# APPLICATIONS OF SELF-ORGANIZING MAPS

Edited by **Magnus Johnsson**

**Applications of Self-Organizing Maps**
http://dx.doi.org/10.5772/3464
Edited by Magnus Johnsson

**Contributors**

Wattanapong Kurdthongmee, Dimitris Kontogiorgos, Hector Benitez-Perez, Hiroshi Dozono, Flavius Gorgonio, Huliane Silva, Cícero Silva, Ji Zhang, Jennifer Francis, Natasa Skific, Ryotaro Kamimura, Lydwine Gross, Arnaud Quesney, Suzanne Angeli, Marina Resta, Yuan-Chao Liu, Masaki Ishii, Fernando Bacao, Victor Lobo, Roberto Henriques, João Carlos Soares de Mello, Eliane Gomes, Lidia Angulo Meza, Sérgio De Zen, Thiago Bernardino De Carvalho, Urbano Gomes Abreu, Luiz Biondi Neto

**Published by InTech**

Janeza Trdine 9, 51000 Rijeka, Croatia



**Publishing Process Manager** Iva Lipovic

**Technical Editor** InTech DTP team

**Cover** InTech Design team

First published November, 2012

Printed in Croatia

A free online edition of this book is available at www.intechopen.com

Additional hard copies can be obtained from orders@intechopen.com

Applications of Self-Organizing Maps, Edited by Magnus Johnsson

p. cm.

ISBN 978-953-51-0862-7

Chapter 4

# Ex-Post Clustering of Brazilian Beef Cattle Farms Using Soms and Cross-Evaluation Dea Models

João Carlos Correia Baptista Soares de Mello,
Eliane Gonçalves Gomes, Lidia Angulo Meza,
Luiz Biondi Neto, Urbano Gomes Pinto de Abreu,
Thiago Bernardino de Carvalho and Sergio de Zen

Additional information is available at the end of the chapter

http://dx.doi.org/10.5772/51324

## 1. Introduction

The beef cattle production system is the set of technologies and management practices, animal type, purpose of breeding, breed group and the eco-region where the activity is developed. The central structure in the beef cattle production chain is the biological system of beef production, including the stages of creation (cow-calf production, stocker production, feedlot beef production) and their combinations. The cow-calf phase is the less profitable activity and the one that has the higher risk. However, it supports the entire structure of the production system.

Although it is undeniable that a systemic view in agriculture is important, it is not yet established in the Brazilian agricultural research. In this study, by using a non parametric technique known as Data Envelopment Analysis (DEA) and Self-Organizing Maps (SOMS), we intend to cluster cattle breeders of some Brazilian municipalities. The objective is to group the farmers according to their efficiency profiles regarding the decisions related to the composition of their production systems, which are focused on the cow-calf phase. The farmers' decisions have direct impacts on the expenditures and on the income reached.

Efficiency is a relative concept: we compare the amount produced by a productive unit or firm given the available resources, to the maximum quantities that would have been produced with the same amount of resources. There are different approaches to measure efficiency. The so-called parametric methods assume a pre-defined functional relationship between resources (inputs) and products (outputs). These methods usually use averages to determine the amount that would have been produced. The non-parametric approaches, among which DEA, do not make any assumption about the functional form of the relationship between inputs and outputs. DEA assumes that the maximum that could have been produced is obtained in the sample under evaluation, by observing the firms that better perform. DEA was initially presented in [1] and computes the efficiency of productive units, the so-called decision making units (DMUs). The DMUs, here the farms, use the same types of inputs to produce the same sort of outputs, and this set must be homogenous.

Mathematically and in the presence of multiple inputs and multiple outputs, efficiency is the ratio between the weighted sum of outputs and the weighted sum of inputs. In the DEA approach, each firm chooses the most appropriate set of weights with the view to maximize that ratio. This choice is neither arbitrary nor subjective, as it is based on some restrictions. The result of the ratio must mean efficiency, i.e., it must be a number between 0 and 1. Thus, it is necessary that the weights that a firm chooses when applied to itself and to the others (in total of *k* firms) cannot give a number bigger than one. These considerations are set in a mathematical programming problem that formalizes the DEA model. In this context, we may say that DEA scores are a benevolent measure of efficiency.

SOMS are a special case of neural networks. In the literature we can find some papers that discuss the joint use of neural networks and data envelopment analysis. SOMS are often used with DEA to perform an ex-ante clustering, in order to allow DEA to evaluate only homogeneous units. For instance, in [2] Neural Networks and DEA were used to determine if the differences among efficiency scores were due to environmental variables or to the management process. The use of Neural Network for clustering and benchmarking container terminals was done in [3]. In [4], authors used Kohonen self-organizing maps to cluster participating countries in the Olympics and then applied DEA for producing a new ranking of participating teams. In references [5, 6] they used the back propagation neural network algorithm to accelerate computations in DEA. In [7] authors applied Neural Networks to estimate missing information for suppliers evaluation using DEA. In [8], SOMS were used to cluster the CEDERJ distance learning centers before applying DEA models to assess the efficiency of these centers. The differences between the clusters were taken into account by a homogenization of the centers for the purpose of comparing them in one single cluster.

In this chapter we propose a different approach. SOMS are used for ex-post evaluation, as discussed in [9]. In our approach we first perform a DEA evaluation and then we cluster the DMUs into groups based on DEA efficiencies. In our case we compute the farmers' efficiency profiles and the SOMS are used to group the DMUs according to their common characteristics. The efficiency profiles are derived from the DEA cross-efficiency matrix. In the cross-evaluation approach [10] each DMU is assessed by its own set of multipliers as well as by the other DMUs' multipliers. Thus, the column vector of efficiencies calculated for each DMU in the cross-evaluation matrix is assumed as the efficiency profile. It is important to note that the DEA cross-efficiency technique is usually used for ranking purposes. Here, however, it is not used as a ranking tool but as the input to a SOM that will be used to cluster the units according to the way their efficiencies were obtained.

Our discussion proceeds as follows. In Section 2 we discuss the theoretical aspects of DEA and of cross-evaluation analysis. Section 3 is about the fundamental aspects of SOMS: Kohonen Neural Networks. In Section 4 we present the proposed approach: DEA and SOMS in an ex-post evaluation. The case study is detailed in section 4. In Section 5 we discuss the results. Finally, we summarize our conclusions and list the references.

## 2. Dea and cross-efficiency evaluation

DEA is a mathematical programming approach developed to compute the comparative efficiency of productive units (firms or DMUs). The DMUs perform similar tasks and use different amounts of inputs to produce different quantities of outputs. In order to maximize the efficiency of each DMU, DEA models allow each one to choose the weight assigned to each variable in complete freedom, subject to some restrictions. In the case the firms under evaluation operate under different scales, it is possible to consider these differences in a DEA formulation. This model is known in the DEA literature as VRS (variable returns to scale) or BCC model (from Banker, Charnes and Cooper), as firstly discussed in [11]. On the other hand, if it is possible to assume that the DMUs operate under constant return to scale, i.e., there is a proportional relationship in increments or decrements between inputs and outputs, the suitable model is the CRS (constant returns to scale) or CCR (from Charnes, Cooper and Rhodes), as presented in the seminal paper [1].

The mathematical linear formulation for DEA CCR model is shown in model (1), where $h_1$ is the efficiency of the DMU '1' under evaluation; $x_{ik}$and $y_{jk}$ are the *i-th* input and *j-th* output of the *k-th* DMU, $k$=1…$n$; $\mu_j$ and $v_i$are the output and the input weights or multipliers, respectively.

$$
\begin{aligned}
&\max h_1 = \sum_{j=1}^{s} \mu_j y_{j1} \\
&\text{subject to} \\
&\sum_{i=1}^{r} v_i x_{i1} = 1 \\
&\sum_{j=1}^{s} \mu_j y_{jk} - \sum_{i=1}^{r} v_i x_{ik} \le 0 \\
&\mu_j,\ v_i \ge 0, \quad k=1...n,\ \ j=1...s,\ \ i=1...r
\end{aligned}
\tag{1}
$$

As previously mentioned, DEA allows each DMU to perform its self-evaluation, in other words, to choose its own set of multipliers in such a way that its efficiency score is maximized. In [10], it was proposed that the optimal weights for each DMU can be used to evaluate its peers, i.e., to compute alternative efficiency scores for every other DMU. This evaluation performed by the complete set of DMUs is called cross-evaluation [12, 13] and the resulting measures of performance are the cross-efficiencies. In (2), $h_{k1}$is the cross-efficiency of DMU '1' using the weights of the DMU '$k$'; $\mu_{jk}$ is the weight of output $j$ obtained for DMU '$k$'; $v_{ik}$ is the weight of input $i$ obtained for DMU '$k$'. The final cross-efficiency index is the average of all peer and self-evaluations.

$$h_{k1} = \frac{\sum_{j=1}^{s} \mu_{jk} y_{j1}}{\sum_{i=1}^{r} v_{ik} x_{i1}} \tag{2}$$

As discussed in [14], the use of cross-evaluation has spread to a number of different areas, as in multiple criteria decision making to improve discrimination among alternatives, in a preferential election, in the ranking and selection of projects and technologies, among others. As discussed in [15-17] cross-evaluation commonly uses the CCR model due to the existence of negative efficiencies when choosing the BCC model. The usage of the BCC model for cross-evaluation imposes the inclusion of an additional set of restrictions to ensure positive efficiencies [16, 17].

As was noticed in [10], in the DEA context the optimal set of weights for a DMU is not unique. Therefore it is necessary to choose one set of weights among all the possibilities. In the original cross-evaluation formulation these approaches led to nonlinear problems. In [12] two linear models were introduced, that are approximations to the original formulation. There are some other alternative models for the cross-efficiency evaluation. We mention the DEA-Game [18], among others [19]. These models may use two approaches. The first one, called "aggressive model", minimizes the cross-evaluation indexes of all DMUs. The second approach, known as "benevolent model", maximizes the cross-evaluation scores of all DMUs. The so-called aggressive $C_k$ formulation is shown in model (3), as presented in [12].

$$\begin{aligned}
&\min C_k = \sum_{j=1}^{s} \mu_{jk} \sum_{m \neq k} y_{jm} \\
&\text{subject to} \\
&\sum_{i=1}^{r} v_{ik} \sum_{m \neq k} x_{im} = 1 \\
&\sum_{j=1}^{s} \mu_{jk} y_{jm} - \sum_{i=1}^{r} v_{ik} x_{im} \leq 0, \quad \forall\, m \neq k \\
&\sum_{j=1}^{s} \mu_{jk} y_{jk} - h_{kk} \sum_{i=1}^{r} v_{ik} x_{ik} = 0 \\
&\mu_{jk},\ v_{ik} \geq 0
\end{aligned} \tag{3}$$

After determining the set of weights to be used in the cross-evaluation, a cross-efficiency matrix is calculated, as shown in Table 1. This matrix contains the self-evaluation ($h_{kk}$) and the peer-evaluation ($h_{kl}$, $k \neq l$) of each DMU. The efficiency scores of a given DMU will be in the column of this matrix.

| DMU | 1 | 2 | 3 | ... | *n* |
|---|---|---|---|---|---|
| 1 | $h_{11}$ | $h_{12}$ | $h_{13}$ | ... | $h_{1n}$ |
| 2 | $h_{21}$ | $h_{22}$ | $h_{23}$ | ... | $h_{2n}$ |
| 3 | $h_{31}$ | $h_{32}$ | $h_{33}$ | ... | $h_{35}$ |
| ... | ... | ... | ... | ... | ... |
| *n* | $h_{n1}$ | $h_{n2}$ | $h_{n3}$ | ... | $h_{n5}$ |

**Table 1.** Cross-efficiencies matrix for *n* DMUs.

We may point out that DMUs sharing similar weights distribution (with similar characteristics) will evaluate each other with high efficiency scores. On the contrary, those who have dissimilar characteristics will evaluate each other with low efficiency scores. The efficiency scores of a DMU (the column of the matrix) are limited by the efficiency score obtained in the self-evaluation (the diagonal of the matrix, $h_{kk}$). In the self-evaluation, provided by the classic DEA models, a DMU is shown in the best possible way as to maximize its efficiency. Therefore, all peer-evaluations will be lower or equal to that efficiency score.

## 3. Fundamental aspects of som: kohonen neural network

The human brain organizes information in a logic way. A paramount aspect of the self-organized networks is motivated by the organization of the human brain in such a way that the sensory inputs are represented by topologically organized maps. The Kohonen self-organizing map emulates the unsupervised learning in a simple and elegant way taking into account the neuron neighborhood [20].

From topographic map development principle came up two feature mapping models: the model presented in [21, 22], having strong neurobiological motivations, and the Kohonen model [23], not as close to neurobiology as the previous one but enabling a simple computing treatment stressing the essential characteristics of the brain maps. Moreover, the Kohonen model or Kohonen Self-Organizing Map yields a low input dimension.

The SOMS are artificial neural networks special structures in a grid form that work in a similar way of the human brain, as far as the information organization is concerned, and are based on competitive learning. The most used SOM is the topologically interconnected two-dimensional, where the neurons are represented by rectangular, hexagonal and random grid knots of neighbor neurons. Higher dimensional maps can also be modeled.

In order to analyze the competitive process, let us suppose that the input space is $m$-dimensional and that $X$ represents a random input pattern [24] such that one can write $X = [X_1, X_2, X_3, ..., X_m]$. Assuming that the weight vector $W$ of each neuron has the same dimension of the input space, for a given neuron $j$ of a total of $l$ neurons the weight vector can be written as $W = [W_{j1}, W_{j2}, W_{j3}, ..., W_{jm}]$, $j$=1…$l$.

For each input vector, the scalar product is evaluated in order to find the $X$ vector which is closest to the weight vector $W$. By comparison, the maximum scalar product $\max\left(W_j^t \cdot X\right)$ is chosen, representing the location in which the topological neighborhood of excited neurons should be centered. Maximizing this scalar product is equivalent to minimizing the Euclidian distance between $X$ and $W$. Other metrics could also be used.

The neuron with the weight vector closest to the input vector $X$ is called the winner neuron, whose index $V(X)$ is given by$V(X)=\min\| X-W_j\|$ . In the cooperative process the winner neuron locates the centre of a topological neighborhood of cooperating neurons. The active winner tends to strongly stimulate its closest neighbor neurons and weakly the farthest ones. It is essential to find a topological neighborhood function $N_{j,V(X)}$ that is independent from the winner neuron location. That neighborhood function should represent the topological neighborhood centered in the winner neuron, denoted by $V$, having as closest lateral neighbors a group of excited and cooperative neurons, from which a representative $j$ neuron can be chosen. The lateral distance $D_{j,V}$ between the winner neuron, indexed by $V$, and the excited neuron, indexed by $j$, can be written as in (4), where is the neighborhood width [24].

$$N_{j,V(X)}=\exp\left(-\frac{D_{j,V}^2}{2\sigma^2}\right) \tag{4}$$

The more dependent is the lateral distance$D_{j,V}$, the greater will be the cooperation among the neighborhood neurons. So, for a two-dimensional output grid, the lateral distance can be defined as in (5), for which the discrete vector $\wp_j$ represents the position of the excited neuron, and $\wp_V$ the position of the neuron that won the competition.

$$D_{j,V}=\sqrt{\| \wp_j-\wp_V\|^2} \tag{5}$$

The topological neighborhood should decrease with discrete time t. In order to accomplish that, the width $\sigma$ of the topological neighborhood $N_{j,V(X)}$ should decrease in time. The width could be written as in (6), where $\sigma_0$ represents the initial value of the neighborhood width and $\tau_1$ a time constant. Usually $\sigma_0$ is adjusted to have the same value as the grid ratio, i.e.,$\tau_1=1,\ 000/\log\sigma_0$.

$$\sigma(t)=\sigma_0\exp\left(-\frac{t}{\tau_1}\right), \quad t=0,\ 1,\ 2,\ 3,\ \ldots \tag{6}$$

The expression of the topological neighborhood in time can be written as in (7).

$$N_{j,V(X)}(t)=\exp\left(-\frac{D_{j,V}^2}{2\sigma^2(t)}\right), \quad t=0,\ 1,\ 2,\ 3,\ \ldots \tag{7}$$

The adaptive process is the last phase of the SOM procedure. During this phase is carried out the adjustment of the connection weights of the neurons. In order the network succeed in the self-organization task, it is necessary to update the weights $W_j$ of the excited j neuron relatively to the input vector X.

The change to the weight vector of the excited neuron j in the grid can be written as $\Delta W_j = \eta\ y_j X - g(y_j) W_j$, where is the learning rate parameter, $\eta\ y_j X$ is the Hebbian term, and $g(y_j) W_j$ is called forgetting term [23-25].

In order to satisfy the requirement, a linear function for $g(y_j)$ is chosen as$g(y_j) = \eta\ y_j$. If $y_j = N_{j,V(X)}$, then the expression can be written as (8).

$$\Delta W_j = \eta\ N_{j,V(X)} (X - W_j) \tag{8}$$

Using discrete-time notation, a weight updating equation can be written, which applies to all neurons that are within the topographic neighborhood equation of the winner neuron [23, 24], as in (9).

$$W_j(n+1) = W_j(n) + \eta(n)\ N_{j,V(X)}(n) (X - W_j(t)) \tag{9}$$

In (9) the learning rate parameter changes in each iteration, with an initial value around 0.1 and decreases with increasing discrete-time t up to values above 0.01 [20]. To that end, equation (10) is written in which decays exponentially and $\tau_2$ is another time-constant of the SOM algorithm. For the fulfillment of the requirements one could choose for instance, $\eta_0 = 0.1$and$\tau_2 = 1,000$.

$$\eta(t) = \eta_0 \exp\left(-\frac{t}{\tau_2}\right), \quad t = 0, 1, 2, 3, \ldots \tag{10}$$

## 4. Dea and soms in an ex-post evaluation

As discussed in Section 2, in the cross-evaluation approach each DMU is assessed by its own set of multipliers (self-evaluation) as well as by the other DMUs' multipliers (peer-evaluation). The cross-efficiency matrix provides the cross-efficiencies of every DMU, each column providing the efficiencies of each DMU, i.e., is the evaluation performed by the other DMUs. As previously discussed, DMUs with the same characteristics will evaluate each other with high efficiencies, or with low efficiencies otherwise. Therefore, we can assume that each column of the cross-evaluation matrix is the efficiency profile of each DMU.

When high value of weights are assigned to the variables of DMU A this implies that DMU A will have a good efficiency score (ratio between the weighted sum of the outputs and the weighted sum of the inputs). In the context of cross-evaluation, when these weights are assigned to DMU B and it also has high efficiency score we may say that DMU B has similar characteristics to DMU A. We can observe the case in which a DMU have a good performance in all variables (or in the majority of them), but it is not unusual the case where the efficiency score is computed based only on two variables (one input and one output). This occurs when the DMU have a low performance on some variables and null weights (or near zero) are assigned to them. Therefore, considering DMUs A and B, DMU A may evaluate DMU B positively, but when DMU B evaluates DMU A, its efficiency may be low. In other words, cross-evaluation may not bring the same results for both.

In this chapter we will use SOMS to cluster the farmers with similar characteristics using the cross-efficiency matrix information. However, the cross-efficiencies are dependent on the DMU self-evaluation. The cross-efficiency scores may be smaller compared to those obtained by other DMUs (lines in the matrix), but they may be consistent or very similar to the efficiency score obtained by the DMU (column in the matrix). Therefore, as our clusters will be defined by the efficiency profile, we need to remove the self-evaluation effect of each DMU. In this approach this is accomplished by normalizing each column of the matrix by the self-evaluation, i.e., all values are divided by the CCR DEA efficiency of the DMU. This is shown in Table 2. The self-evaluation is the highest efficiency score of the DMU, located in the diagonal of the cross-evaluation matrix. As a consequence, we will have unitary values in the matrix diagonal. The objective of this operation is to group the DMUs by the way they are evaluated by the others rather than by their efficiency level. This minimizes the benevolent characteristic of the classic DEA models.

| **DMU** | **1** | **2** | **3** | **...** | ***n*** |
|---|---|---|---|---|---|
| 1 | $\frac{h_{11}}{h_{11}}$ | $\frac{h_{12}}{h_{22}}$ | $\frac{h_{13}}{h_{33}}$ | ... | $\frac{h_{1n}}{h_{nn}}$ |
| 2 | $\frac{h_{21}}{h_{11}}$ | $\frac{h_{22}}{h_{22}}$ | $\frac{h_{23}}{h_{33}}$ | ... | $\frac{h_{2n}}{h_{nn}}$ |
| 3 | $\frac{h_{31}}{h_{11}}$ | $\frac{h_{32}}{h_{22}}$ | $\frac{h_{33}}{h_{33}}$ | ... | $\frac{h_{3n}}{h_{nn}}$ |
| ... | ... | ... | ... | ... | ... |
| $n$ | $\frac{h_{n1}}{h_{11}}$ | $\frac{h_{n2}}{h_{22}}$ | $\frac{h_{n13}}{h_{33}}$ | ... | $\frac{h_{nn}}{h_{nn}}$ |

**Table 2.** Normalized cross-efficiency matrix.

In the traditional cross-evaluation applications, the resulting average cross-efficiency index is commonly used for ranking purposes. Here, in contrast, the square matrix of normalized efficiencies will be submitted to the SOMS procedure (as its input) to generate the homogeneous clusters. In other words, the normalized cross-evaluation matrix will be used to cluster the DMUs according to their efficiency profile.

An approach based on cross-evaluation and cluster technique was used in [26] to group participating countries in Summer Olympics Games. The authors used the average cross-efficiencies and cluster analysis to group the nations and to select more appropriate targets for poorly performing countries to use as benchmarks.

## 5. Case study

The beef complex in Brazil is consolidated as an important link in the production and in the international trade: Brazil is the largest exporter and the second largest producer of beef. Considering this scenario, the study and the evaluation of beef cattle production systems are important tools for enhancing the performance of this sector.

In Brazil, the cow-calf beef cattle phase occurs predominantly in extensive continuous grazing, with native or cultivated pastures, encompassing: calves (until weaning or even one year old), cows, heifers and bulls. The cow-calf phase supports the entire structure of the beef production chain.

This case study seeks to assess the comparative performance of extensive livestock modal production systems in its cow-calf phase, in some Brazilian municipalities. The objective is to measure their performance and to group them according to the decisions regarding the composition of the production system. This has a direct impact on the expenditures and on the income generated.

### 5.1 Data Source

Primary data were collected through the panel system, which allows the definition of representative farms, as proposed in [27].

In this approach, based on the experience of the participating farmers, it is characterized the property that is the most commonly found in the region, i.e., a property or a production system that is representative of the locality under study. In some cases it is not possible to determine this typology and more than one representative property or production system are specified.

The panel is a less costly procedure of obtaining information than the census or the sampling of farms. The technique is applied during a meeting with a group of one or more researchers, one technician and eight regional farmers. Meetings are scheduled in advance, with the support of rural unions and regional contacts. At the end of that debate one can say that any characterization of the typical farm in the region has the consent of the farmers. Thus, productivity rates, establishment costs, fixed and variable costs, i.e., all the numbers resulting from the panel, tend to be fairly close to the regional reality.

It is noteworthy that the rates and the costs reported by each participant are not related to their properties, but with a single farm, declared at the beginning of the panel as the one that best represents the scale of operation and the production system of most of the local properties.

This study evaluated 21 beef cattle modal production systems that performed only the cow-calf phase, in seven states of Brazil. The data, derived from the indicators of the project developed by the Centro de Estudos Avançados em Economia Aplicada and the Confederação da Agricultura e Pecuária do Brasil, were collected in municipalities of these seven states: Mato Grosso do Sul - MS (eight), Goiás - GO (four), Rio Grande do Sul - RS (one), Minas Gerais - MG (four), Tocantins - TO (two), São Paulo - SP (one) and Bahia - BA (one). Panels with the farmers, with the support of the local rural technical assistance, were performed to collect the data, according to the methodology described in [28].

## 5.2 Modeling

### 5.2.1 DMUs

The objective of the DEA model is to measure the performance of the farmers' decision regarding the composition of the rearing production system. Thereby, the DMUs are the 21 modal systems, identified from the panel discussions in 21 cities in seven Brazilian states. Table 3 presents the dataset.

| DMUs | | | Breeders | Calves | Cull cows |
|---|---|---|---|---|---|
| Municipality | State | Code | (input) | (output) | (output) |
| Alvorada | TO | #1 | 12 | 147 | 30 |
| Amambai | MS | #2 | 15 | 143 | 40 |
| Aquidauana | MS | #3 | 92 | 713 | 214 |
| Bonito | MS | #4 | 14 | 166 | 75 |
| Brasilândia | MS | #5 | 31 | 290 | 178 |
| Camapuã | MS | #6 | 9 | 65 | 33 |
| Carlos Chagas | MG | #7 | 19 | 297 | 160 |
| Catalão | GO | #8 | 8 | 81 | 42 |
| Corumbá | MS | #9 | 69 | 455 | 200 |
| Itamarajú | BA | #10 | 4 | 44 | 18 |
| Lavras do Sul | RS | #11 | 5 | 58 | 30 |
| Montes Claros | MG | #12 | 5 | 47 | 28 |
| Niquelândia | GO | #13 | 4 | 35 | 18 |
| Paraíso do Tocantins | TO | #14 | 12 | 123 | 35 |
| Porangatu | GO | #15 | 5 | 46 | 23 |
| Ribas Rio Pardo | MS | #16 | 15 | 143 | 70 |
| Rio Verde | GO | #17 | 23 | 196 | 82 |
| São Gabriel d'Oeste | MS | #18 | 11 | 95 | 40 |
| Tupã | SP | #19 | 5 | 46 | 30 |
| Uberaba | MG | #20 | 5 | 66 | 36 |
| Uberlândia | MG | #21 | 2 | 20 | 10 |

**Table 3.** DMUs, inputs and outputs.

### 5.2.2 Variables

The technicians and the researchers mentioned in 5.1, analyzed the variables set and immediately identified those relevant to our study. They selected "number of bulls" as the input variable, since this variable represents a significant portion of all total expenditures of the ranchers that produce calves. It is directly linked to the quality of animals that will be sold in these systems. This is also the only category that is purchased from other herd, especially in ranches with herds of genetic selection.

The products of the system that are responsible for the main revenue from the cow-calf systems were chosen as the output variables. These are the "number of calves on the herd" and the "number of cull cows". All calves produced are sold on the property and generate income. Cull cows are those that are sold, as they are not part of the production system either by higher age or by reproductive performance lower than desired. The decision to fit between these variables is important because it will provide the dynamics structure of the breeding herd. This is the fundamental factor for keeping the economic viability of the beef cattle production system.

The variables indicated by experts need to be examined by analysts to determine whether they conform to the properties required by the DEA models. In particular, there must be a causal relationship between each input-output pair [29]. There is a clear causal relationship between the output "number of calves on the herd" and the input "number of bulls". The same cannot be said of the relationship between the input and the output "number of cull cows". Actually, there is no direct causal relationship between these variables; however there is a cost-benefit relationship. In the case the rancher has a big number of bulls (that represent an expense) he must earn more, either through the sale of calves or cows. Therefore, the "bulls – cull cows" ratio makes sense when using DEA to analyze cost-benefit ratios, and not just pure productive relations. This interpretation of DEA was introduced in [30] and was used in [31, 32]. Generalizations of this usage can be seen in [33].

## 6. Results and discussion

Figure 1 shows the cross-evaluation matrix. It was computed using the SIAD software [34]. As previously mentioned, we used the DEA CCR as the BCC model can generate negative efficiency scores [17].

The values of the Figure 1, cross-evaluation matrix, were normalized before using SOMS, as explained in Section 4. Four different grid dimensions were tested: the 5x5 grid, the 3x3 grid, the 2x2 grid and the 2x1 grid. We began our analysis with the 5x5 grid since it yields 25 neurons against the 21 DMUs of the data set.

| | #1 | #2 | #3 | #4 | #5 | #6 | #7 | #8 | #9 | #10 | #11 | #12 | #13 | #14 | #15 | #16 | #17 | #18 | #19 | #20 | #21 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| #1 | 0.7837 | 0.6099 | 0.4958 | 0.7585 | 0.5985 | 0.4620 | 1.0000 | 0.6477 | 0.4219 | 0.7037 | 0.7421 | 0.6013 | 0.5598 | 0.6557 | 0.5886 | 0.6099 | 0.5452 | 0.5525 | 0.5886 | 0.8444 | 0.6397 |
| #2 | 0.7837 | 0.6099 | 0.4958 | 0.7585 | 0.5985 | 0.4620 | 1.0000 | 0.6477 | 0.4219 | 0.7037 | 0.7421 | 0.6013 | 0.5598 | 0.6557 | 0.5886 | 0.6099 | 0.5452 | 0.5525 | 0.5886 | 0.8444 | 0.6397 |
| #3 | 0.7837 | 0.6099 | 0.4958 | 0.7585 | 0.5985 | 0.4620 | 1.0000 | 0.6477 | 0.4219 | 0.7037 | 0.7421 | 0.6013 | 0.5598 | 0.6557 | 0.5886 | 0.6099 | 0.5452 | 0.5525 | 0.5886 | 0.8444 | 0.6397 |
| #4 | 0.7837 | 0.6099 | 0.4958 | 0.7585 | 0.5985 | 0.4620 | 1.0000 | 0.6477 | 0.4219 | 0.7037 | 0.7421 | 0.6013 | 0.5598 | 0.6557 | 0.5886 | 0.6099 | 0.5452 | 0.5525 | 0.5886 | 0.8444 | 0.6397 |
| #5 | 0.2969 | 0.3167 | 0.2762 | 0.6362 | 0.6819 | 0.4354 | 1.0000 | 0.6234 | 0.3442 | 0.5344 | 0.7125 | 0.6650 | 0.5344 | 0.3464 | 0.5463 | 0.5542 | 0.4234 | 0.4318 | 0.7125 | 0.8550 | 0.5938 |
| #6 | 0.7837 | 0.6099 | 0.4958 | 0.7585 | 0.5985 | 0.4620 | 1.0000 | 0.6477 | 0.4219 | 0.7037 | 0.7421 | 0.6013 | 0.5598 | 0.6557 | 0.5886 | 0.6099 | 0.5452 | 0.5525 | 0.5886 | 0.8444 | 0.6397 |
| #7 | 0.2969 | 0.3167 | 0.2762 | 0.6362 | 0.6819 | 0.4354 | 1.0000 | 0.6234 | 0.3442 | 0.5344 | 0.7125 | 0.6650 | 0.5344 | 0.3464 | 0.5463 | 0.5542 | 0.4234 | 0.4318 | 0.7125 | 0.8550 | 0.5938 |
| #8 | 0.7837 | 0.6099 | 0.4958 | 0.7585 | 0.5985 | 0.4620 | 1.0000 | 0.6477 | 0.4219 | 0.7037 | 0.7421 | 0.6013 | 0.5598 | 0.6557 | 0.5886 | 0.6099 | 0.5452 | 0.5525 | 0.5886 | 0.8444 | 0.6397 |
| #9 | 0.7837 | 0.6099 | 0.4958 | 0.7585 | 0.5985 | 0.4620 | 1.0000 | 0.6477 | 0.4219 | 0.7037 | 0.7421 | 0.6013 | 0.5598 | 0.6557 | 0.5886 | 0.6099 | 0.5452 | 0.5525 | 0.5886 | 0.8444 | 0.6397 |
| #10 | 0.7837 | 0.6099 | 0.4958 | 0.7585 | 0.5985 | 0.4620 | 1.0000 | 0.6477 | 0.4219 | 0.7037 | 0.7421 | 0.6013 | 0.5598 | 0.6557 | 0.5886 | 0.6099 | 0.5452 | 0.5525 | 0.5886 | 0.8444 | 0.6397 |
| #11 | 0.7837 | 0.6099 | 0.4958 | 0.7585 | 0.5985 | 0.4620 | 1.0000 | 0.6477 | 0.4219 | 0.7037 | 0.7421 | 0.6013 | 0.5598 | 0.6557 | 0.5886 | 0.6099 | 0.5452 | 0.5525 | 0.5886 | 0.8444 | 0.6397 |
| #12 | 0.2969 | 0.3167 | 0.2762 | 0.6362 | 0.6819 | 0.4354 | 1.0000 | 0.6234 | 0.3442 | 0.5344 | 0.7125 | 0.6650 | 0.5344 | 0.3464 | 0.5463 | 0.5542 | 0.4234 | 0.4318 | 0.7125 | 0.8550 | 0.5938 |
| #13 | 0.7837 | 0.6099 | 0.4958 | 0.7585 | 0.5985 | 0.4620 | 1.0000 | 0.6477 | 0.4219 | 0.7037 | 0.7421 | 0.6013 | 0.5598 | 0.6557 | 0.5886 | 0.6099 | 0.5452 | 0.5525 | 0.5886 | 0.8444 | 0.6397 |
| #14 | 0.7837 | 0.6099 | 0.4958 | 0.7585 | 0.5985 | 0.4620 | 1.0000 | 0.6477 | 0.4219 | 0.7037 | 0.7421 | 0.6013 | 0.5598 | 0.6557 | 0.5886 | 0.6099 | 0.5452 | 0.5525 | 0.5886 | 0.8444 | 0.6397 |
| #15 | 0.7837 | 0.6099 | 0.4958 | 0.7585 | 0.5985 | 0.4620 | 1.0000 | 0.6477 | 0.4219 | 0.7037 | 0.7421 | 0.6013 | 0.5598 | 0.6557 | 0.5886 | 0.6099 | 0.5452 | 0.5525 | 0.5886 | 0.8444 | 0.6397 |
| #16 | 0.7837 | 0.6099 | 0.4958 | 0.7585 | 0.5985 | 0.4620 | 1.0000 | 0.6477 | 0.4219 | 0.7037 | 0.7421 | 0.6013 | 0.5598 | 0.6557 | 0.5886 | 0.6099 | 0.5452 | 0.5525 | 0.5886 | 0.8444 | 0.6397 |
| #17 | 0.7837 | 0.6099 | 0.4958 | 0.7585 | 0.5985 | 0.4620 | 1.0000 | 0.6477 | 0.4219 | 0.7037 | 0.7421 | 0.6013 | 0.5598 | 0.6557 | 0.5886 | 0.6099 | 0.5452 | 0.5525 | 0.5886 | 0.8444 | 0.6397 |
| #18 | 0.7837 | 0.6099 | 0.4958 | 0.7585 | 0.5985 | 0.4620 | 1.0000 | 0.6477 | 0.4219 | 0.7037 | 0.7421 | 0.6013 | 0.5598 | 0.6557 | 0.5886 | 0.6099 | 0.5452 | 0.5525 | 0.5886 | 0.8444 | 0.6397 |
| #19 | 0.2969 | 0.3167 | 0.2762 | 0.6362 | 0.6819 | 0.4354 | 1.0000 | 0.6234 | 0.3442 | 0.5344 | 0.7125 | 0.6650 | 0.5344 | 0.3464 | 0.5463 | 0.5542 | 0.4234 | 0.4318 | 0.7125 | 0.8550 | 0.5938 |
| #20 | 0.2969 | 0.3167 | 0.2762 | 0.6362 | 0.6819 | 0.4354 | 1.0000 | 0.6234 | 0.3442 | 0.5344 | 0.7125 | 0.6650 | 0.5344 | 0.3464 | 0.5463 | 0.5542 | 0.4234 | 0.4318 | 0.7125 | 0.8550 | 0.5938 |
| #21 | 0.7837 | 0.6099 | 0.4958 | 0.7585 | 0.5985 | 0.4620 | 1.0000 | 0.6477 | 0.4219 | 0.7037 | 0.7421 | 0.6013 | 0.5598 | 0.6557 | 0.5886 | 0.6099 | 0.5452 | 0.5525 | 0.5886 | 0.8444 | 0.6397 |

**Figure 1.** Cross-evaluation matrix. In the diagonal are the scores computed by the self-evaluations. Other cells show the peer-evaluations.

| Cluster | DMUs | | |
|---|---|---|---|
| | **Municipality** | **State** | **Code** |
| 1 | Alvorada | TO | #1 |
| 2 | Amambai | MS | #2 |
| 3 | Aquidauana | MS | #3 |
| 4 | Bonito | MS | #4 |
| 5 | Brasilândia | MS | #5 |
| 5 | Montes Claros | MG | #12 |
| 6 | Camapuã | MS | #6 |
| 7 | Carlos Chagas | MG | #7 |
| 8 | Catalão | GO | #8 |
| 9 | Corumbá | MS | #9 |
| 10 | Itamarajú | BA | #10 |
| 11 | Lavras do Sul | RS | #11 |
| 12 | Niquelândia | GO | #13 |
| 13 | Paraíso do Tocantins | TO | #14 |
| 14 | Porangatu | GO | #15 |
| 14 | Ribas Rio Pardo | MS | #16 |
| 15 | Rio Verde | GO | #17 |
| 15 | São Gabriel d'Oeste | MS | #18 |
| 16 | Tupã | SP | #19 |
| 17 | Uberaba | MG | #20 |
| 18 | Uberlândia | MG | #21 |

**Table 4.** Clusters obtained with the 5x5 grid.

http://dx.doi.org/10.5772/51324

The 5x5 grid provided the results shown in Table 4. There were obtained 18 clusters and almost all contained one municipality (DMU). The exceptions were cluster 5 (which contains Brasilândia and Montes Claros), cluster 14 (compounded of Porangatu and Ribas Rio Pardo), and cluster 15 (with Rio Verde and São Gabriel d'Oeste).

Figure 2 shows the resulting allocation for the 5x5 grid. The clusters with two DMUs were the ones in which the municipalities had the most similar efficiency profiles, i.e., strongest similarities. If these six DMUs hadn't met their similar pairs they would had been allocated in the empty clusters, each DMU in one cluster, instead of being grouped in pairs in three clusters.

These three clusters were formed mainly due to the similarity of the decision profile in regards to the cows' culling age in the production systems of Brasilândia and Montes Claros (8.2 and 8.5 years), Ribas Rio Pardo and Porangatu (9.2 and 9.0 years), Rio Verde and São Gabriel d'Oeste (10.3 and 10.7 years). In practice, however, the 5x5 grid didn't perform well, as it hadn't grouped other production systems with similar profiles. Thus, this grid didn't provide the basis for an accurate synthesis of convergence points among the other systems under assessment.

The 5x5 grid allocated many single municipalities in the clusters, i.e., one DMU in one cluster. Therefore, this topology was not suitable for this study. We also analyzed the 4x4 grid. We inferred that the 4x4 typology was also not adequate for this study as we obtained similar results to the 5x5 grid.

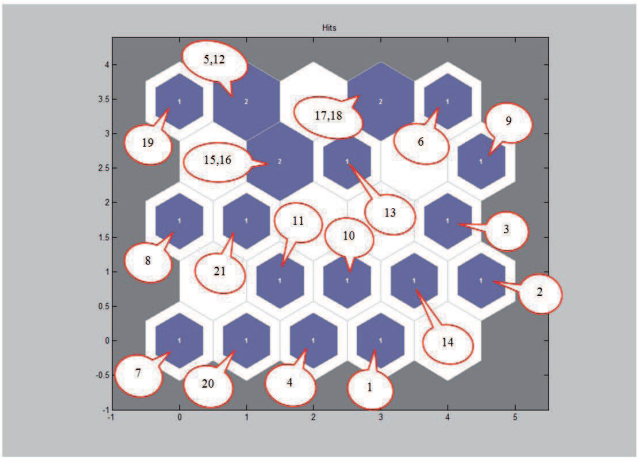


**Figure 2.** DMUs' distribution for the 5x5 grid.

The set shown in Table 5 is from the 3x3 grid. We observe that some DMUs were in the same clusters they were allocated in the 5x5 grid. This was the case of the cluster composed of Brasilândia and Montes Claros. In this cluster were added the DMUs Catalão, Ribas Rio Pardo, Tupã and Uberlândia. The same occurred with the group Rio Verde – São Gabriel d'Oeste that in the 3x3 grid was also consisted of Niquelândia and Porangatu municipalities. We may point out that Porangatu and Ribas Rio Pardo were no longer in the same cluster, which was a consequence of this new grid.

| Cluster | DMUs | | |
|---|---|---|---|
| | Municipality | State | Code |
| 1 | Alvorada | TO | #1 |
| 2 | Amambai | MS | #2 |
| 2 | Paraíso do Tocantins | TO | #14 |
| 3 | Aquidauana | MS | #3 |
| 3 | Camapuã | MS | #6 |
| 3 | Corumbá | MS | #9 |
| 4 | Bonito | MS | #4 |
| 4 | Lavras do Sul | RS | #11 |
| 5 | Brasilândia | MS | #5 |
| 5 | Catalão | GO | #8 |
| 5 | Montes Claros | MG | #12 |
| 5 | Ribas Rio Pardo | MS | #16 |
| 5 | Tupã | SP | #19 |
| 5 | Uberlândia | MG | #21 |
| 6 | Carlos Chagas | MG | #7 |
| 7 | Itamarajú | BA | #10 |
| 8 | Niquelândia | GO | #13 |
| 8 | Porangatu | GO | #15 |
| 8 | Rio Verde | GO | #17 |
| 8 | São Gabriel d'Oeste | MS | #18 |
| 9 | Uberaba | MG | #20 |

**Table 5.** Clusters obtained with the 3x3 grid.

This situation was probably a result of the same decision-making profile of the municipalities that were grouped, in regards the dispose of cows and bulls in their production systems. The 3x3 grid provided balanced groups concerning the variables, especially in relation to the profile of discarding cows and bulls. Bulls' disposal is directly related to purchase: the higher is the annual discard, it increases the need for annual replacements.

By limiting the number of clusters to nine (3x3 grid) we obtained three clusters with only one municipality: cluster 1 (Alvorada), cluster 6 (Carlos Chagas) and cluster 7 (Itamarajú). These three municipalities had unique profiles of decision- making with respect to the culling of cows and bulls. In Alvorada, cows and bulls culling rates were the lowest among the municipalities analyzed (8 and 13%, respectively). In Carlos Chagas, the rates were higher (18 and 21%). In Itamarajú, the cows disposal rate was small (12%), but the bulls disposal rate was the highest among the production systems evaluated (24%).

Figure 3 shows the DMUs' distribution for the 3x3 dimension grid.

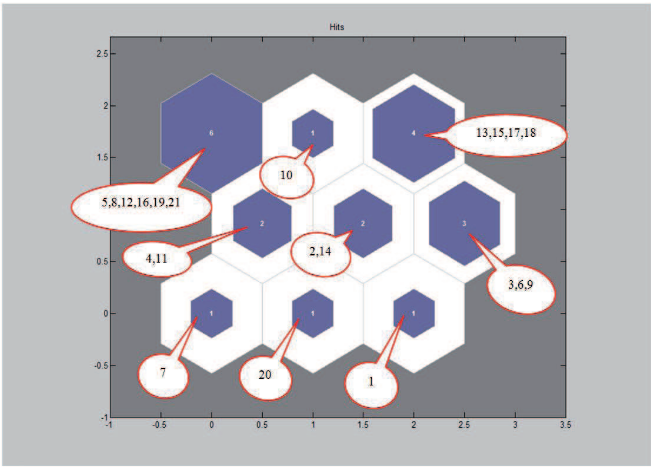


**Figure 3.** DMUs' distribution for the 3x3 grid.

The third grid used was the 2x2. The results are displayed in Table 6. Comparing these results with the ones from the 3x3 grid we observe that most of the previous allocation remained unchanged and there were mergers according to the clusters similarities. The exception is cluster 8 in the 3x3 grid (Niquelândia, Porangatu, Rio Verde and São Gabriel d'Oeste): in the current configuration its DMUs were inserted into other clusters (Niquelândia and Porangatu are in cluster 3 along with Bonito, Brasilândia, Catalão, Itamarajú, Lavras do Sul, Montes Claros, Ribas Rio Pardo, Tupã and Uberlândia; Rio Verde and São Gabriel d'Oeste were in cluster 2 along with Aquidauana, Camapuã and Corumbá).

Although these clusters were different from the ones obtained with the 3x3 grid, they were compatible with the results from 5x5 configuration, where the pairs Rio Verde – Gabriel d'Oeste and Porangatu – Ribas Rio Pardo were allocated in different groups.

| Cluster | DMUs | | |
|---|---|---|---|
| | Municipality | State | Code |
| 1 | Alvorada | TO | #1 |
| 1 | Amambai | MS | #2 |
| 1 | Paraíso do Tocantins | TO | #14 |
| 2 | Aquidauana | MS | #3 |
| 2 | Camapuã | MS | #6 |
| 2 | Corumbá | MS | #9 |
| 2 | Rio Verde | GO | #17 |
| 2 | São Gabriel d'Oeste | MS | #18 |
| 3 | Bonito | MS | #4 |
| 3 | Brasilândia | MS | #5 |
| 3 | Catalão | GO | #8 |
| 3 | Itamarajú | BA | #10 |
| 3 | Lavras do Sul | RS | #11 |
| 3 | Montes Claros | MG | #12 |
| 3 | Niquelândia | GO | #13 |
| 3 | Porangatu | GO | #15 |
| 3 | Ribas Rio Pardo | MS | #16 |
| 3 | Tupã | SP | #19 |
| 3 | Uberlândia | MG | #21 |
| 4 | Carlos Chagas | MG | #7 |
| 4 | Uberaba | MG | #20 |

**Table 6.** Clusters obtained with the 2x2 grid.

The DMUs' allocation in regards to the 2x2 grid is shown in Figure 4. The biggest cluster had 11 municipalities: Bonito, Brasilândia, Catalão, Itamarajú, Lavras do Sul, Montes Claros, Niquelândia, Porangatu, Ribas Rio Pardo, Tupã and Uberlândia. In this group, cows and bulls discard rates had medium values (19 and 14%, respectively).

The smallest group was composed of two municipalities: Carlos Chagas and Uberaba. These municipalities exhibited higher rates for the discarding of cows and bulls (18% and 21%), which is a characteristic of production systems with a more dynamics profile, when compared to other clusters.

The clusters composed of three (Alvorada, Amambai and Paraíso do Tocantins) and five municipalities (Rio Verde, Aquidauana, Camapuã, São Gabriel d'Oeste and Corumbá) had discard rates of cows and bulls of 10% and 13%, and 18% and 15%, respectively. This implies that the former group (with lower rates) probably had herds with a lower replacement dynamics. In the latter cluster, which had a lower rate of bulls' replacement, there was less investment in improving the genetic quality of the herd.

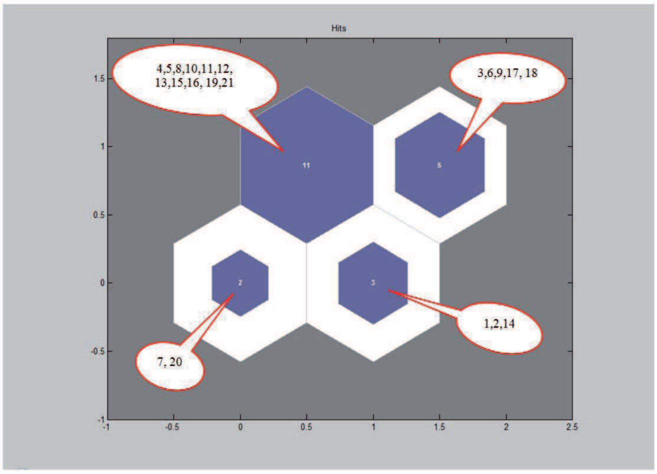


**Figure 4.** DMUs' distribution for the 2x2 grid.

Finally, the 2x1 dimension grid was used and the results are depicted in Table 7 and Figure 5 shows its DMUs' distribution.

| Cluster | DMUs | | |
|---|---|---|---|
| | Municipality | State | Code |
| 1 | Alvorada | TO | #1 |
| 1 | Amambai | MS | #2 |
| 1 | Aquidauana | MS | #3 |
| 1 | Brasilândia | MS | #5 |
| 1 | Camapuã | MS | #6 |
| 1 | Catalão | GO | #8 |
| 1 | Corumbá | MS | #9 |
| 1 | Itamarajú | BA | #10 |
| 1 | Montes Claros | MG | #12 |
| 1 | Niquelândia | GO | #13 |

| Cluster | DMUs | | |
|---|---|---|---|
| | Municipality | State | Code |
| 1 | Paraíso do Tocantins | TO | #14 |
| 1 | Porangatu | GO | #15 |
| 1 | Ribas Rio Pardo | MS | #16 |
| 1 | Rio Verde | GO | #17 |
| 1 | São Gabriel d'Oeste | MS | #18 |
| 1 | Tupã | SP | #19 |
| 1 | Uberlândia | MG | #21 |
| 2 | Bonito | MS | #4 |
| 2 | Carlos Chagas | MG | #7 |
| 2 | Lavras do Sul | RS | #11 |
| 2 | Uberaba | MG | #20 |

**Table 7.** Clusters obtained with the 2x1 grid.

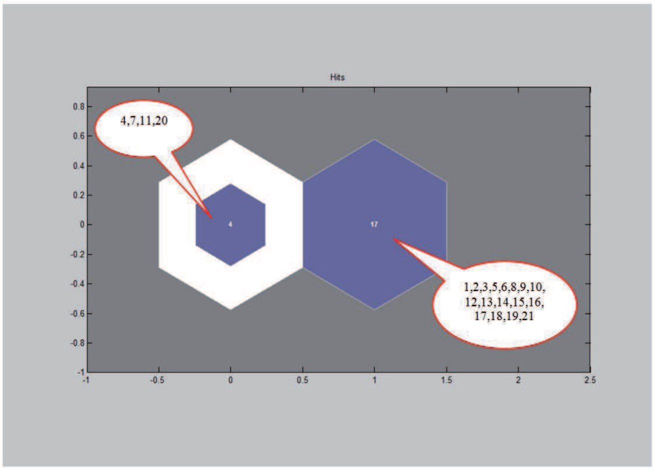


**Figure 5.** DMUs' distribution for the 2x1 grid.

We notice that a small cluster was formed with Bonito, Carlos Chagas, Lavras do Sul and Uberaba. Bonito and Lavras do Sul were part of one cluster in the 3x3 grid and in the 2x2 grid they were included in cluster 3 along with Brasilândia, Catalão, Itamarajú, Montes Claros, Niquelândia, Porangatu, Ribas Rio Pardo, Tupã and Uberlândia, which were in cluster 1. This was a result of the new grid used. Both groups had no outstanding difference between the culling rates of cows and bulls (14 and 17% in cluster 1, 14 and 18% in cluster 2).

This grid was not satisfactory to support the analysis of similarities and dissimilarities of the beef cattle production systems profiles evaluated.

## 7. Conclusions

In this paper we used Data Envelopment Analysis with SOMS to cluster the 21 beef cattle modal production systems into homogeneous groups according to their efficiency profiles. The efficiency profiles are derived from the cross-evaluation DEA approach. This approach allows a peer-evaluation instead of only considering the self-evaluation performed by the classic DEA models.

The approach used here is different from the previous ones in the literature, in which SOMS are used to group the DMUs into homogeneous sets and then use DEA models to evaluate the DMUs belonging to each cluster. In our approach rather than performing ex-ante clustering we carried out ex-post grouping. We used the DEA cross-efficiencies as inputs to the SOMS procedure. It is important to stress that we didn't use the classic DEA efficiency scores, but the efficiency profiles of each DMU derived from the cross-evaluation matrix.

The proposed approach yielded interesting results for 3x3 and 2x2 grids. In using these grids, the production systems developed in the 21 assessed municipalities were grouped, particularly, in relation to the inlet-outlet dynamics of important animals' categories of the breeding herd. The different decision-making profiles of culling of cows and bulls, and consequently of the purchase of bulls, were satisfactorily grouped relative to the efficiency profile observed in each production system.

## Acknowledgements

To National Council for Scientific and Technological Development (CNPq) for the financial support.

## Author details

João Carlos Correia Baptista Soares de Mello[1*], Eliane Gonçalves Gomes[2],
Lidia Angulo Meza[1], Luiz Biondi Neto[3], Urbano Gomes Pinto de Abreu[4],
Thiago Bernardino de Carvalho[5] and Sergio de Zen[5]

*Address all correspondence to: jcsmello@producao.uff.br

1 UFF. Passo da Pátria 156, Brazil

2 Brazilian Agricultural Research Corporation (Embrapa), Brazil

3 UERJ. Rua São Francisco Xavier 524, Brazil

4 Embrapa Pantanal. Caixa Postal 109, Brazil

5 Center for Advanced Studies on Applied Economics (CEPEA/ESALQ/USP), Brazil

## References

[1] Charnes, A., Cooper, W. W., & Rhodes, E. (1978). Measuring the efficiency of decision-making units. *European Journal of Operational Research*, 2, 429-44.

[2] Samoilenko, S., & Osei-Bryson, K. M. (2010). Determining sources of relative inefficiency in heterogeneous samples: Methodology using Cluster Analysis, DEA and Neural Networks. *European Journal of Operational Research*, 206(2), 479-87.

[3] Sharma, M. J., & Yu, S. J. (2009). Performance based stratification and clustering for benchmarking of container terminals. Expert Systems with Applications PART 1) , 5016-5022.

[4] Churilov, L., & Flitman, A. (2006). Towards fair ranking of olympics achievements: The case of Sydney 2000. *Computers and Operations Research*, 33(7), 2057-2082.

[5] Biondi, Neto. L., Lins, M. P. E., Gomes, E. G., Soares de Mello., J. C. C. B., & Oliveira, F. S. (2004). Neural data envelopment analysis: a simulation. *International Journal of Industrial Engineering*, 11, 14-24.

[6] Emrouznejad, A., & Shale, E. (2009). A combined neural network and DEA for measuring efficiency of large scale datasets. *Computers and Industrial Engineering*, 56(1), 249-54.

[7] Çelebi, D., & Bayraktar, D. (2008). An integrated neural network and data envelopment analysis for supplier evaluation under incomplete information. *Expert Systems with Applications*, 35(4), 1698-710.

[8] Angulo-Meza, L., Biondi Neto., L., Brandão, L. C., Andrade, F. V. S., Soares de Mello, J. C. C. B., & Coelho, P. H. G. (2011). Modelling with self-organising maps and data envelopment analysis: A case study in educational evaluation. *Self organizing maps, new achievements*, Vienna, Intech, 71-88.

[9] Paschoalino, F. F., & Soares de, Mello. J. C. C. B., Angulo-Meza L., Biondi Neto L. (2011). DMU clustering based on Cross Efficiency Evaluation. International Conference on Data Envelopment Analysis and Its Applications to Management. Lima (Peru).

[10] Sexton, T. R., Silkman, R. H., & Logan, A. J. (1986). Data Envelopment Analysis: Critique and extensions. In: Silkman H, ed. Measuring efficiency: An assessment of data envelopment analysis. San Francisco Jossey-BassEditor: , 73-105.

[11] Banker, R. D., Charnes, A., & Cooper, W. W. (1984). Some models for estimating technical scale inefficiencies in data envelopment analysis. *Mngmt Sc.*, 30(9), 1078-92.

[12] Doyle, J. R., & Green, R. H. (1994). Efficiency and cross-efficiency in DEA derivations, meanings and uses. *Journal of the Operational Research Society*, 45, 567-78.

[13] Doyle, J. R., & Green, R. H. (1995). Cross-Evaluation in DEA: Improving Discrimination Among DMUs. *Information Systems and Operational Research*, 33(3), 205-22.

[14] Appa, G., Argyris, N., & Williams, H. P. (2006). A methodology for cross-evaluation in DEA. Working Paper: London School of Economics and Political Science November 7.

[15] Soares de, Mello. J. C. C. B., Lins, M. P. E., & Gomes, E. G. (2002). Construction of a smoothed DEA frontier. *Pesquisa Operacional.*, 28(2), 183-201.

[16] Wu, J., Liang, L., & Chen, Y. (2009). DEA game cross-efficiency approach to Olympic rankings. *Omega*, 37(4), 909-18.

[17] Gomes, E. G., Soares de Mello, J. C. C. B., Angulo-Meza, L., Silveira , J. Q., Biondi Neto, L., & Abreu, U. G. P. d. (2012). Some Remarks About Negative Efficiencies in DEA Models. In: Holtzman Y, ed. Advanced Topics in Applied Operations Management. Rijeka InTech , 113-32.

[18] Liang, L., Wu, J., Cook, W. D., & Zhu, J. (2008). The DEA game cross-efficiency model and its nash equilibrium. *Operations Research*, 56(5), 1278-88.

[19] 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), 1025-30.

[20] Mitra, P., Murthy, C. A., & Pal, S. K. (2002). Unsupervised feature selection using feature similarity. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24(3), 301-12.

[21] Willshaw, D. J., Buneman, O. P., & Longuet-Higgins, H. C. (1969). Non-holographic associative memory. *Nature*, 222, 960-2.

[22] Willshaw, D. J., & Von der, Malsburg. C. (1976). How patterned neural connections can be set up by self-organization. *Proceedings of the Royal Society of London Series B*, 194, 431-45.

[23] Kohonen, T. (2001). Self-organizing maps. rd ed. ed. Berlin: Springer-Verlag.

[24] Haykin, S. (1999). Neural networks: a comprehensive foundation. nd ed. ed. New Jersey Prentice Hall

[25] Bishop, C. M. (1995). Neural networks for pattern recognition. New York, Oxford University Press.

[26] Wu, J., Liang, L., Wu, D., & Yang, F. (2008). Olympics ranking and benchmarking based on cross efficiency evaluation method and cluster analysis: the case of Sydney 2000. *International Journal of Enterprise Network Management*, 2(4), 377-92.

[27] Plaxico, J. S., & Tweeten, L. G. (1963). Representative farms for policy and projection research. *Journal of Farm Economics*, 45, 1458-65.

[28] Centro de Estudos Avançados em Economia Aplicada. (2010). Metodologia do índice de preços dos insumos utilizados na produção pecuária brasileira. [cited 2010 March, 24]; Available from: http://www.cepea.esalq.usp.br/boi/metodologiacna.pdf

[29] Gomes, E. G., Soares de Mello, J. C. C. B., Souza, L., Angulo-Meza, G. D. S., & Mangabeira, J. A.d. C. (2009). Efficiency and sustainability assessment for a group of farmers in the Brazilian Amazon. *Annals of Operations Research*.

[30] Womer, N. K., Bougnol, M. L., Dulá, J. H., & Retzlaff-Roberts, D. (2006). Benefit-cost analysis using data envelopment analysis. *Annals of Operations Research*, 145(1), 229-50.

[31] Kuosmanen, T., & Kortelainen, M. (2007). Valuing environmental factors in cost-benefit analysis using data envelopment analysis. *Ecological Economics*, 62(1), 56-65.

[32] Kuosmanen, T., Bijsterbosch, N., & Dellink, R. (2009). Environmental cost-benefit analysis of alternative timing strategies in greenhouse gas abatement: A data envelopment analysis approach. *Ecological Economics*, 68(6), 1633-42.

[33] Bougnol, M. L., Dulá, J. H., Estellita, Lins. M. P., & Moreira da Silva, A. C. (2010). Enhancing standard performance practices with DEA. *Omega*, 38(1-2), 33-45.

[34] 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.