

CHOOSING WEIGHTS IN OPTIMAL SOLUTIONS FOR DEA-BCC MODELS BY MEANS OF A N-DIMENSIONAL SMOOTH FRONTIER

Flávia Badini Nacif

Production Engineering Master
Universidade Federal Fluminense (UFF)
Niterói – RJ
flavianacif@predialnet.com.br

João Carlos Correia Baptista Soares de Mello*

Production Engineering Department
Universidade Federal Fluminense (UFF)
Niterói – RJ
jcsmello@producao.uff.br

Lidia Angulo Meza

Materials Sciences Department
Universidade Federal Fluminense (UFF)
Volta Redonda – RJ
lidia@metal.ecimvr.uff.br

* *Corresponding author* / autor para quem as correspondências devem ser encaminhadas

Recebido em 08/2007; aceito em 10/2008
Received August 2007; accepted October 2008

Abstract

The DEA (Data Envelopment Analysis) smoothed frontier was introduced to solve the problem of multiple optimal solutions in the extreme efficient DMUs (Decision Making Units), which hinders the knowledge of the substitution rates (tradeoffs). It consists of changing the original frontier (piecewise linear) for a smoothed one, being as close as possible to the original one, and having continuous partial derivatives at every point. First, a solution was developed only for the BCC (Banker, Charnes and Cooper) model with either a single *input* or a single *output*. Then, it was generalized for the N-dimensional BCC model with simultaneous multiplicity of *inputs* and *outputs*, but limited by the fact that the polynomial of the *output* needs to be a linear one. The present article presents a general model, which not only expunges the limitations of the previous models but also includes them.

Keywords: DEA; smoothing; polynomial approaching.

Resumo

A suavização da fronteira DEA (Data Envelopment Analysis – Análise Envoltória de Dados) surgiu como uma solução do problema das múltiplas soluções ótimas nas DMUs (Decision Making Units – Unidades Tomadoras de Decisão) extremo-eficientes, o que impossibilita o conhecimento das razões de substituição (*tradeoffs*). Ela consiste na substituição da fronteira original (linear por partes) por outra suavizada, de modo que esta fronteira suavizada seja próxima da original, e que tenha derivadas contínuas em todos os pontos. Inicialmente foi desenvolvida solução apenas para o caso do modelo BCC (Banker, Charnes e Cooper) com apenas um *input*, ou apenas um *output*. Em seguida obteve-se uma generalização da solução para o caso BCC N-dimensional com multiplicidade simultânea dos *inputs* e dos *outputs*, porém com a limitação da linearidade do polinômio dos *outputs*. O presente artigo vem apresentar um modelo geral, que elimina as limitações dos modelos anteriores, e também os engloba.

Palavras-chave: DEA; suavização; aproximações polinomiais.

1. Introduction

Data Envelopment Analysis – DEA was developed by Charnes, Cooper & Rhodes (1978) to compute the efficiency of productive units (Decision Making Units – DMUs), whenever the financial viewpoint is not the dominant consideration. DEA methodology evaluates the efficiency of each DMU taking into account its resources (inputs) and the results it obtains (outputs).

Classic DEA models always present dual formulations as any mathematical programming model. There are, thus, two equivalent DEA formulations (Cooper *et al.*, 2000). To put it simply, one of the formulations (the Envelope model) defines a feasible production area and works with the distance of each DMU to the frontier of that area. The other formulation (the Multipliers model) works with the ratio of weighted sums of products and resources. The weighing factors are the most favourable that can be chosen for each DMU under certain circumstances.

Being dual problems, the two formulations will obviously provide the same efficiency for each DMU. Furthermore, other information beyond efficiency can be obtained from the two models. The envelope model provides weight coefficients for each efficient DMU (designated as λ_i) to create a virtual DMU that will be used as benchmark for each inefficient DMU. For any given orientation, these coefficients are computed in single given way for each DMU. It is of particular interest to see what happens in the efficient DMUs: they are their own benchmark and, so, the envelope LPP (Linear Programming Problem) returns 1 for λ referring to that DMU and nought for all others. A highly degenerated LPP obtains thus.

The multipliers model provides the weighing coefficients that each DMU will allocate to each input or output. The fact that each DMU provides different values for these multipliers is the very essence of DEA. Each DMU is free to value better whatever it is best at, and to ignore the variables in which it does not perform well. All and every DEA model should preserve this freedom in greater or smaller measure.

Beside this freedom for each DMU to choose its own weights, certain DMUs may choose more than a unique set of weights. This can cause problems and the aim of this paper is to propose a way of choosing a unique set of multipliers for all DMUs.

This paper is organised thus: Section 2 presents a justification for this study. A bibliographic review is presented in Section 3. The general theory of smoothing the DEA frontier is presented in Section 4. The generalisation of the BCC smoothing model and the original contribution of this study are presented in section 5. A numerical example is given in section 6 and, finally, section 7 presents the conclusions of this paper.

2. Reasons for this study

The need to know the value of the multipliers (also known as weights) stems from the economical interpretations associated to them. The commonest of these interpretations is that each weight is an indicator of the importance each DMU gives to its value to determine its efficiency. The validity of this interpretation, without further calculations, requires the variables to be previously standardised.

From an economic viewpoint, multipliers may have two interpretations. The first is as components of the weight ratio between variables (tradeoffs), meaning how much should a

DMU increase a given input when another decreases by one unit or how much does the increase in a given input increase an output. A second and more important interpretation is that weights are standardised shadow prices (Coelli *et al.*, 1998). This is useful to determine the prices of quantities that have no market value; one only needs to know the effective cost of some of the variables. Knowing this, shadow prices cease to be standardised. Reinhard *et al.* (2000), for instance, have used this approach to calculate the price of pollution in a study on environmental efficiency. One ought to emphasise once more that there is for each DMU a different set of prices, which means that prices calculated for each DMU are valid only for that DMU.

The practical use of the multipliers interpretation hits a serious snag that is inherent to the multipliers model LPP. The complementary slack theorem may also mean that multipliers are the coefficients of an equation that defines a hyperplane tangent to the frontier on the DMU projection point (Cooper *et al.*, 2000). As efficient DMUs (more exactly, the extreme efficient DMUs) are vertices, where there are neither derivatives nor tangent hyperplanes but an infinity of support hyperplanes, there is then an infinity of multiplier sets for each extreme efficient hyperplane all of them leading to efficiency 1 for those DMUs. Therefore, besides each DMU having the freedom to choose its own weights, which is desirable, it is impossible for the benchmark DMUs, performers of good management practices, to know the weights they have really allocated to each variable. To determine the importance of each input or output or even to calculate the shadow weights becomes compromised when dealing with extreme efficient DMUs.

The lack of unique values for extreme efficient DMUs' weights has further different consequences. From a theoretical point of view, it prevents the calculation of directional derivatives along the whole frontier. From the practical point of view, it is an obstacle against the use of DEA as an auxiliary tool in multicriteria problems. Some circumstances in multicriteria problems may render it desirable that weights be allocated independently of the decision maker's judgement as, for instance, when several decision makers do not agree on them. DEA would be an excellent tool for that if it were not that the weights allocated by some DMUs are not known.

3. Bibliographic review

With greater or lesser sophistication, available literature includes several approaches to deal with the problem of more than one set of multipliers. If the numbers of extreme efficient DMUs is small when compared with its total number, it is obvious the weights allocated by the extreme efficient DMUs can be ignored and work can proceed with the weights allocated by the remaining DMUs alone (as done by Lins *et al.*, 2003; Soares de Mello *et al.*, 2008). The problem of more than one set of multipliers for the extreme efficient DMUs has been approached on several occasions, but their solutions have been unsatisfactory. Charnes *et al.* (1985) had already accepted the problem existed when they proposed the arbitrary use of a single value for derivatives obtained from the computation of an average weight based on the barycentres of the concurring hyper-surfaces. This method has several disadvantages. One is the need to know the equations for all the faces, which requires an intense load of computer work (Dulá, 2002). Another is the sudden variation owing to derivative discontinuities. Finally, it is not applicable to either the DMUs at the start of a Pareto inefficient region or those that are adjacent to an incomplete dimension face (Olesen & Petersen, 1996). Cooper *et al.* (2007) have proposed to refine the method put forward by Charnes *et al.* (1985).

Some of these drawbacks may have caused the method not to be commonly used. Several authors simply ignore the problem. Occasionally, they mention it and then proceed with the solutions found for the first optimal solution by the algorithm they have used (Thanassoulis, 1993; Chilingirian, 1995).

In some specific circumstances, partial solutions are possible for the problem of more than one set of multipliers. Doyle & Green (1994), for instance, needed to calculate a unique value for each DMU's multipliers vector in their crossed evaluation model. They did that by solving lexicographically the traditional DEA model and one of two other models, which the authors named "aggressive" or "benevolent". These auxiliary models determine the multipliers to be used in the cross evaluation model. There are other ways to calculate these multipliers for the cross evaluation model as proposed by Liang *et al.* (2008). An alternative way to calculate multipliers can be found in a practical application by Soares de Mello *et al.* (2009).

On the other hand, the super efficiency model (Andersen & Petersen, 1993) does not suffer from having more than one set of multipliers because of its formulation. However, it has other disadvantages. To start with, it does not limit efficiencies to the interval $[0,1]$. Furthermore, it eliminates a different constraint for each DMU LPP, which means that the efficiency frontier has a different shape depending on which DMU is being studied.

According to Rosen *et al.* (1998), the multipliers values can vary between the value based on the derivative at left and that on the right. These authors state that it is impossible to get round this multiplicity of values and propose a modified SIMPLEX table to calculate the multipliers variation limits.

The impossibility referred to by Rosen *et al.* (1998) arises from the linear nature of some segments of the DEA frontier. Soares de Mello *et al.* (2002, 2004) have shown that it is possible to get around this impossibility by substituting a new frontier with similar properties, but with derivatives at all points, for the original DEA frontier. Among the identical properties, the allocation of unitary efficiencies to the original DEA model extreme efficient DMUs is included. This technique is described here in general terms and exemplified for two-dimensional cases. It consists of smoothing the original DEA frontier while respecting the DEA Basic properties: convexity, throughput monotonicity (outputs growing together with inputs), the same efficient DMUs and allocation of different weights by each DMU. An application of this technique can be found in Gomes *et al.* (2004).

Another type of smoothed frontier appears in the Hyperbolic and Spherical DEA models (Kozyreff Filho & Millionni, 2004; Avellar *et al.*, 2004; Gomes & Avellar, 2005; Avellar *et al.*, 2005; Avellar *et al.*, 2007). These models have a different objective to that studied herein. In those articles, the authors deal simultaneously with uniformity and smoothing of the frontier. The results obtained by them can be classed as a smoothed variant of the Zero Sum Gains DEA model (Lins *et al.*, 2003; Gomes & Lins, 2008).

4. Smoothing the DEA frontier

A pseudo-metric topology was used in Soares de Mello *et al.* (2002, 2004) to smooth the two and three-dimensional DEA frontiers. The pseudo-metric used here measures both how close the smoothed function is to the original frontier and that of their derivatives where they exist.

In the two-dimensional case, it was proposed that the difference in the arc length of the function's curve between two points was taken as the measure of proximity of the two

functions as well as of their derivatives. The frontier between two consecutive efficient DMUs is a straight line. This is the shortest length between two points and, therefore, any other curve that connects those two points will have a greater arc length. The arc length will be an increasing function of the divergences from that curve to the straight line. Therefore, this arc length is able to measure the proximity of the smoothed and the original frontiers. It may also measure the proximity between the frontier derivatives as shown in Soares de Mello *et al.* (2002). It is well known that for any given curve $y = f(x)$, the length (IL) of an arc is given by (1).

$$L = \int_{x_1}^{x_2} \sqrt{1 + \left(\frac{dy}{dx}\right)^2} dx \tag{1}$$

where x_1 and x_2 are the minimum and maximal input values.

The same arguments can be generalised for higher dimension problems if a hyper-plane is substituted for the straight line and a multiple integral for the simple one.

4.1 General Formulation of the Smoothing Model

For a single input DEA models, smoothing is but looking for a function that minimises the curve's arc length (or its n-dimension generalization), that contains the Pareto efficient DMUs and that have partial derivatives in every point. For computational ease, the square of the curve's arc length can be minimized with no change of the result. Then, after finding the extreme efficient DMUs in the DEA classic model, smoothing can be obtained through the variational problem (2).

$$\begin{aligned} \text{Min } L &= \int_R \left[1 + \sum_i \left(\frac{\partial F}{\partial x_i} \right)^2 \right] dS \\ \text{subject to} & \\ F(\bar{X}_j) &= \text{output}(\bar{X}_j), \forall \bar{X}_j \in E = \{ \bar{X} : \bar{X} \text{ is Pareto efficient DMU} \} \\ \forall \bar{X}_j, \exists & \frac{\partial F}{\partial x_i} \end{aligned} \tag{2}$$

In this model, F is a function of the inputs in the output and x_i is the input vector.

This is a BCC model (Banker *et al.*, 1984) and, so, it requires the frontier to be convex, i. e., $\frac{\partial^2 F}{\partial x_i^2} \leq 0$. With this additional constraint, the No Optimal Solution Theorem is obtained. Its

proof is in Soares de Mello *et al.* (2002). It guarantees there is no “closest solution” although there are close enough solutions that can be used. This calculation needs a similar approach to that of the Finite Elements method (Reddy, 1993).

At the beginning, when the two-dimensional case was first dealt with, Soares de Mello *et al.* (2002) did use an approximation for each frontier area. Owing to the two-dimension geometry, it was possible to calculate a single optimal approximate function for each side. However, for a larger number of dimensions, this becomes a computationally intensive problem (Dulá, 2002).

Later, Soares de Mello *et al.* (2004) dealt with the frontier-smoothing problem for a larger number of dimensions and one single input or output. As opposed to the two dimensions model, a single approximate polynomial was used for the whole frontier. Higher degree polynomial functions were used for that purpose.

The authors have shown that in the particular case of two inputs and one output the lowest polynomial degree can be obtained as a function of the number of extreme efficient DMUs. This function is described in table 1. This table was arrived at so the number of decision variables is larger than the number of equality constraints. As these variables are the coefficients of the approximate polynomial, its degree will depend on the number of decision variables.

Table 1 – Finding the polynomial degree.

Number of extreme efficient DMUs	Polynomial degree
3 – 5	2
6 – 9	3
10 – 14	4
15 – 20	5
21 – 27	6
...	...

The approximate polynomial for a BCC DEA problem with two inputs, a single output and three efficient DMUs, for instance, can be:

$$Z = a + bx + cy + dx^2 + exy + fy^2 + gx^3 + hx^2y + ixy^2 + jy^3 + kx^4 + lx^3y + mx^2y^2 + nxy^3 + oy^4 \tag{3}$$

In this model, Z stands for the output and x and y for the inputs.

The DEA BCC smoothing model with two inputs and one output is shown in (4).

$$\begin{aligned} & \text{Min} \left\{ \int_{x_{\min}}^{x_{\max}} \int_{y_{\min}}^{y_{\max}} \left[1 + \left(\frac{\partial Z}{\partial x} \right)^2 + \left(\frac{\partial Z}{\partial y} \right)^2 \right] dydx \right\} \\ & \text{subject to} \\ & Z(x_{efi}, y_{efi}) = Z_{efi} \\ & \frac{\partial Z}{\partial x}(x_{\max}, y_{\max}) \geq 0 \\ & \frac{\partial Z}{\partial y}(x_{\max}, y_{\max}) \geq 0 \\ & \frac{\partial^2 Z}{\partial x^2} \leq 0, \forall x, y \\ & \frac{\partial^2 Z}{\partial y^2} \leq 0, \forall x, y \end{aligned} \tag{4}$$

In this model, the variables are the same as before.

The objective function ensures that the smoothed frontier and its derivatives are close to the original frontier. The first constraint ensures that the smoothed frontier includes the same efficient DMUs of the original frontier. The constraints that include first derivatives ensure that the output is an increasing function of the inputs. Finally, the constraints that include second-degree derivatives ensure the frontier is convex.

Model (4) becomes model (5) if it is taken into account that Z is a polynomial function of x and y.

$$\begin{aligned}
 & \text{Min} \left\{ \int_{x_{\min}}^{x_{\max}} \int_{y_{\min}}^{y_{\max}} \left[1 + \left(\frac{\partial Z}{\partial x} \right)^2 + \left(\frac{\partial Z}{\partial y} \right)^2 \right] dy dx \right\} \\
 & \text{subject to} \\
 & Z(x_{efi}, y_{efi}) = Z_{efi} \\
 & \frac{\partial Z}{\partial x}(x_{\max}, y_{\max}) \geq 0 \\
 & \frac{\partial Z}{\partial y}(x_{\max}, y_{\max}) \geq 0 \\
 & d, f, g, h, i, \dots \leq 0
 \end{aligned} \tag{5}$$

In this model, the variables Z, x and y stand for respectively the output and the inputs. Besides, x_{efin} , y_{efin} represent the inputs of any extreme efficient DMU, Z_{efin} represents the output of the same extreme efficient DMU n. Variables d, f, g, h, and i are decision variables.

Once again, it should be noted that the function Z is polynomial; the result of the double integral in the objective function is a quadratic function of the polynomial coefficients. As the constraints are linear, the smoothing problem becomes a quadratic programming problem.

Figure 1 shows the smoothed frontier obtained by Soares de Mello *et al.* (2004) for two inputs and one single output.

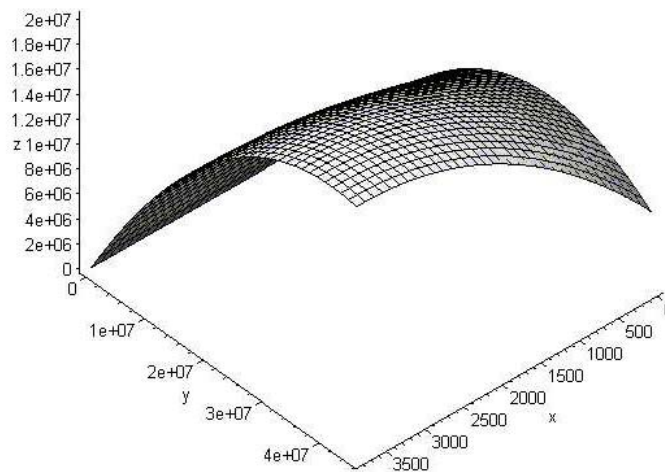


Figure 1 – Smoothed Tridimensional BCC Frontier.

For the case of a single output and n inputs model (6) obtains.

$$\min \left\{ \int_{x_{1\min}}^{x_{1\max}} \dots \int_{x_{n\min}}^{x_{n\max}} \left[1 + \sum_{i=1}^n \left(\frac{\partial F}{\partial x_i} \right)^2 \right] dx_n \dots dx_1 \right\}$$

subject to

$$z_{1\text{efi}} = F(x_{1\text{efi}} \dots x_{n\text{efi}}), \quad \forall \text{ extreme-efficient DMU} \tag{6}$$

$$\frac{\partial F}{\partial x_i}(x_{1\max}, \dots, x_{n\max}) \geq 0, \quad \forall i = 1..n$$

$$\frac{\partial^2 F}{\partial x_i^2} \leq 0, \quad \forall i = 1..n$$

There was a change of notation for this model caused by the larger number of variables involved. For the next models, with at least more than two inputs, the notation will be the same. So, x_1, \dots, x_n are the DMU's inputs; Z will be the DMU's output.

For the inverse case of a single input and m outputs, model (7) obtains.

$$\min \left\{ \int_{z_{1\min}}^{z_{1\max}} \dots \int_{z_{m\min}}^{z_{m\max}} \left[1 + \sum_{i=1}^m \left(\frac{\partial H}{\partial z_i} \right)^2 \right] dz_m \dots dz_1 \right\}$$

subject to

$$x_{1\text{efi}} = H(z_{1\text{efi}} \dots z_{m\text{efi}}), \quad \forall \text{ extreme-efficient DMU} \tag{7}$$

$$\frac{\partial H}{\partial z_i}(z_{1\min}, \dots, z_{n\min}) \geq 0, \quad \forall i = 1..m$$

$$\frac{\partial^2 H}{\partial z_i^2} \geq 0, \quad \forall i = 1..m$$

The second derivative is now positive. In this case, the input is a function of the outputs so the theorem of the inverse function ensures that this constraint makes the frontier to be convex.

5. Generalization of the BCC smoothing model

To deal simultaneously with both output and input multiplicity in the smoothing problem, a U function is defined as follows:

$$U(x_1 \dots x_n, z_1 \dots z_m) = F(x_1 \dots x_n) - H(z_1 \dots z_m) \tag{8}$$

When function U equals a constant, equation (8) represents a level. Its multidimensional generalization of the arc length ought to be minimised. As opposed to the model previously shown the new solution will obtain taking into account the differentiation of all inputs relative to all outputs and vice-versa.

$$\min \left\{ \int_{x_{1min}}^{x_{1max}} \dots \int_{x_{nmin}}^{x_{nmax}} \int_{z_{1min}}^{z_{1max}} \dots \int_{z_{mmin}}^{z_{mmax}} \left[I + \sum_{i=1}^n \left(\frac{\partial U}{\partial x_i} \right)^2 + \sum_{i=1}^m \left(\frac{\partial U}{\partial z_i} \right)^2 \right] dz_m \dots dz_1 dx_n \dots dx_1 \right\}$$

subject to

$$U(x_{1efi} \dots x_{nefi}, z_{1efi} \dots z_{mefi}) = 0, \forall \text{ extreme-efficient DMU} \quad (9.1)$$

$$\frac{\partial U}{\partial x_i}(x_{1max}, \dots, x_{nmax}) \geq 0, \forall i = 1..n \quad (9.2)$$

$$\frac{\partial U}{\partial z_i}(z_{1min}, \dots, z_{nmin}) \leq 0, \forall i = 1..m \quad (9.3) \quad (9)$$

$$\frac{\partial^2 U}{\partial x_i^2} \leq 0, \forall i = 1..n \quad (9.4)$$

$$\frac{\partial^2 U}{\partial z_i^2} \leq 0, \forall i = 1..m \quad (9.5)$$

Formulation (9) represents the general DEA model smoothed for any number of input and output variables. The objective function is given by the (n+m)-uple in which $x_{i min}$ and $x_{i max}$ indicate the minimum and maximum value for each input, whereas $z_{i min}$ and $z_{i max}$ do the same for each output. Constraint (9.1) ensures that DMUs extreme efficient be contained in the smoothed frontier. Constraint (9.2) ensures the frontier’s increasing monotonicity where inputs hit their maxima, whereas constraint (9.3) does the same for decreasing monotonicity when outputs hit their minima. Convexity obtains from (9.4) and (9.5) respectively for inputs and outputs. Note should be taken that, as in (7), it should always be $\frac{\partial^2 H}{\partial z_i^2} \geq 0$. As $U = F - H$, the minus sign of this equation leads to constraint (9.5).

Attention is drawn for a fact relative to the convexity of the smoothed frontier (9.4) and (9.5). Similarly to the single input or output model (Soares de Mello, 2002), it will not be possible to keep the signal constraint of the second derivatives at all points. The truth is that this will be only possible for polynomials up to the second degree. In this case, the second derivative for any input or output will be a constant as it represents a parabola with a constant convexity throughout. Therefore, for the remaining cases, a stronger constraint must be imposed: the coefficient of the polynomial term for which the second derivative is not nil, should be:

- Lesser or equal to zero if it’s a coefficient of the input polynomial;
- Greater or equal to zero if it’s a coefficient of the output polynomial.

Finally, the equation of the smoothed frontier will be as follows:

$$U(x_1 \dots x_n, z_1 \dots z_m) = 0 \quad (10)$$

5.1 Model Properties

A demonstration will now follow that the proposed model has all required conditions to be a smoothed BCC frontier: it must contain all the extreme efficient DMUs, be convex and outputs must be increasing functions of the inputs.

Constraint (9.1) ensures that the frontier contains all the extreme efficient DMUs because it is equivalent to “the equality of functions F and H calculated at each of the extreme efficient DMUs values.

Next, it will be shown that in this model, outputs are increasing functions of the inputs, i.e., the model (9) constraints lead to (11).

$$\frac{\partial z_i}{\partial x_j} \geq 0, \forall i = 1..m, j = 1..n \tag{11}$$

For that purpose, the Implicit Function Theorem will be used. It allows the expression $\frac{\partial z_i}{\partial x_j}$ to be calculated from the partial derivatives of function U (8):

$$\frac{\partial z_i}{\partial x_j} = - \frac{\left(\frac{\partial U}{\partial x_j} \right)}{\left(\frac{\partial U}{\partial z_i} \right)} \tag{12}$$

From constraints (9.2) and (9.3) the numerator of expression (12) must be non-negative, its denominator non-positive. This leads to expression (13), valid for the whole frontier with the possible exception of a finite number of points where the derivative may not exist.

$$\frac{\partial z_i}{\partial x_j} \geq 0 \tag{13}$$

As these isolated points do not affect monotonicity, it is demonstrated that in the smoothed frontier, outputs are increasing functions of the inputs.

Lastly, to prove that the BCC frontier is convex the Convexity expression in (14) must be checked:

$$\frac{\partial^2 z_i}{\partial x_j^2} \leq 0, \forall i = 1..m, j = 1..n \tag{14}$$

Having recourse to the Implicit Function Theorem, the Chain Rule, and the Product Derivation, expression (15) obtains:

$$\frac{\partial^2 z_i}{\partial x_j^2} = - \left(\frac{\frac{\partial^2 U}{\partial x_j^2} \times \frac{\partial U}{\partial z_i} - \frac{\partial U}{\partial x_j} \times \frac{\partial^2 U}{\partial x_j \partial z_i}}{\left(\frac{\partial U}{\partial z_i} \right)^2} \right) \tag{15}$$

As the polynomials are independent, i.e., there is no term in U that multiplies x and z , function (16) is valid.

$$\frac{\partial^2 U}{\partial z_i \partial x_j} = 0 \tag{16}$$

So, equation (15) can be simplified becoming equation (17).

$$\frac{\partial^2 z_i}{\partial x_j^2} = - \frac{\left(\frac{\partial^2 U}{\partial x_j^2} \times \frac{\partial U}{\partial z_i} \right)}{\left(\frac{\partial U}{\partial z_i} \right)^2} \tag{17}$$

From model constraints (9.1) to (9.5), the two numerator factors are always non-positive while the denominator is non-negative. This leads to expression (18), valid for the whole frontier with the possible exception of a finite number of points where the derivative may not exist.

$$\frac{\partial^2 z_i}{\partial x_j^2} \leq 0 \tag{18}$$

As these isolated points do not affect convexity, it is demonstrated that the smoothed BCC frontier has all the desired properties.

It is also easy to show through similar calculations that the one output/several inputs and vice versa cases dealt with in Soares de Mello *et al.* (2004) are but particular cases of this general model.

5.2 Finding the polynomial degree

As there is a polynomial for inputs and another for outputs, degrees must be found for both. To find those values it should taken into account that:

- Both the input and output polynomials must have the same degree (called *g*) so symmetry is maintained when smoothing.
- One of the coefficients can be deleted because the smoothing model equality constraint will remain true if it is divided by the coefficient's value. This property allows us to establish the convention, with no loss of generalization, that the polynomial independent term will be 1. With this in mind, function U is defined by (19).

$$U = F_{g,n} - H_{g,m} = 1 + \sum_k a_k x_1^{ci_{1k}} x_2^{ci_{2k}} \dots x_n^{ci_{nk}} - \sum_k b_k z_1^{co_{1k}} z_2^{co_{2k}} \dots z_m^{co_{mk}} \quad \forall ci_{ik}, co_{ik} \text{ int} \tag{19}$$

such as it satisfies (20)

$$\sum_{i=1}^n ci_{ik} \leq g, \quad \sum_{i=1}^m co_{ik} \leq g \tag{20}$$

- The total number of coefficients, or decision variables, is the sum of the number of coefficients in output and input polynomials.
- The number of output and input variables, respectively *m* and *n*, are provided by the real problem.

Table 2 can now be built. From the number of variables and the number of polynomial terms, the polynomial degree is determined for both inputs and outputs. The table should be used twice: for inputs and outputs.

Table 2 – Polynomial degrees in function of number of variables and their terms.

		Nvar (number of variables)				
		1	2	3	4	5
g (polynomial degree)	1	1	2	3	4	5
	2	2	5	9	14	20
	3	3	9	19	34	55
	4	4	14	34	69	125
	5	5	20	55	125	251

This is a symmetrical matrix, meaning the number of polynomial coefficients is the same for either degree g and m variables or vice-versa.

If no and ni are respectively, the number of outputs and inputs, condition (21) has to apply.

$$N^\circ \text{ of extreme-efficient DMUs} < N^\circ \text{ of coefficients} = no + ni \tag{21}$$

This expression ensures that the number of decision variables is greater than the number of equality constraints in the smoothing problem.

Degree g will be found on the line where the minimum sum for $no + ni$ that satisfies the above inequality is found.

Take an example: with two inputs, three outputs and sixteen extreme efficient DMUs, there must be at least sixteen model decision variables. For two inputs and three outputs, columns 2 and 3 of table 2 must be checked. Then, the corresponding lines in table 2 determine the polynomial degree. Line 1 corresponds to a polynomial of one degree and shows five model decision variables, which is insufficient. Line 2 gives a different reading: two inputs correspond to five coefficients and three outputs nine coefficients. Their sum of fourteen is still insufficient. If polynomials are of third degree, the corresponding line 3 will show nine coefficients for two inputs and nineteen for three outputs. Their sum of twenty-eight is more than enough to avoid equality constraints to generate unfeasibility. Therefore, this example requires degree 3 for both input and output polynomials.

5.3 Finding multipliers

From the smoothed frontier equation:

$$F(x_1 \dots x_n) - H(z_1 \dots z_m) = 0 \tag{22}$$

multipliers for each input and output can be obtained.

To start with, the hyperplane tangent to the smoothed frontier on a point correspondent to an extreme efficient DMU must be determined. Let $(x_1, \dots, x_n, z_1, \dots, z_m)$ be a generic point of that hyperplane.

Take

$$G(x_{1o}, \dots, x_{no}, z_{1o}, \dots, z_{mo}) = F(x_{1o}, \dots, x_{no}) - H(z_{1o}, \dots, z_{mo}) \tag{23}$$

where DMU o is an extreme efficient DMU.

The tangent hyperplane equation is the internal product

$$\left[\bar{V} G(x_{1o}, \dots, x_{no}, z_{1o}, \dots, z_{mo}) \right] \cdot (x_1 - x_{1o}, x_2 - x_{2o}, \dots, z_1 - z_{1o}, z_2 - z_{2o}, \dots) = 0 \quad (24)$$

On the other hand, from the DEA theory the support plane equation on DMU o is (Lins & Angulo-Meza, 2000):

$$\sum u_{jo}^* z_j - \sum v_{io}^* x_i = u_* \quad (25)$$

Should it exist, the support plane is also a tangent plane. As the smoothed curve is designed to ensure that tangent hyperplanes do exist, the multipliers values are obtained from the equivalence of equations (24) and (25) from the constraint that the inputs weighted sum must be one.

For notation's simplicity sake, further development will follow a two input, two output case that can be easily generalised.

The equations to be taken into account are:

$$u_{1o}^* z_1 + u_{2o}^* z_2 - v_{1o}^* x_1 - v_{2o}^* x_2 = u_* \quad (26)$$

$$(x_1 - x_{1o}) \frac{\partial H}{\partial x_1} + (x_2 - x_{2o}) \frac{\partial H}{\partial x_2} - (z_1 - z_{1o}) \frac{\partial F}{\partial z_1} - (z_2 - z_{2o}) \frac{\partial F}{\partial z_2} = 0 \quad (27)$$

$$x_{1o} v_{1o} + x_{2o} v_{2o} = 1 \quad (28)$$

If expressions (26) and (27) are compared term-to-term, multipliers values that do not obey equation (28) are obtained. For that term-to-term comparison, please note that equation (27) is equivalent to equation (29):

$$x_1 \frac{\partial H}{\partial x_1} + x_2 \frac{\partial H}{\partial x_2} - z_1 \frac{\partial F}{\partial z_1} - z_2 \frac{\partial F}{\partial z_2} = x_{1o} \frac{\partial H}{\partial x_1} + x_{2o} \frac{\partial H}{\partial x_2} - z_{1o} \frac{\partial F}{\partial z_1} - z_{2o} \frac{\partial F}{\partial z_2} \quad (29)$$

To solve the compatibility problem with equation (28) it should be noted that equation (29) is equivalent to equation (30) for any value of α that is not nil.

$$\alpha x_1 \frac{\partial H}{\partial x_1} + \alpha x_2 \frac{\partial H}{\partial x_2} - \alpha z_1 \frac{\partial F}{\partial z_1} - \alpha z_2 \frac{\partial F}{\partial z_2} = \alpha x_{1o} \frac{\partial H}{\partial x_1} + \alpha x_{2o} \frac{\partial H}{\partial x_2} - \alpha z_{1o} \frac{\partial F}{\partial z_1} - \alpha z_{2o} \frac{\partial F}{\partial z_2} \quad (30)$$

Term-to-term comparison between equations (30) and (26) constrained by equation (28) leads to:

$$\alpha = - \frac{1}{x_{1o} \frac{\partial H}{\partial x_1} + x_{2o} \frac{\partial H}{\partial x_2}} \quad (31)$$

and to multipliers:

$$v_{1o} = \frac{\frac{\partial H}{\partial x_1}}{x_{1o} \frac{\partial H}{\partial x_1} + x_{2o} \frac{\partial H}{\partial x_2}} \quad (32)$$

$$v_{2o} = \frac{\frac{\partial H}{\partial x_2}}{x_{1o} \frac{\partial H}{\partial x_1} + x_{2o} \frac{\partial H}{\partial x_2}} \quad (33)$$

$$u_{1o} = \frac{\frac{\partial F}{\partial z_1}}{x_{1o} \frac{\partial H}{\partial x_1} + x_{2o} \frac{\partial H}{\partial x_2}} \quad (34)$$

$$u_{2o} = \frac{\frac{\partial F}{\partial z_2}}{x_{1o} \frac{\partial H}{\partial x_1} + x_{2o} \frac{\partial H}{\partial x_2}} \quad (35)$$

$$u_* = - \frac{x_{1o} \frac{\partial H}{\partial x_1} + x_{2o} \frac{\partial H}{\partial x_2} + z_{1o} \frac{\partial F}{\partial z_1} + z_{2o} \frac{\partial F}{\partial z_2}}{x_{1o} \frac{\partial H}{\partial x_1} + x_{2o} \frac{\partial H}{\partial x_2}} \quad (36)$$

5.4 Calculation of the relative importance of each variable

The multipliers calculated hereabove might not have any meaning for the decision maker as they simultaneously incorporate scale effects and subjective importance. A more meaningful quantity is the relative importance of each input (output) to establish the virtual input (output). This can be used to increase discrimination in DEA (Angulo-Meza & Lins, 2002; Adler *et al.*, 2002).

The importance of any input i is given by (37).

$$Im p(x_i) = \frac{v_{io} x_i}{\sum_i v_{io} x_i} \quad (37)$$

Therefore, entering the multipliers' values in expression (37), the importance of input i is given by expression (38).

$$Im p(x_i) = \frac{x_i \frac{\partial H}{\partial x_i}}{\sum_i x_i \frac{\partial H}{\partial x_i}} \quad (38)$$

Similar calculations lead to the importance of output j being given by expression (39)

$$Im p(z_j) = - \frac{z_j \frac{\partial F}{\partial z_j}}{\sum_i x_i \frac{\partial H}{\partial x_i}} \quad (39)$$

6. Numerical example

The theoretical concepts hereabove will be now exemplified. The data are found on Table 3. X_1 and X_2 are inputs; Z_1 and Z_2 are outputs. These data were obtained from Nacif (2005).

Table 3 – Data for the numerical example.

DMUs	X_1	X_2	Z_1	Z_2
A	680	1	50	120
B	1500	1	50	240
C	610	2	50	130
D	860	2	90	250
E	1400	2	90	500
F	220	6	75	220
G	290	6	75	1,5
H	1800	0	100	130
I	2800	0	100	260
J	4000	0	100	510

Using the SIAD software (Angulo-Meza *et al.*, 2005), the extreme efficient DMUs are found to be A, D, E, F, H and J. Only the data for these DMUs are needed to calculate the smoothed frontier.

To determine the input and output polynomial degree account must be taken that there are six extreme efficient DMUs, two inputs, and two outputs. From table 2, only the column corresponding to two variables for both input and output polynomials should be taken into account. Table 4 is then built from the second column of Table 2.

Table 4 – Number of coefficients for each degree g for two inputs and two outputs.

		Number of coefficients for the input polynomial	Number of coefficients for the output polynomial	Number of coefficients for the joint polynomial
g (input and output polynomials degree)	1	2	2	4
	2	5	5	10
	3	9	9	18
	4	14	14	28
	5	20	20	40

The number of extreme efficient DMUs should be lower than the total number of polynomial coefficients to avoid unfeasibility. Thus, Table 5, that relates the number of extreme efficient DMUs with the polynomial degree, is built. To do it, it must be noted that the number of extreme efficient DMUs must be lower than that shown on the last column of Table 4.

Table 5 – Polynomials degrees versus extreme efficient DMUs.

Nº of extreme efficient DMUs	Degree of each polynomial
1-3	1
4-9	2
10-17	3
18-27	4
28-39	5

Both the input and output polynomials should be second degree, for the proposed problem.

Function U shall be thus of the following type:

$$U(x_1, x_2, z_1, z_2) = 1 + ax_1^2 + bx_1 + cx_1x_2 + dx_2 + ex_2^2 - fz_1^2 - gz_1 - hz_1z_2 - iz_2 - jz_2^2 \quad (40)$$

whereas the smoothing equation will be:

$$U(x_1, x_2, z_1, z_2) = 0 \quad (41)$$

After calculating the input and output values as well as function U, the above smoothing equation (40) becomes model (42).

$$Min \left\{ \int_{220}^{4000} \int_0^6 \int_{50}^{100} \int_{1.5}^{510} \left[1 + (2ax_1 + b + cx_2)^2 + (cx_1 + d + 2ex_2)^2 + (-2fz_1 - g - hz_2)^2 \right] dz_2 dz_1 dx_2 dx_1 \right\}$$

subject to

$$1 + 680^2.a + 680.b + 680.1.c + 1.d + 1^2.e - 50^2.f - 50.g - 50.120.h - 120.i - 120^2.j = 0$$

$$1 + 860^2.a + 860.b + 860.2.c + 2.d + 2^2.e - 90^2.f - 90.g - 90.250.h - 250.i - 250^2.j = 0$$

$$1 + 1400^2.a + 1400.b + 1400.2.c + 2.d + 2^2.e - 90^2.f - 90.g - 90.500.h - 500.i - 500^2.j = 0$$

$$1 + 220^2.a + 220.b + 220.6.c + 6.d + 6^2.e - 75^2.f - 75.g - 75.220.h - 220.i - 220^2.j = 0$$

$$1 + 1800^2.a + 1800.b + 1800.0.c + 0.d + 0^2.e - 100^2.f - 100.g - 100.130.h - 130.i - 130^2.j = 0$$

$$1 + 4000^2.a + 4000.b + 4000.0.c + 0.d + 0^2.e - 100^2.f - 100.g - 100.510.h - 510.i - 510^2.j = 0$$

$$2.4000.a + b + 6.c \geq 0$$

$$4000.c + d + 2.6.c \geq 0$$

$$-2.100.f - g - 510.h \leq 0$$

$$-100.h - i - 2.510.j \leq 0$$

$$2.a \leq 0$$

$$2.c \leq 0$$

$$-2.f \leq 0$$

$$-2.j \leq 0$$

(42)

Calculating the decision variable values and placing them on the smoothing equation, equation (43) obtains.

$$1000 - 0,00008x_1^2 + 0,39109x_1 + 0,04183x_1x_2 + 79,71422x_2 - 14,99935z_1 + 0,10645z_1z_2 - 10,21461z_2 \quad (43)$$

This is the smoothing equation for the given problem. It represents a hyper-surface in \mathbb{R}^4 with derivatives on all points that replaces the original DEA frontier, which was piecewise linear. It should be noted that some quadratic terms have disappeared because their coefficients were nil.

7. Conclusions

A methodology to solve the unique weights calculation problem for DEA BCC models with simultaneous multiplicity of inputs and outputs has been proposed. The methodology is an extension of the one developed by Soares de Mello *et al.* (2004) based on the replacement of the original DEA frontier (piecewise linear) by another smoothed frontier (with continuous derivatives). The smoothed frontier will be represented by a polynomial equation bringing together the input and output polynomials.

A model was initially developed by Nacif *et al.* (2004) that dealt with the n-dimensional but restrained the output polynomial to be linear. Together with the model, the algorithm to determine the polynomial degree was generalised so it could be calculated for all cases.

To generalise smoothing even further, a new model, called General n-Dimensional Smoothing Model (model 9) was proposed. This model was gathered all those previously proposed as shown in section 3.1. For shortness sake, calculation details have been omitted; they can be seen in Nacif (2005). This new generalisation is needed to avoid asymmetric handling of inputs and outputs. In the previous model, either the input or the output polynomial had to be arbitrarily made linear. Either option would change completely the problem and, therefore, would not lead to a unique value for each multiplier, the main target of this paper.

The smoothing model proposed here has more algebraic calculi than the n-dimensional model by Soares de Mello *et al.* (2004). It will be recalled that the latter model had no simultaneous input and output multiplicity. The number of algebraic calculi increases, in the model presented here, as the number of inputs, outputs, and extreme efficient DMUs increases. This becomes apparent when tables 1 and 2 are compared and makes for the model high complexity when a large number of variables and extreme efficient DMUs are involved. The calculations need several steps, some of which use different programmes. The longest operation is the transposition of data among them; the slow whole procedure brings to the fore the need for a specific software for the smoothing problem. This will allow cases that are more complex to be studied.

It is emphasised that this study has satisfied the need for development for multiple inputs and outputs cases mentioned by Soares de Mello *et al.* (2004). However, models with constant scale returns (CCR) have yet to be looked into.

Finally, it should be mentioned that this smoothing model eliminates two of the worst DEA problems: inefficient Pareto regions on the efficient frontier and non-complete dimension

facets. The first because of a monotonicity constraint: they do not exist because output partial derivatives with respect to inputs are never nil. Non-complete dimension facets are eliminated, as the frontier will be defined by a single polynomial equation. This avoids the problems raised by Gonzalez-Araya (2003).

Acknowledgements

To CNPq for its financial support.

References

- (1) Adler, N.; Friedman, L. & Sinuany-Stern, Z. (2002). Review of ranking methods in the data envelopment analysis context. *European Journal of Operational Research*, **140**(2), 249-265.
- (2) Angulo-Meza, L. & Lins, M.P.E. (2002). Review of methods for increasing discrimination in data envelopment analysis. *Annals of Operations Research*, **116**(1-4), 225-242.
- (3) Angulo-Meza, L.; Biondi Neto, L.; Soares de Mello, J.C.C.B. & Gomes, E.G. (2005). ISYDS – Integrated System for Decision Support (SIAD – Sistema Integrado de Apoio a Decisão): A Software Package for Data Envelopment Analysis Model. *Pesquisa Operacional*, **25**(3), 493-503.
- (4) Andersen, P. & Petersen, N.C. (1993). A procedure for ranking efficient units in Data Envelopment Analysis. *Management Science*, **39**(10), 1261-1264.
- (5) Avellar, J.V.G.; Milioni, A.Z. & Rabello, T.N. (2007). Spherical frontier DEA model based on a constant sum of inputs. *Journal of the Operational Research Society*, **58**(9), 1246-1251.
- (6) Avellar, J.V.G.; Milioni, A.Z.; Marchi, M.M. & Pinto, M.J. (2004). Distribuição de Horas de Vôo no COMAER utilizando um modelo DEA de fronteira hiperbólica. *Revista Pesquisa Naval*, **17**, 45-50.
- (7) Avellar, J.V.G.; Milioni, A.Z. & Rabello, T.N. (2005). Modelos DEA com variáveis limitadas ou soma constante. *Pesquisa Operacional*, **25**(1), 135-150.
- (8) Banker, R.D.; Charnes, A. & Cooper, W.W. (1984). Some models for estimating technical scale inefficiencies in Data Envelopment Analysis. *Management Science*, **30**(9), 1078-1092.
- (9) Charnes, A.; Cooper, W.W.; Golany, B.; Seiford, L. & Stutz, J. (1985). Foundations of Data Envelopment Analysis for Pareto-Koopmans efficient empirical production functions. *Journal of Econometrics*, **30**(1-2), 91-107.
- (10) Charnes, A.; Cooper, W.W. & Rhodes, E. (1978). Measuring the Efficiency of Decision-Making Units. *European Journal of Operational Research*, **2**(6), 429-444.
- (11) Chilingerian, J.A. (1995). Evaluating physician efficiency in Hospitals: a multivariate analysis of best practice. *European Journal of Operational Research*, **80**(3), 548-574.
- (12) Coelli, T. ; Rao, D.S.P. & Battese, G.E. (1998). *An introduction to efficiency and productivity analysis*. Kluwer Academic Publishers.

- (13) Cooper, W.W.; Seiford, L.M. & Tone, K. (2000). *Data Envelopment Analysis: A Comprehensive Text with Models, Applications, References and DEA-Solver Software*. Kluwer Academic Publishers, USA.
- (14) Doyle, J.R. & Green, R.H. (1994). Efficiency and Cross-efficiency in DEA: Derivations, Meanings and Uses. *Journal of the Operational Research Society*, **45**(5), 567-578.
- (15) Dulá, J.H. (2002). Computations in DEA. *Pesquisa Operacional*, **22**(2), 165-182.
- (16) Gomes, E.G. & Lins, M.P.E. (2008). Modelling undesirable outputs with zero sum gains data envelopment analysis models. *Journal of the Operational Research Society*, **59**(5), 616-623.
- (17) Gomes, E.G. & Avellar, J.V.G. (2005). Modelos de análise de envoltória de dados com variáveis de soma constante: teoria e aplicação. *Pesquisa naval*, **18**, 99-104.
- (18) Gomes, E.G.; Soares de Mello, J.C.C.B. & Lins, M.P.E. (2004). Redistribuição de inputs e outputs em modelos de análise de envoltória de dados com ganhos de soma zero. *Pesquisa Operacional*, **24**(2), 269-284.
- (19) Gonzalez-Araya, M.C. (2003). Projeções Não Radiais em Regiões Fortemente Eficientes da Fronteira DEA – Algoritmos e Aplicações. Tese de Doutorado, Engenharia de Produção, COPPE/UFRJ.
- (20) Kozyreff Filho, E. & Milioni, A.Z. (2004). Um Método para Estimativa de Metas DEA. *Produção*, **14**(2), 90-101.
- (21) Liang, L.; Wu, J.; Cook, W.D. & Zhu, J. (2008). Alternative secondary goals in DEA cross-efficiency evaluation. *International Journal of Production Economics*, **113**(2), 1025-1030.
- (22) Lins, M.P.E.; Gomes, E.G.; Soares de Mello, J.C.C.B. & Soares de Mello, A.J.R. (2003). Olympic ranking based on a Zero Sum Gains DEA model. *European Journal of Operational Research*, **148**(2) 85-95.
- (23) Nacif, F.B. (2005). Suavização da fronteira DEA: Modelo geral BCC n-dimensional. Dissertação de Mestrado, Engenharia de Produção, Universidade Federal Fluminense, Niterói.
- (24) Nacif, F.B.; Pimenta, H.L.N. & Soares de Mello, J.C.C.B. (2004). Suavização da Fronteira DEA: O Caso BCC N-Dimensional com Multiplicidade Simultânea dos Inputs e dos Outputs. *Anais do XXXVI SBPO*, São João del Rei.
- (25) Olesen, O.P. & Petersen, N.C. (1996). Indicators of ill-conditioned data sets and model misspecifications in Data Envelopment Analysis: an extended facet approach. *Management Science*, **42**(2), 205-219.
- (26) Reddy, J.N. (1993). *Introduction to the Finite Element Method*. McGraw-Hill.
- (27) Rosen, D.; Schaffnit, C. & Paradi, J.C. (1998). Marginal rates and two dimensional level curves in DEA. *Journal of Productivity Analysis*, **9**(3), 205-232.
- (28) Reinhard, S.; Knox Lovell, C.A. & Thijssen, G.J. (2000). Environmental efficiency with multiple environmentally detrimental variables; estimated with SFA and DEA. *European Journal of Operational Research*, **121**(2), 287-303.

- (29) Soares de Mello, J.C.C.B.; Gomes, E.G.; Biondi Neto, L. & Lins, M.P.E. (2004). Suavização da Fronteira DEA: O caso BCC tridimensional. *Investigação Operacional*, **24**(1), 89-107.
- (30) Soares de Mello, J.C.C.B.; Lins, M.P.E. & Gomes, E.G. (2002). Construction of a smoothed DEA frontier. *Pesquisa Operacional*, **22**(2), 183-201.
- (31) Soares de Mello, J.C.C.B.; Angulo-Meza, L. & Branco da Silva, B.P. (2009). A ranking for the Olympic Games with unitary input DEA models. *IMA Journal of Management Mathematics*, **20**(2), 201-211.
- (32) Soares de Mello, J.C.C.B.; Gomes, E.G.; Angulo-Meza, L. & Biondi Neto, L. (2008). Cross Evaluation using Weight Restrictions in Unitary Input DEA Models: Theoretical Aspects and Application to Olympic Games Ranking. *WSEAS Transactions on Systems*, **7**(1), 31-39.
- (33) Thanassoulis, E. (1993). A comparison of regression analysis and Data Envelopment Analysis as alternative methods of performance assessment. *Journal of the Operational Research Society*, **44**(11), 1129-1144.