.933	2.035	1.065
21	2 683.032	1 522.066	1 107.011	628.000	1.888	1.071
22	1 331.480	1 546.134	983.607	1 142.178	0.961	1.116
23	1 815.422	1 060.299	909.091	530.955	1.314	0.768
24	238.030	972.222	220.892	902.222	0.175	0.713
25	326.726	910.051	370.844	1 032.934	0.238	0.663
26	215.415	776.641	230.476	830.943	0.163	0.589
27	967.473	924.996	767.306	733.617	0.715	0.683
28	1 051.796	1 777.230	1 365.962	2 308.078	0.727	1.228
29	541.403	1 450.509	95.969	257.118	0.384	1.029
30	575.414	959.090	0.000	0.000	0.426	0.710
31	606.061	973.056	378.788	608.160	0.464	0.745
32	434.164	703.549	797.383	1 292.132	0.334	0.541
33	281.450	666.980	438.141	1 038.308	0.201	0.476
34	3 782.516	1 160.557	2 025.586	621.493	2.740	0.841
35	665.000	1 005.947	89.644	135.605	0.495	0.748
36	1 438.356	1 440.257	919.823	921.039	0.996	0.998
37	1 865.707	1 286.238	180.396	124.367	1.329	0.916
38	966.605	1 128.946	122.638	143.234	0.713	0.833
39	596.763	806.948	582.192	787.244	0.468	0.632
40	401.017	1 104.740	131.338	361.815	0.275	0.759
41	3 563.288	1 368.967	205.479	78.942	2.537	0.975

Table 3. Category definitions for each indicator

Category	g_1 (max) Elec./Labor	g_2 (max) Elec./ODS	g_3 (max) Heat/Labor	g_4 (max) Heat/ODS	g_5 (min) Meth./Labor	g_6 (min) Meth./ODS
C^1 - Poor	< 575	< 950	< 150	< 130	> 1.80	> 0.97
C^2 - Fair	[575, 1000[[950, 1125[[150, 375[[130, 530[]0.9, 1.8]]0.82, 0.97]
C^3 - Good	[1000, 2300[[1125, 1300[[375, 950[[530, 880[]0.45, 0.9]]0.71, 0.82]
C^4 - Very good	≥ 2300	≥ 1300	≥ 950	≥ 880	≤ 0.45	≤ 0.71

Table 4. Pessimistic, 50% majority, and pessimistic classifications

DMU	50%			DMU	50%		
	Optimistic	majority	Pessimistic		Optimistic	majority	Pessimistic
1	3	2	1	22	4	4	1
2	4	3	1	23	3	3	2
3	4	3	2	24	4	3	1
4	4	3	1	25	4	4	1
5	4	2	1	26	4	3	1
6	4	3	1	27	4	3	1
7	3	2	2	28	4	4	1
8	4	2	1	29	4	2	1
9	4	3	2	30	4	2	1
10	3	2	1	31	3	3	2
11	4	2	1	32	4	4	1
12	4	2	1	33	4	4	1
13	4	2	1	34	4	3	1
14	4	3	1	35	3	2	1
15	4	4	1	36	4	3	1
16	4	3	1	37	3	2	1
17	4	4	2	38	3	2	1
18	4	3	1	39	4	3	1
19	4	2	1	40	4	2	1
20	4	4	1	41	4	2	1
21	4	4	1				

Table 4 presents some conclusions that may be drawn without making any distinction between the relative importance of each indicator. The left column (“Optimistic”) indicates the classification that would result if the DMU was allowed to choose an indicator, i.e., if the DMU was allowed to specify the ELECTRE TRI weights of the indicators, setting one of them to have a weight equal to 1 and all the remaining indicators as weighing 0. These are the classifications that are more in accordance with the spirit behind DEA. The middle column (“50% majority”) indicates the classification that would result if we required the support of at least half of the indicators: a DMU is classified into category C^h if and only if three out of six indicators place it in that category (or in a better one). Finally, the rightmost column (“Pessimistic”) indicates the worst category suggested by some indicator. This means that a DMU is classified into category C^h if and only if all the six indicators place it in that category (or in a better one). Let us note that it may happen that a DMU can be classified into categories C^1 or C^4 , but not C^2 or C^3 . This may occur when a DMU is evaluated as belonging to C^1 according to some indicators and belonging to C^4 according to all the remaining indicators. If the indicators placing it in C^4 are

sufficient for the required majority (cutting level), then the DMU is sorted into C^4 , while otherwise it is sorted into C^1 .

If the DMUs were entirely free to choose the weights to assign to the indicators, then 33 of them would be sorted into C^4 (including all the efficient DMUs according to DEA) and the remaining eight into C^3 . This means that all of the DMUs would be sorted into the top two categories if each DMU was allowed to be judged according to only one of the six indicators, that single indicator being chosen by the DMU. In order to decrease the number of DMUs in the best categories several options can be envisaged: to make the category bound more demanding (i.e., to increase the b^j performances), and/or to require the support of more than one indicator (e.g. the support of half of the indicators, as depicted in table 4, and/or to add information about the relative power of the indicators.

ELECTRE TRI allows to incorporate the managerial judgment about how important each indicator is and whether a very low performance in some indicators may be an impediment to reach the highest categories. Using IRIS, it is possible to compute the range of categories for a DMU that is compatible with a set of parameter constraints. Let us first consider, for instance, the following constraints as an illustrative case:

(a) The manager states that the most important output is electricity, followed by methane emissions (to be minimized), and lastly by heat. This implies that the importance of g_1 (electricity/labor) cannot be lower than the importance of g_5 (methane/labor), which in turn cannot be lower than the importance of g_3 (heat/labor). Analogously, the importance of g_2 (electricity/ODS) cannot be lower than the importance of g_6 (methane/ODS), which in turn cannot be lower than the importance of g_4 (heat/ODS).

(b) Concerning the inputs, the manager states that ODS is more important than labor. This implies that the importance of g_2 (electricity/ODS) cannot be lower than the importance of g_1 (electricity/labor). Analogously the importance of g_4 (heat/ODS) cannot be lower than the importance of g_3 (heat/labor), and the importance of g_6 (methane/ODS) cannot be lower than the importance of g_5 (methane/labor).

The results corresponding to these requirements are shown in table 5. The column "Suggested" indicates the classification corresponding to the weight values inferred by IRIS ($k_2=0.2667$, $k_1=k_6=0.2$, $k_4=k_5=0.1333$, $k_3=0.0667$). The column "Optimistic" corresponds to a situation where the DMUs could choose their weights (provided that the imposed constraints were satisfied). Optionally, the ELECTRE TRI models also allow incorporating veto thresholds, such that, for instance, a DMU that is classified as C^1 according to a given indicator will not be able to reach category C^4 in a multi-criteria evaluation. Supposing the manager would consider that the most important indicator is g_2 (electricity/ODS) should also have some veto power, such that a DMU with a ratio less than 950 could not achieve category C^4 , then the only change would be that DMUs 32 and 33 could no longer reach C^4 due to their low performance under that indicator.

Similar types of ad-hoc robustness analysis could easily be carried out in order to capture the imprecision associated with the specification of some parameters and identify those for which small changes reveal to have more impact on the results. A form which is easily perceived by managers is to ask for intervals for some parameters (for instance, the weights), aimed at capturing information that is not precisely known but can be taken as bounded within some acceptable limits.

When taking the suggested parameters inferred by IRIS, taking into account the constraints on the indicator (criterion) weights, the results may be rather different from those obtained with DEA. However, the MCDA analysis may complement the DEA analysis by providing another perspective from which the conclusions of DEA may be strengthened or weakened. In our illustrative example, among DEA-efficient DMUs, only DMUs 17 and 20 reach C^4 using the sorting of IRIS, and these are the two only DMUs among the 41 to reach that category. Despite being DEA-efficient, IRIS sorts DMUs 32 and 33 in C^1 , reflecting their poor performances in two of the indicators that weigh more (g_1 and g_2). Concerning the worst DMUs according to DEA, DMUs 5 and 25 are sorted into the worst category, whereas DMUs 14, 31 and 39 are sorted into C^2 .

Table 5. Classification subject to importance constraints

DMU	IRIS suggestion	Optimistic	DMU	IRIS suggestion	Optimistic
1	2	3	22	3	4
2	3	4	23	3	3
3	3	3	24	2	3
4	2	3	25	1	4
5	1	2	26	1	3
6	3	3	27	2	3
7	2	3	28	3	4
8	2	3	29	1	4
9	2	4	30	2	2
10	2	3	31	2	3
11	2	3	32	1	4
12	2	4	33	1	4
13	2	2	34	3	3
14	2	3	35	2	2
15	3	4	36	3	4
16	3	3	37	2	3
17	4	4	38	2	3
18	2	3	39	2	3
19	2	3	40	2	2
20	4	4	41	1	4
21	3	4			

5. DISCUSSION AND CONCLUSIONS

DEA is a data oriented approach and it requires no a priori specification of the functional form of the production model converting inputs into outputs. Units are then free to choose their most favorable weights for becoming efficient when compared with their peers. On the other hand, this can present a disadvantage whenever over-specialization must be avoided in the consumption of inputs or the production of outputs, which amounts to practically ignore some inputs and outputs. Moreover, managerial preference information is often required since inputs and outputs do not have generally the same importance in assessing the efficiency of operational units. Therefore, models for efficiency evaluation must explicitly incorporate meaningful techniques to take weights into account, understood as coefficients of relative importance of inputs and outputs. This has been the main motivation for the use of MCDA techniques, in order to assess the extent in which these could overcome those

characteristics of DEA, and what adaptations would be needed. Therefore, we are not proposing MCDA as an approach to replace DEA as a performance evaluation tool but rather as a complementary technique, namely as far as the meaningful introduction of managerial preferences is concerned.

ACKNOWLEDGMENTS

This research has been partially funded by the Austrian Federal Ministry of Transport, Innovation, and Technology (BMVIT) under Project Grant No. 807742, and the Portuguese Foundation for Science and Technology under Project Grants POCI/EGE/58371/2004 and POCTI/ESE/38422/2001. Fruitful comments provided by Michael Laaber of IFA-Tulln are gratefully acknowledged.

REFERENCES

- Ball VE, Lovell CAK, Nehring RF, Somwaru A, Incorporating undesirable outputs into models of production: An application to US agriculture, *Cahiers d'Economie et Sociologie Rurales* 31, 59-73, 1994.
- Banker RD, Charnes A, Cooper W.W., Some models for estimating technical and scale inefficiencies in data envelopment analysis, *Management Science*, 30, 1078-1092, 1984.
- Braschkat J, SO Gaertner, GA Reinhart (2003). Biogas versus other fuels: a comparative environmental assessment, in: *The Future of Biogas in Europe II*. Esbjerg. <<http://websvr5.sdu.dk/bio/konf.rapport.pdf>>
- Braun R, Madlener R, Laaber M, Efficiency evaluation of energy crop digestion plants. 7th FAO/SREN Workshop on “The Future of Biogas for Sustainable Energy Production in Europe”, Uppsala, Sweden, December 2005.
- Charnes A, Cooper WW, Rhodes E, Measuring the efficiency of decision making units, *European Journal of Operational Research*, 2(6), 429-444, 1978.
- Charnes A, Cooper WW, Golany B, Seiford L, Stutz J., Foundations of Data Envelopment Analysis for Pareto-Koopmans Efficient Empirical Production Functions, *Journal of Econometrics* (Netherlands) 30(1-2), 91-107, 1985.
- Coelli T, Rao DSP, and Battese GE, *An Introduction to Efficiency and Productivity Analysis*, Kluwer Academic Publishers, Boston/Dordrecht/London, 1998.
- Cooper WW, Seiford LM, Tone K (2000). *Data Envelopment Analysis: A Comprehensive Text with Models, Applications, References and DEA-Solver Software*, Kluwer Academic Publishers, Boston/Dordrecht/London.
- Cooper WW, Seiford LM, Zhu J (Eds.) (2004). *Handbook on Data Envelopment Analysis*, Kluwer Academic Publishers.
- Dias L, Mousseau V (2003). IRIS – Interactive Robustness Analysis and Parameters Interference for Multicriteria Sorting Problems (Version 2.0), INESC, Coimbra.
- Dias L, Mousseau V, Figureira J, Clímaco J, An aggregation/disaggregation approach to obtain robust conclusions with ELECTRE TRI. *European Journal of Operational Research* 138, 332-348, 2002.
- Farrel MJ, The measurement of productive efficiency, *Journal of the Royal Statistical Society, Series A*, General 120, 253-281, 1957.

- Färe R, Lovell, CAK (1978). Measuring the technical efficiency of production, *Journal of Economic Theory*, 19(1), 150-162.
- Färe R, Grosskopf S, Lovell CAK, and Pasurka C, Multilateral productivity comparisons when some outputs are undesirable: A nonparametric approach, *The Review of Economics and Statistics*, 71(1), 90-98, 1989.
- Green Electricity Act (2002). Bundesgesetz, mit dem Neuregelungen auf dem Gebiet der Elektrizitätserzeugung aus erneuerbaren Energieträgern und auf dem Gebiet der Wärme-Kraft-Kopplung erlassen werden (Ökostromgesetz), BGBl. I Nr. 149/2002 (accompanying ordinance: BGBl. II Nr. 508/2002), Vienna.
- Joro T, Viitala EJ, Weight-restricted DEA in action: from expert opinions to mathematical models. *Journal of the Operational Research Society* 55, 814-821, 2004.
- Madlener R, Productivity Benchmarking with Data Envelopment Analysis: Towards an Austrian Quality Standard for Biogas CHP Plants. Paper presented at the 29th IAEE International Conference, Potsdam, Germany, 7-10 June 2006.
- Madlener R, Markard J, Truffer B, Umbach-Daniel A, User-led vs policy-driven innovation processes: lessons from the development and diffusion of biogas plants in Austria and Switzerland, April 2006 (unpublished mimeo).
- Mousseau V, Slowinski R, Zielniewicz P, A user-oriented implementation of the ELECTRE-TRI method integrating preference elicitation support. *Computers & Operations Research* 27(7-8), 757-777, 2000.
- Podinovski VV, Production trade-offs and weight restrictions in data envelopment analysis, *Journal of the Operational Research Society* 55, 1311-1322, 2004.
- Reinhard S, Lovell CAK, and Thijssen GJ, Environmental efficiency with multiple environmentally detrimental variables; estimated with SFA and DEA, *European Journal of Operational Research*, 121, 287-303, 2000.
- Walla, C. (2005). Wirtschaftlichkeit von Biogasanlagen. Dissertation, Department für Wirtschafts- und Sozialwissenschaften, Institut für Agrar- und Forstökonomie, Universität für Bodenkultur Wien, Wien, Februar.
- Yu W, ELECTRE TRI: Aspects méthodologiques et manuel d'utilisation. Document du LAMSADE 80, Université Paris-Dauphine, 1992.