CONFIDENCE INTERVALS FOR DEA EFFICIENCY MEASUREMENTS APPLIED TO EMBRAPA'S RESEARCH SYSTEM: A BOOTSTRAP APPROACH

Geraldo da Silva e Souza
Mirian Oliveira de Souza
Eliane Gonçalves Gomes
Brazilian Agricultural Research Corporation (Embrapa)
Parque Estação Biológica – PqEB, Av. W3 Norte final, 70770-901, Brasília, DF
{geraldo.souza; mirian.souza; eliane.gomes}@embrapa.br

ABSTRACT

We define and model the research production at Embrapa, the major Brazilian institution responsible for applied agricultural research. The main theoretical framework is Data Envelopment Analysis – DEA. We explore the economic interpretation and the statistical properties of these models to compute confidence intervals for output oriented efficiency measurements. It was based on a parametric flexible model, defined by the truncated normal distribution. Intervals are calculated exploring a bootstrap approach. These results provide a better insight on the efficiency classification and allow comparisons among the DMUs involved in the evaluation process, taking into account random errors and inefficiency random variation.

KEYWORDS. Data envelopment analysis. Confidence intervals. Bootstrap. Agricultural research.

RESUMO

Neste artigo foi estudado o sistema de produção de pesquisa da Embrapa, a maior instituição brasileira de pesquisa agropecuária. A principal ferramenta teórica usada foi a Análise de Envoltória de Dados – DEA. Exploraram-se a interpretação econômica e as propriedades estatísticas desses modelos, para calcular intervalos de confiança para medidas de eficiência orientadas a *output*. Tomou-se como base um modelo paramétrico flexível, definido pela distribuição normal truncada. Os intervalos foram calculados por reamostragem. Estes resultados geraram melhores entendimentos sobre as medidas de eficiência e permitiram comparações entre as DMUs envolvidas na análise. O modelo considerou erros de ineficiência e erros aleatórios.

PALAVRAS-CHAVE. Análise de Envoltória de Dados. Intervalos de confiança. Reamostragem. Pesquisa Agropecuária.

1. Introduction

It is of importance to administrators of research institutions to have at their disposal measures and procedures making feasible an evaluation of the quantum of production, as well as the technical efficiency of the production process of their institutions. In times of competition and budget constraints, a research institution needs to know by how much it may increase its production, without absorbing additional resources. The quantitative monitoring of the production process allows for an effective administration of the resources available and the observation of predefined research patterns and goals. In this context, the Brazilian Agricultural Research Corporation (Embrapa) developed a production model based on the input-output data of its research units. The theoretical framework for this model is the analysis of production frontiers, known as Data Envelopment Analysis (DEA).

Several uses are made of the efficiency measurements by Embrapa's administration. Those include monitoring of production targets, resource allocation and rewarding. Administrative actions regarding a given ranking of units will have more impact if they take into account the stochastic variation imbedded in the measurements of production variables. This leads to the consideration of statistical production models, from which one may infer statistical properties for efficiency estimates. For the stochastic frontier analysis, with proper parametric specifications of the production or cost functions, this is a natural process, as can be seen in Kumbhakar and Lovell (2000) and Coelli et al. (2005).

For the nonparametric frontier approaches induced by classical DEA (Coelli et al., 2005) or the Free Disposal Hull of Deprins et al. (1984), some technical issues arise and a proper approach has to be put forward to guarantee the derivation of sound statistical results. This is the line of work carried out by Banker (1993), Banker and Natarajan (2004, 2008), Simar and Wilson (2004, 2007), Daraio and Simar (2007), Souza and Staub (2007) and Souza et al. (2009a).

In this article we combine the results of Banker (1993), Banker and Natarajan (2008), Simar and Wilson (2007) and Souza and Staub (2007) to come up with confidence intervals for DEA efficiency measurements, robust relative to production function choices and efficiency distributions within reason. These intervals are more appealing than those generated by the bootstrap of Simar and Wilson (2004, 2007) that may produce unexpected results, like one unit being regarded as inefficient after being observed as a benchmark or generating confidence limits that do not include observed efficiency measurements.

Our discussion proceeds as follows. In Section 2 we review the concepts leading to the models for which one may view DEA estimates as nonparametric maximum likelihood, and for which statistical properties may be derived for efficiency estimates. In Section 3 we review Embrapa's production model. Section 4 deals with the statistical results of our application and, finally, in Section 5 we summarize our findings.

2. DEA Production Models

Consider a production process composed of *n* decision making units (DMUs). Each DMU uses varying quantities of *m* different inputs to produce varying quantities of *s* different outputs.

Denote by $Y = (y_1, y_2, ..., y_n)$ the $s \times n$ production matrix of the *n* DMUs. The *r*th column of *Y* is the output vector of DMU *r*. Denote by $X = (x_1, x_2, ..., x_n)$ the $m \times n$ input matrix. The *r*th column of *X* is the input vector of DMU *r*. The matrices $Y = (y_{ij})$ and $X = (x_{ij})$ must satisfy: $p_{ij} \geq 0, \sum_i p_{ij} > 0$ and $\sum_j p_{ij} > 0$, where p is x or y.

The measure of technical efficiency of production (under constant returns to scale) for DMU $o \in \{1, 2, ..., n\}$, denoted $\mathrm{E}^{CR}(o)$, is the solution of the linear programming problem (1).

$$\mathrm{E}^{\mathrm{CR}}(o) = \max_{u,v} \frac{y_o' u}{x_o' v}$$

subject to (1)

i) $x_o' v = 1$, ii) $y_j' u - x_j' v \leq 0,\ j = 1,2,\ldots,n$ and iii) $u \geq 0, v \geq 0$

If we look at the coefficients u and v as input and output prices, we see that the measure of technical efficiency of production is very close to the notion of productivity (output income/input expenditure). Technical efficiency, in this context, basically, is looking for the price system (u,v) for which DMU *o* achieves the best relative productivity ratio.

The dual problem of the linear programming problem (1) has an important economic interpretation, which we will explore. This is $\min_{\theta,\lambda} \theta$, subject to $\begin{pmatrix} 0 & Y \\ x_o & -X \\ 0 & I \end{pmatrix} \begin{pmatrix} \theta \\ \lambda \end{pmatrix} \geq \begin{pmatrix} y_o \\ 0 \\ 0 \end{pmatrix}$ or, equivalently, formulation (2).

$$\min_{\theta,\lambda} \theta$$

subject to (2)

i) $Y\lambda \geq y_o$, ii) $X\lambda \leq \theta x_o$ and iii) $\lambda \geq 0, \theta$ free

The matrix products $Y\lambda$ and $X\lambda$, with $\lambda \geq 0$, represent linear combinations of the columns of *Y* and *X*, respectively, i.e., a sort of weighted averages of output and input vectors. In this way, for each λ we can generate a new production relation, a new “pseudo” producer. Trivially, the set of DMUs 1, 2,..., *n* are included among those new producers. Making allowance for these newly defined production relationships, the question that the dual intends to answer is: What proportional reduction of inputs θx_o it is possible to achieve for DMU *o* and still produce at least output vector y_o? The solution $\theta^*(x_o, y_o)$ is the smallest θ with this property.

We can define the concept of technical efficiency of production in a context of fixed inputs instead of fixed outputs, i.e., in a program of output augmentation. In this environment the measure of technical efficiency of production of DMU *o*, under constant returns to scale (CCR model – Charnes et al., 1978), is the one defined in (3).

$$\phi_o^* = \max_{\phi,\lambda} \phi$$

subject to (3)

i) $Y\lambda \geq \phi y_o$, ii) $X\lambda \leq x_o$ and iii) $\lambda \geq 0$, ϕ free

In the output augmentation program the question we ask is: what proportional rate ϕ can be uniformly applied to augment the output vector y_o, without increasing the input vector x_o? The solution ϕ^* is the largest ϕ with this property. This is the approach we will explore here.

Questions of scale can be dealt properly imposing proper restrictions in the linear programming problem. One obtains the variable returns DEA imposing the additional condition $1'\lambda = 1$ on the weight vector λ. This is called the BCC model (Banker et al., 1984).

We now turn our attention to production statistical models. We follow along the lines of Banker (1993). Suppose *s* = 1 (a single output) and assume the existence of a continuous frontier production function $g: K \to R$ defined on the convex and compact subset *K* of the positive orthant of R^s. For each DMU *o*, the input observations x_o are points in *K*. Let (4).

$$K^* = \left\{ x \in K; x \geq \sum_{i=1}^{n} \lambda_j x_j, \lambda_j \geq 0, \sum_{i=1}^{n} \lambda_j = 1 \right\} \quad (4)$$

The DEA frontier production function is defined for $x \in K^*$ by (5), and it can be shown that for DMU *o*, $g^*(x_o) = \phi_o^* y_o$, where we are assuming variable returns to scale.

$$g^*(x) = \sup_{\lambda} \left\{ \sum_{j=1}^{n} \lambda_j y_j ; x \geq \sum_{i=1}^{n} \lambda_j x_j, \lambda_j \geq 0, \sum_{i=1}^{n} \lambda_j = 1 \right\} \quad (5)$$

Suppose that observations (x_o, y_o) are interior points to K and that they are generated in accordance with the statistical model (6) (Banker, 1993).

$$y_o = g(x_o) - \varepsilon_o \quad (6)$$

In (6):

a) The inefficiencies ε_o are *iid* with a common density $f(\varepsilon)$.
b) The common distribution function $F(x)$ of the inefficiencies is strictly positive in $(0,+\infty)$.
c) The inputs x_o represent a random sample from a density $h(x)$ strictly positive in the interior of K.
d) The inputs x_o and the inefficiencies are independent.

Then (Banker, 1993):

1. $g^*(x_o)$ is the nonparametric maximum likelihood estimate of $g(x_o)$ if $f(\varepsilon)$ is monotonically decreasing in $(0,+\infty)$.
2. $g^*(x_o)$ is weakly consistent for $g(x_o)$.
3. Let M be any fixed subset of DMUs. If n is large, the joint distribution of the estimated inefficiencies $\hat{\varepsilon}_j = y_j - g^*(x_j),\ j \in M$, is, approximately, the joint distribution of the true inefficiencies $\varepsilon_j,\ j \in M$.

These results can be used to test hypothesis about the nature of the production process. An example is the verification of whether the technology shows constant returns to scale or variable returns to scale. We may perform this test comparing the empirical distribution functions of the $\hat{\varepsilon}_o$ (estimated inefficiency errors) under the assumptions of constant and variable returns to scale computing Kolmogorov-Smirnov test statistic (Conover, 1998).

The statistical properties of univariate DEA estimates were extended by Souza and Staub (2007) to allow for non *iid* inefficient components. Suppose in this context that the sequence of pairs (x_j, ε_j), satisfying the statistical model $y_j = g(x_j) - \varepsilon_j$, are drawn independently from the product probability density functions $h_j(x) f_j(\varepsilon)$, where the sequence of input densities $h_j(x)$ satisfies $0 < l(x) \leq \inf_j h_j(x) \leq \sup_j h_j(x) \leq L(x)$, for integrable functions $l(x)$ and $L(x)$, and x interior to K. The inefficiency densities $f_j(\varepsilon)$ are such that $F(u) = inf_j F_j(u) > 0,\ u > 0$, where $F_j(u) = \int_0^u f_j(\varepsilon) d\varepsilon$.

Then if x_0 is a point in K^* interior to K, $g^*(x_0)$ converges almost surely to $g(x_0)$. The property stated in 3. above is still true. In this context let $\hat{q}$ be the quantile of order $100(1-\alpha)\%$ for residual distribution corresponding to production observation (x_j, y_j). The interval

$[g^*(x_j), g^*(x_j)+\hat{q}_i]$ has asymptotically level $1-\alpha$ for $g(x_j)=\phi_j y_j$. Dividing the one sided interval by y_j provides an interval for ϕ_j.

Following Banker and Natarajan (2004, 2008) we assume now that observations $\left(x_j, y_j\right)$ on production follow the statistical model (7),

$$y_j = g\left(x_j\right) + v_j - u_j \quad j = 1 \ldots n \tag{7}$$

where $g(.)$ is a continuous production function. The random variables v_j and u_j represent random and inefficiency errors, respectively. The random errors have a two sided continuous distribution with a non null density function on $\left(-V^M, V^M\right)$. The inefficiency error component is positive. Thus (8):

$$\begin{aligned} y_j &= g\left(x_j\right) + V^M - (V^M - v_j + u_j) \\ y_j^a &= y_j - V^M = g\left(x_j\right) - \varepsilon_j \end{aligned} \tag{8}$$

The component ε_j is strictly positive. Since identical distributions are not required and one may let the mean μ of the inefficiency distribution be dependent on a linear function $\delta' z$ of covariates or contextual variables. We consider a two-stage statistical procedure to estimate δ, V^M, as well as any other parameters indexing the inefficiency distribution. Motivated by Simar and Wilson (2007), we use maximum likelihood methods to estimate inefficiency errors parameters from the model $y_j = \tilde{g}(x_j) - \varepsilon_j$, where $\tilde{g}(x_j) = g(x_j) + V^M$. Notice that $\tilde{g}(x)$ is an unknown production function. Only the inefficient firms enter in this analysis. For example one may fit a gamma distribution $\Gamma(p, \lambda_j)$ with mean $\mu_j = p/\lambda_j$, where $\lambda_j = \exp(-\delta' z_j)$, by maximum likelihood, to DEA residuals $\hat{\varepsilon}_j = \left(\phi_j^* - 1\right) y_j$, or alternatively use the truncated normal $N^+(\mu_j, \sigma^2)$. We obtain information on the constant V^M assuming that the efficient units are producing on the technological frontier. In this context an optimum estimate would be $\hat{V}^M = \sum_{l=1}^{n_l} \hat{\mu}_l / n_l$, where $\hat{\mu}_l$ is the maximum likelihood estimate of μ_l and the sum is over the efficient units.

The maximum likelihood estimate of μ_l is computed from the inefficient units. This is a subtle modification on the methods proposed by Banker and Natarajan (2008). The use of the gamma and truncated normal distributions and the adaptation of the procedure of Simar and Wilson (2007) to the present instance are original contributions. Notice that once V^M is estimated, the output observations should be adjusted for the inefficient units by subtracting this quantity, and new efficiency measurements are then computed to estimate the unknown production function $g(x)$ and the corresponding set of technical efficiency scores.

Checking whether or not the production model fits the data is a matter of verifying if the postulated inefficiency distribution fits the efficiency observations. As in stochastic frontier analysis, the three commonly used family of distributions used to model inefficiency errors are the exponential, the half-normal and the truncated normal, the latter having flexibility properties (Coelli et al., 2005). Goodness of fit statistics can be used to assess the best fit.

To compute standard error for the technical efficiency estimates one should use bootstrap estimates. Shorter confidence limits are obtained using this approach. In this context we suggest here an adaptation of the algorithm proposed by Löthgren and Tambour (1999). This is achieved following the steps:

1. Let $\hat{\phi}_i^{*,b} = 1 + \varepsilon_i^b / y_i$, where $\varepsilon_i^{*,b}$ is generated from the error model fitted by maximum likelihood.
2. Let the bootstrap pseudo-data be given by $\left(x_i, [\phi_i^{*,b} / \hat{\phi}_i] y_i\right)$, $i = 1 \ldots n$.
3. Estimate bootstrap efficiencies $\zeta_i^{*,b}$ using the BCC version of the linear programming (3) for the pseudo data.
4. Repeat 1-3 B times to create a set of B firm-specific bootstrapped efficiency estimate $\zeta_i^{*,b}$, $b = 1 \ldots B$. Compute the efficiency standard errors with equation (9) (Efron and Tibshirani, 1986, 1993).

$$s_i = \sqrt{\frac{1}{B-1} \sum_{b=1}^{B} \left(\zeta_i^{*,b} - \frac{1}{B} \sum_{b=1}^{B} \zeta_i^{*,b} \right)^2} \tag{9}$$

With knowledge of the standard errors, one may compute the one-sided confidence intervals as $[\hat{\phi}_i, \hat{\phi}_i + z_{1-\alpha} s_i]$, where $z_{1-\alpha}$, $\alpha \in (0,1)$ is the standard normal quantile of order $1-\alpha$.

3. Embrapa's Production System

Embrapa's research system comprises 37 units (DMUs) of research centers. Input and output variables are defined from a set of performance indicators known to the company since 1991. The set comprises 28 outputs and 3 inputs.

We begin our discussion with the output. The output variables are classified into four categories: Scientific production; Production of technical publications; Development of technologies, products, and processes; Diffusion of technologies and image.

By scientific production we mean the publication of articles and book chapters aimed mainly to the academic world. We require each item to be specified with complete bibliographical reference.

The category of technical publications groups publications produced by research centers aiming, primarily, at agricultural businesses and agricultural production.

The category of development of technologies, products and processes groups indicators related to the effort made by a research unit to make its production available to society in the form of a final product. Only new technologies, products and processes are considered. Those must be already tested at the client's level in the form of prototypes, or through demonstration units, or be already patented.

The category of diffusion of technologies and image encompasses production variables related to Embrapa's effort to make its products known to the public and to market its image.

The input side of Embrapa's production process is composed of three variables. Personnel expenditure, Operational costs (consumption materials, travel and services less income from production projects), and Capital (measured by depreciation).

All output variables are measured as counts and normalized by the mean. Likewise, the inputs are normalized by the mean. As a final output we take a weighted average of all variables in all categories of production. The weights are user defined and reflect the administration perception of the relative importance of each variable to each research center or DMU. A single output allows the use of the statistical tests described in the previous section.

Defining weights is a hard and questionable task. In our application in Embrapa we followed an approach based on the Law of Categorical Judgments of Thurstone (Torgerson, 1958). More details on this issue may be seen in Gomes and Souza (2008). The model is competitive with the

AHP method of Saaty (1994) and is well suited when several judges are involved in the evaluation process. Basically, we sent out about 500 questionnaires to researchers and administrators and asked them to rank in importance – scale from 1 to 5, each production category and each production variable within the corresponding production category. A set of weights was determined under the assumption that the psychological continuum of the responses projects onto a normal distribution.

DEA models implicitly assume that the DMUs are comparable. This is not strictly the case in Embrapa. To make them comparable it is necessary an effort to define an output measure adjusted for differences in operation and perceptions. At the level of the partial production categories we induced this measure allowing a distinct set of weights for each DMU. In principle one could go ahead and use DEA with multiple outputs. This would minimize the effort of defining weights leaving to DEA the task of finding these coefficients. The problem with such approach is that there is a kind of dimensionality curse in DEA models. As the number of factors (inputs and outputs) increases, the ability to discriminate between DMUs decreases. Thus we found convenient to extend the weight system to produce a single measure of output y_o.

A personnel score was created for each unit dividing its number of employees by the company's mean. Outputs and inputs were normalized by this variable. This further established a common basis to compare research units (regarding scale) and avoided the incidence of spurious efficient units and zero output (shadow) prices, another common occurrence in multiple output models, and also a disturbing fact for management interpretation. It's important to mention that at Embrapa the dedication is the same for all employees, and they are grouped into two categories: Research and Research Support.

We see the use of ratios to define production variables in our application as unavoidable. Different denominators are used with the virtue of being independent of sizes of the units. Therefore facilitate comparisons between units and the postulation of a common production function. In the context of a pure DEA analysis, the problem of efficiency comparisons may be solved imposing the BCC assumption. See Hollingsworth and Smith (2003) and Emrouznejad and Amin (2007). These authors state that when using ratio variables, the constant returns to scale assumption is not valid.

DEA models are known to be sensitive to outliers. In our application, the control of outliers is particularly important for output variables. In this context we use box plot fences to adjust the values of outlying observations. Values above $Q3+1.5(Q3-Q1)$ are reduced to this value for any variable. Here $Q1$ and $Q3$ denote the first and third quartiles, respectively. This approach led to the exclusion of a research unit from our analysis (DMU 25), leaving in 36 the total number of research centers.

4. Statistical Results

Table 1 shows Embrapa's production data for 2008. Table 2 shows goodness of fit statistics related to the fit of the truncated normal and gamma distributions. The best fit is the truncated normal. The models were estimated by maximum likelihood including two contextual variables: size and type of unit. Statistical results for the truncated normal are shown in Table 3. We see that size and type are non significant effects.

Notice that the log likelihood function for the gamma distribution is given by (10). Here $\Gamma(.)$ is the gamma function (11), where z_1 and z_2 are size dummy variables, and z_3 and z_4 are type dummy variables.

$$L(\alpha,\beta_1,\beta_2,\beta_3,\beta_4,p) = p\sum_{j=1}^{m}\log(\lambda_j) + (p-1)\sum_{j=1}^{m}\log(\varepsilon_j) - \sum_{j=1}^{m}\varepsilon_j\lambda_j - m\log(\Gamma(p)) \quad (10)$$

$$\lambda_j = \exp(-\mu_j), \quad \mu_j = \alpha + \beta_1 z_{1j} + \beta_2 z_{2j} + \beta_3 z_3 j + \beta_4 z_4 j \quad (11)$$

Table 1. Embrapa's production variables, 2008. *Y* is the output, *X*1 is personnel expenditures, *X*2 is other expenses and *X*3 is capital

Research center	*X1*	*X2*	*X3*	*Y*
DMU1	1.4285	1.7996	1.9869	1.4648
DMU2	1.0406	0.8825	0.7318	1.1044
DMU3	0.9636	1.2486	1.2576	1.5316
DMU4	0.8268	0.8671	0.8701	0.5759
DMU5	1.1502	1.4076	1.3486	1.6424
DMU6	0.9569	1.1446	0.8559	0.8694
DMU7	1.2329	1.4955	3.2180	1.6424
DMU8	1.3082	0.9091	0.8137	1.2856
DMU9	0.9549	1.2279	1.7311	0.6554
DMU10	1.0142	0.6820	1.1481	0.8453
DMU11	0.8218	0.8497	0.8237	0.7458
DMU12	1.1466	0.7993	1.2703	0.9808
DMU13	1.0392	1.0231	0.7336	0.9043
DMU14	0.9563	1.1583	0.9729	1.0867
DMU15	1.2210	0.8054	1.0168	0.6387
DMU16	0.9203	0.8012	0.9762	0.7622
DMU17	0.9700	1.3267	1.0159	0.8086
DMU18	1.0105	0.9849	0.9225	0.8708
DMU19	1.3536	1.1611	1.3020	1.6424
DMU20	0.8867	1.1050	1.4015	0.9909
DMU21	0.9520	0.6988	0.6651	0.5428
DMU22	0.8707	0.8008	0.7752	0.6733
DMU23	0.9221	0.8977	0.9196	1.3127
DMU24	1.0654	1.2284	1.0347	0.6907
DMU25	1.0959	1.2664	0.6881	0.3206
DMU26	0.8993	0.6251	0.5248	0.8848
DMU27	0.9960	1.0876	1.2457	1.1711
DMU28	0.7806	0.7354	0.4612	0.6676
DMU29	0.9650	1.0785	0.8520	0.9645
DMU30	1.1376	0.9332	0.8409	0.9248
DMU31	1.0952	1.0729	0.7107	0.8512
DMU32	0.7762	0.9055	0.6025	1.0238
DMU33	1.0156	0.8010	0.6685	0.6344
DMU34	1.1596	1.1730	1.2072	1.0884
DMU35	0.9298	1.2288	1.0984	0.9716
DMU36	1.0945	1.2404	1.7440	1.0571
DMU37	1.2123	1.2335	3.0235	1.2627

Table 2. Goodness of fit statistics for inefficiency errors

Distribution	-2ll	AIC	BIC
Truncated normal	-128.3	-116.3	-108.6
Gamma	-20.3	-8.3	-0.6

Table 3. Maximum likelihood estimates for the truncated normal distribution

Parameter	Estimate	Standard error	p-value
z1 (size)	0.0524	0.0634	0.4161
z2 (size)	0.0611	0.0587	0.3069
z3 (type)	0.0335	0.0589	0.5738
z4 (type)	-0.0384	0.0615	0.5374
α (constant)	0.3412	0.0533	<0.0001
ρ	0.1170	0.0114	<0.0001
V^M	0.3907	0.0347	

The log likelihood function for the product of truncated normals $N^+(\mu_j,\sigma)$ is given by (12). Here $\phi(.)$ denotes the probability density function of the standard normal distribution and $\Phi(.)$ the distribution function.

$$L(\alpha,\beta_1,\beta_2,\beta_3,\beta_4,\sigma)=\sum_{j=1}^{m}\ln\left(\frac{\phi\left(\frac{\varepsilon_j-\mu_j}{\sigma}\right)}{\sigma\,\Phi\left(\mu_j/\sigma\right)}\right) \tag{12}$$

Table 4 shows final efficiency estimates under the assumption of variable returns to scale, their standard errors and 95% parametric bootstrap individual confidence intervals. The confidence intervals are shown within [0,1], since efficiency measurements are usually presented between zero and 1.

The constant B, of the firm-specific bootstrapped efficiency estimates $\zeta_i^{*,b}$, was chosen to be 2,000 in our study. According to Efron and Tibshirani (1986, 1993), setting B=1,000 is enough to ensure that the confidence intervals converge adequately.

The main advantage of the technique used here relative to other confidence intervals suggested in the literature, as for example the proposal of Simar and Wilson (2004, 2007), is that the actual efficiency estimates are lower bounds and they define real possibilities for the corresponding population values. The specification of random errors actually reduces efficiency estimates for inefficient units. This would not happen for the purely deterministic frontier specification.

The evaluation process at Embrapa is carried out on a year basis with many changes making it difficult to properly model dependency through time. To assess the evolution of raw efficiency measurements through time we considered a dynamic panel data model of Arellano and Bond (1991), Arellano and Bover (1995) e Blundell and Bond (1998), as in (13).

$$eff_{it}=c+\beta_1 eff_{i(t-1)}+\beta_2 z_{1t}+\beta_3 z_{2t}+\beta_4 z_{3t}+\beta_5 z_{4t}+\beta_6 time_t+\tau_i+\varepsilon_{it} \tag{13}$$

Table 4. Adjusted product oriented measures of technical and [0,1] scaled 95% confidence limits. Italic rows mean unitary upper bound

Research center	Standard deviation (*)	Efficiency (**)	Lower bound	Upper bound	Interval length
DMU1	0.079	1.529	0.603	0.654	0.051
DMU2	0.087	1.634	0.563	0.612	0.049
DMU3	*0.064*	*1.000*	*0.906*	*1.000*	*0.094*
DMU4	0.415	5.800	0.154	0.172	0.018
DMU5	0.054	1.000	0.919	1.000	0.081
DMU6	0.139	2.643	0.348	0.378	0.030
DMU7	*0.051*	*1.000*	*0.923*	*1.000*	*0.077*
DMU8	*0.067*	*1.000*	*0.901*	*1.000*	*0.099*
DMU9	0.345	5.691	0.160	0.176	0.016
DMU10	0.212	2.143	0.402	0.467	0.065
DMU11	0.217	2.927	0.305	0.342	0.037
DMU12	0.124	1.963	0.462	0.510	0.048
DMU13	0.134	2.289	0.399	0.437	0.038
DMU14	0.099	1.946	0.474	0.514	0.040
DMU15	0.296	4.709	0.193	0.212	0.019
DMU16	0.205	3.126	0.289	0.320	0.031
DMU17	0.160	3.316	0.280	0.302	0.022
DMU18	0.143	2.762	0.334	0.362	0.028
DMU19	*0.057*	*1.000*	*0.914*	*1.000*	*0.086*
DMU20	0.123	2.203	0.416	0.454	0.038
DMU21	0.601	6.577	0.132	0.152	0.020
DMU22	0.226	3.791	0.240	0.264	0.024
DMU23	*0.073*	*1.000*	*0.894*	*1.000*	*0.106*
DMU24	0.204	4.713	0.198	0.212	0.014
DMU26	*0.131*	*1.000*	*0.823*	*1.000*	*0.177*
DMU27	0.085	1.861	0.500	0.537	0.037
DMU28	*0.176*	*1.000*	*0.776*	*1.000*	*0.224*
DMU29	0.115	2.201	0.419	0.454	0.035
DMU30	0.121	2.399	0.385	0.417	0.032
DMU31	0.155	2.514	0.361	0.398	0.037
DMU32	*0.116*	*1.000*	*0.841*	*1.000*	*0.159*
DMU33	0.272	4.441	0.205	0.225	0.020
DMU34	0.090	2.217	0.423	0.451	0.028
DMU35	0.133	2.440	0.376	0.410	0.034
DMU36	0.105	2.363	0.394	0.423	0.029
DMU37	0.077	1.851	0.506	0.540	0.034
Average	-	2.529	0.484	0.539	0.055

(*) Based on "Efficiency" column.
(**) Output oriented DEA BCC model. Output adjusted.

In this model i denotes a research center, t is the year under analysis and *time* is a time trend. The z's are dummies for size (z_1 and z_2) and type (z_3 and z_4). The components τ and ε are independent random errors, possibly showing heteroscedasticity and first order autocorrelation in the component ε. Table 5 shows the statistical results found. We see that there is no second order residual autocorrelation, and size and type are not significant effects. There is no permanent residual efficiency effect, since the lag coefficient is not significant, and efficiency grows at a rate of 1.5% per year. These results will spill over the adjusted efficiency values.

Table 5. Efficiency evolution through time. Dynamic panel analysis

	Estimate	Standard error	z	P>|z|
lag	0.1234317	0.1234853	1.00	0.318
time	0.0153581	0.0061681	2.49	0.013
z1	0.8387450	1.028.493	0.82	0.415
z2	0.2780564	1.283.786	0.22	0.829
z3	-0.1965680	1.230.432	-0.16	0.873
z4	-0.0136601	0.582840	-0.02	0.981
c	0.2462017	0.913219	0.27	0.787

5. Summary and Conclusions

Under the assumption of a non parametric production model containing random and inefficiency errors we compute DEA based efficiency measurements. The application studied is aimed to Embrapa's research system. Embrapa, the Brazilian Agricultural Research Corporation, is a state company responsible for most of applied agricultural research in Brazil.

Inefficiency error components follow a truncated normal distribution, which is used to compute the 95% upper limits for the individual efficient measurements via bootstrap. Only non efficient units are used for parameter estimation via maximum likelihood. The gamma distribution fails to provide a good fit. The intervals provided include actual adjusted efficiency measurements as possible values. The analysis carried out indicates that size and type of research centers are not significant factors, and that raw efficient measurements grow at a rate of 1.5% a year.

6. Acknowledgment

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

References

Arellano, M. e Bond, S. (1991), Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations, *The Review of Economic Studies*, 58, 277-297.

Arellano, M. e Bover, O. (1995). Another look at the instrumental variable estimation of the error-components models. *Journal of Econometrics*, 68, 29-51.

Banker, R. D. (1993), Maximum likelihood, consistency and DEA: a statistical foundation, *Management Science*, 39(10), 1265-1273.

Banker, R. D., Charnes, A. e Cooper, W. W. (1984), Some models for estimating technical scale inefficiencies in Data Envelopment Analysis, *Management Science*, 30 (9), 1078-1092.

Banker, R. D. e Natarajan, R. Statistical tests based on DEA efficiency scores. Cooper, W W, Seiford, L M and Zhu, J (Eds.), *Handbook on Data Envelopment Analysis*, Kluwer International Series, Boston, 299-321, 2004.

Banker, R. D. e Natarajan, R. (2008), Evaluating contextual variables affecting productivity using data envelopment analysis, *Operations Research*, 56, 48-58.

Blundell, R. e Bond, S. (1998), Initial conditions and moment restrictions in dynamic panel-data models, *Journal of Econometrics*, 87, 115-143.

Charnes, A., Cooper, W. W. e Rhodes, E. (1978), Measuring the efficiency of decision-making units, *European Journal of Operational Research*, 2, 429-444.

Coelli, T. J., Prasada Rao, D. S., O'Donnell, C. J. e Battese, G. E. *An Introduction to Efficiency and Productivity Analysis*, 2nd Edition, Springer, New York, 2005.
Conover, W. J. *Practical Nonparametric Statistics*, Wiley, New York, 1998.
Daraio, C. e Simar, L. *Advanced Robust and Nonparametric Methods in Efficiency Analysis*, Springer, New York, 2007.
Deprins, D., Simar, L. e Tulkens, H. Measuring labor inefficiency in post offices, Marchand, M, Pestieau, P and Tulkens, H (Eds.), *The Performance of Public Enterprises: Concepts and Measurements*, North-Holland, Amsterdam, 243-267, 1984.
Efron, B. e Tibshirani, R. (1986), Bootstrap methods for standard errors, confidence intervals and other measures of statistical accuracy, *Statistical Science*, 1, 54-77.
Efron, B. e Tibshirani, R. An *Introduction to the Bootstrap*, Chapman & Hall, New York, 1993.
Emrouznejad, A. e Gholam, R. A. (2009), DEA models for ratio data: convexity consideration, *Applied Mathematical Modelling*, 33 (1), 486-498.
Gomes, E. G. e Souza, G. S. (2008), Percepções psicométricas sobre a importância de atividades de pesquisa. *Anais do XL Simpósio Brasileiro de Pesquisa Operacional*.
Hollingsworth, B. e Smith, P. (2003), Use of ratios in data envelopment analysis, *Applied Economics Letters*, 10 (11), 733-735.
Kumbhakar, S. e Lovell, C. A. K. *Stochastic Frontier Analysis*, Cambrigde University Press, New York, 2000.
Löthgren, M. e Tambour, M. (1999), Bootstrapping the data envelopment analysis Malmquist productivity index, *Applied Economics*, 31(4), 417-423.
Saaty, T. L. *The Fundamentals of Decision Making and Priority Theory with the Analytic Hierarchy Process*, RWS Publication, Pittsburgh, 1994.
Simar, L. e Wilson, P. W. (2004), Performance of the bootstrap for DEA estimators and iterating the principle. In: Cooper, W W, Seiford, L M and Zhu, J (eds.). *Handbook on Data Envelopment Analysis*. Kluwer International Series: Boston, pp. 265-298.
Simar, L. e Wilson, P. W. (2007), Estimation and inference in two-stage, semi-parametric models of production processes, *Journal of Econometrics*, 136 (1), 31-64.
Souza, G. S. e Staub, R. B. (2007), Two-stage inference using data envelopment analysis efficiency measurements in univariate production models, *International Transactions in Operational Research*, 14, 245-258.
Souza, G. S., Gomes, E. G. e Staub, R. B. (2009a). Influence of contextual variables: an application to agricultural research evaluation in Brazil. *Proceedings International Data Envelopment Analysis Symposium*.
Souza, G. S., Moreira, T. B. e Gomes, E. G. (2009b). An efficiency approach for analyzing the major agricultural economies. *Proceedings International Data Envelopment Analysis Symposium*.
Torgerson, W. S. *Theory and Methods of Scaling*, Wiley, New York, 1958.