TECHNICAL EFFICIENCY IN BRAZILIAN AGRICULTURE: A STOCHASTIC FRONTIER APPROACH

rosaura.gazzola@embrapa.br

Apresentação Oral-Evolução e estrutura da agropecuária no Brasil

GERALDO DA SILVA E SOUZA[1]; ELIANE GONÇALVES GOMES[2]; ROSAURA GAZZOLA[3]; RODRIGO DA SILVA SOUZA[4]; MICHELLE PINTO OLIVEIRA[5].
1,2,3.EMBRAPA SEDE, BRASILIA - DF - BRASIL; 4.EMBRAPA ARROZ E FEIJAO, GOIANIA - GO - BRASIL; 5.UNIVERSIDADE CATÓLICA DE BRASILIA, BRASILIA - DF - BRASIL.

EFICIÊNCIA TÉCNICA DA AGRICULTURA BRASILEIRA: UMA ABORDAGEM DE FRONTEIRA ESTOCÁSTICA

TECHNICAL EFFICIENCY IN BRAZILIAN AGRICULTURE: A STOCHASTIC FRONTIER APPROACH

Grupo de Pesquisa: Evolução e estrutura da agropecuária no Brasil

Resumo
Neste estudo avaliou-se a eficiência técnica dos 27 estados brasileiros nos anos 1995/96 e 2006. Os dados de terra e trabalho utilizados foram os censos agropecuários dos respectivos anos e a informação de custeio e capital foram obtidos no Banco Central. Para a análise dos dados foi usado um modelo de fronteira estocástica, com ajuste aos dados de 99% de correlação entre os valores observados e previstos. Os resultados mostram o Distrito Federal como o estado com maior eficiência técnica no ano de 2006 (0,95) e o segundo melhor em 1995/96 (0,89). O resultado mais baixo foi para o Piauí em 2006 (0,26) e para o Tocantins em 1995/96 (0,22). A elasticidade calculada mostra que ao aumentar em 1% a renda *per capita*, a eficiência técnica aumentaria em 0,77% para a região Nordeste, em 0,76% para a região Norte, 0,59% para o Centro-Oeste, 0,56% para a região Sul e 0,49% para a região Sudeste.
Palavras-chave: Eficiência técnica, Fronteira estocástica de produção, Elasticidade, Agricultura; Estatísticas rurais e agropecuárias.

Abstract
In this study we assessed the technical efficiency of the agricultural sector in the 27 Brazilian states in the years 1995/96 and 2006. The data on land and labor were obtained from the agricultural census of the two considered years. Data on credit for investment and running costs were obtained at the Brazilian Central Bank. In the analysis we used a stochastic frontier model. The model adjusted the data with 99% of correlation between predicted and observed values. The results show Distrito Federal with the highest technical efficiency in agriculture in 2006 (0.95) and the second highest in 1995/96 (0.89). The

lowest technical efficiency was found in Piauí in 2006 (0.26) and in Tocantins in 1995/96 (0.22). The estimated elasticities show that increases of 1% in *per capita* income would increase the technical efficiency by 0.77% in the North, by 0.76% in the Northeast, by 0.59% in the Central-West, by 0.56% in the South and by 0.49% in the Southeast region.
Key Words: Technical efficiency, Stochastic frontier production, Elasticity, Agriculture, Rural Statistics.

1. INTRODUCTION

Brazil is one of the most important countries in relation to agribusiness. Agribusiness represents about 25% of Brazilian GDP, 36% of exports in 2008 and 37% of jobs in 2008.

The objective of this study was to determine the technical efficiency of agriculture and livestock production of the 27 Brazilian states. Since there are regional variations regarding the way the agribusiness is organized in Brazil, it seems to be plausible that the technical efficiency shall differ from state to state. The topic is delicate, since there are considerable differences among states. The Human Development Index (HDI) in 2005 had considerable differences. Distrito Federal (0.874), Santa Catarina (0.840), São Paulo (0.833) and Rio de Janeiro (0.832) are states with higher HDI, while states like Alagoas (0.677) and Maranhão (0.683) have lower HDI (IDH, 2009).

The states of the South and Southeast historically, and more recently, the Central-West use more technology, such as improved varieties of plants, fertilizers, irrigation (Central-West), mechanization and chemicals. Brazilian agriculture differs regionally, due, primarily, to the differences in geographical area, such as climate and natural resources, and thus production characteristics. For example, in South region soybeans, maize, poultry and pork have particular significance, but in Northern region, rubber (hevea), nuts, wood extraction is important. These regional differences cause different technical efficiency among regions. Thus, because of the peculiarities and differences of the states among the regions, further analyses were necessary.

For our analysis, production data were extracted from the agricultural census of 1995/96 and 2006. Together with production data, information on official credit used by farmers for investment and running costs in both mentioned years. We used a stochastic frontier model to estimate the technical efficiency of the agricultural sector in the 27 Brazilian states, in the years 1995/96 and 2006. The literature on technical efficiency measures has some examples of studies using stochastic frontier models to assess the efficiency of agricultural activities, considering regional aggregation levels. In this frame, studies like Chen and Song (2008), Onishi et al. (2008), Kaneko et al. (2004), Bhattacharayya and Parker (1999), Battese and Broca (1997), Hofler and Payne (1995) can me mentioned.

This paper is organized in five sections. Following the introduction, the stochastic frontier analysis is described. Next, data and their sources are described. The last two sections cover the results of this study, and finally, the conclusions.

2. STOCHASTIC FRONTIER ANALYSIS

Basically, two approaches are available in the literature about efficiency analysis: the stochastic efficiency frontier analysis and the deterministic frontier analysis. In the context of deterministic frontiers, Data Envelopment Analysis (DEA) is by far the most used technique.

With a single output, for the stochastic frontier, typically, one specifies a parametric log cost function $C(\ln p, \ln y, \theta)$ dependent on log factor input prices $\ln p$ and log output level $\ln y$, and postulates the model (1), for cost data C_{it} for a panel of N producing units and T time periods.

$$\ln C_{it} = C(\ln p_{it}, \ln y_{it}, \theta) + v_{it} + u_{it} \quad i = 1,\ldots, N \quad t = 1,\ldots,T \tag{1}$$

For a production function one specifies (2), for log inputs $\ln x$.

$$\ln y_{it} = f(\ln x_{it}, \theta) + v_{it} - u_{it} \tag{2}$$

In these formulations, θ is an unknown parameter, $C(.)$ and $f(.)$ have known functional forms, and the stochastic components v_{it} and u_{it} represent random errors and inefficiency errors, respectively.

Typical parametric log cost families are provided by the Translog form (Coelli et al., 2005), the CES (Gallant, 1982), and the Fourier Flexible Form (Gallant, 1982). The latter endows the analysis with nonparametric properties. The random errors v_{it} are assumed to be uncorrelated across time and panel, and normally distributed with mean zero and variance $\sigma_v^{\,2} > 0$. A flexible family of distributions to model the u_{it} (Kumbhakar and Lovell, 2000; Coelli et al., 2005) is provided by truncation of the normal.

In this context one may postulate $u_{it} = z_{it}\delta + w_{it}$, where z_{it} is a vector of specific inefficiency variables (covariates), δ is a vector of unknown coefficients of the firm specific inefficiency variables, and w_{it} is the truncation at $-z_{it}\delta$ of the normal with mean zero and variance σ_u^2. Here we use the production function approach. Here we will follow the production approach using the Cobb-Douglas representation (3), which leads to (4), where $\alpha = \ln\theta_0$.

$$y_{it} = \theta_0 x_{1it}^{\theta_1} x_{2it}^{\theta_2} x_{3it}^{\theta_3} x_{4it}^{\theta_4} \exp(v_{it}) \exp(-u_{it}) \tag{3}$$

$$\ln y_{it} = \alpha + \theta_1 \ln x_{1it} + \theta_2 \ln x_{2it} + \theta_3 \ln x_{3it} + \theta_4 \ln x_{4it} + v_{it} - u_{it} \tag{4}$$

As Coelli et al. (2005) put it, much of stochastic efficiency analysis is directed towards the prediction of inefficiency (efficiency) effects. The most common output-oriented measure of technical efficiency for firm *o* is estimated in the stochastic frontier case by (5), where $\varepsilon_{it} = q_{it} - w_{it}\theta, \mu_{*it} = (-\varepsilon_{it}\,\sigma_u^2 + \mu_{it}\sigma_v^2) \big/ \sigma_S^2, \mu_{it} = z_{it}\delta, \sigma_S^2 = \sigma_u^2 + \sigma_v^2, \sigma_* =$ $= \sigma_u\sigma_v / \sigma_S$.

$$E\left[\exp(-u_{it}) \mid \varepsilon_{it}\right] = \left\{ [1 - \Phi(\sigma_* - \mu_{*it} / \sigma_*)] \big/ \Phi(\mu_{*it} / \sigma_*) \right\} \exp(-\mu_{*it} + 0.5\sigma_*^2) \tag{5}$$

Assuming a normal-truncated normal specification the parameters are obtained maximizing the log likelihood function (Battese and Coelli, 1995) as in (6), where $d_{it} = \mu_{it}/\sigma_u$, $d_{*it} = \mu_{*it}/\sigma_*$, $\sigma_* = \sigma_u\sigma_v/\sigma_S$.

$$
\begin{aligned}
&-\frac{1}{2}\sum_{i=1}^{N}\sum_{t=1}^{T}\left\{\ln(2\pi)+\ln(\sigma_S^2)\right\}\sum_{i=1}^{N}\sum_{t=1}^{T}\left\{\ln(2\pi)+\ln(\sigma_S^2)\right\}\\
&-\frac{1}{2}\sum_{i=1}^{N}\sum_{t=1}^{T}\left\{[(q_{it}-r_{it}\theta)+z_{it}\delta]/\sigma_S^2\right\} \qquad (6)\\
&-\sum_{i=1}^{N}\sum_{t=1}^{T}\left\{\ln\Phi(d_{it})-\ln\Phi(d_{it}^*)\right\}
\end{aligned}
$$

A convenient model re-parameterization, making $\gamma = \sigma_u^2/\sigma_S^2$, leads to the log likelihood as a function $L(\beta,\delta,\gamma,\sigma_S^2)$, where $\sigma_u^2 = \gamma\sigma_S^2$, $\sigma_*^2 = \gamma(1-\gamma)\sigma_S^2$. A classical production model is implied by $\gamma = 0$.

The elasticity for firm *i* in period *t* relative to a contextual variable measured in logs with parameter estimate *b* is computed using the formula (7).

$$
b\left(1-\frac{d_{it}\phi(d_{it})\Phi(d_{it})+[\phi(d_{it})]^2}{[\Phi(d_{it})]^2}\right) \qquad (7)
$$

3. DATA

In this model we used the value of agricultural production as dependent variable and land, labor, capital and running costs as independents variables.

The data on value of agricultural and livestock production of all 27 Brazilian states in the years 1995/96 and 2006 was used. The two years correspond to the two last available agricultural census data in Brazil. The output variable used was the total value of Brazilian agricultural and livestock production in the years 1995/96 and 2006 (total value of production, in R$). The inputs of the model were: total land area used (planted area, in hectares), labor force (employment in agriculture and livestock, number of persons), investment and running costs (in monetary value, R$).

The data on area, labor force and value of production were obtained from the agricultural census (IPEA, 2008). Credit data on investments or capital and running costs or other inputs were extracted from the “Anuários Estatísticos do Crédito Rural” (BACEN, 1995, 2006), representing all official credit taken by farmers in all 27 states in the two years of consideration.

The credit for running costs includes annual expenditures on annual and permanent crops, as well as livestock. Those expenditures are used for maintenance and field operations and for livestock related activities. These expenditures include costs of seeds, fertilizers, pesticides, and field operations. In livestock production, the running expenditures include maintenance of pastures, vaccines, salt and medicaments. Depending on the state, the running expenditures for agriculture and livestock are composed by different crops and livestock types.

Agricultural credit for running costs of livestock activities include the costs of keeping wild animals, bee keeping, poultry, cattle, buffalo, goats, rabbits, horses, sheep, fishery, pig, pastures, vaccines, mineral salts and medicaments.

Agricultural investments include several items and monetary units are related in R$. The establishment of perennial crops, the renovation of plantations, the improvement of agricultural enterprises, rural electricity, storage building, machinery and implements are among the items covered by credit for investment. In livestock production, the official credit for investment covers the acquisition of production animals (domestically or imported).

4. RESULTS

For the analysis all variables were measured in log form. The actual model used postulates a linear relationship between the log of the agricultural production *y* and the log of the inputs *l*, *k*, *t*, and *c*, denoting labor, capital, land and other inputs, respectively.

Table 1 shows the statistical results of maximum likelihood estimation of the stochastic frontier model using Stata 10.1 software.

Table 1. Stochastic frontier estimation.

	Coefficient	Standard error	z	P>|z|	[95% Confidence interval]	
Production (*y*)						
Labor (*l*)	0.3238	0.0442	7.33	0.000	0.2373	0.4103
Capital (*k*)	0.1413	0.0529	2.67	0.008	0.0376	0.2449
Land (*t*)	0.2672	0.0286	9.34	0.000	0.2111	0.3232
Other Inputs (*c*)*	0.2633	0.0470	5.60	0.000	0.1712	0.3554
Constant	-0.9182	0.4623	-1.99	0.047	-1.8243	-0.0120
Technical effect						
l (GDP pc)	-0.7699	0.1114	-6.91	0.000	-0.9881	-0.5516
constant	1.8278	0.2726	6.71	0.000	1.2935	2.3620
sigma_S2	0.0604				0.0396	0.0921
gamma	0.4627				0.0319	0.9574

*Other inputs are running costs for expenditures on annual and perennial cultivates and livestock.

The likelihood ratio test statistic for the joint hypothesis implying the presence of HDI (Human Development Index), time, and regional effects has a value of 7.15 with a *p*-value of about 31%. For this reason we dropped log HDI and all the other categorical variables, and use the more parsimonious model presented in Table 1.

Table 1 also shows that a 1% increase in different inputs would have different impacts on production: in capital the gains in production would be 0.14% *ceteris paribus*, in labor 0.32%, in land 0.27%, and in other inputs 0.26%.

Figure 1 represents the results of the stochastic frontier model. We can see clearly differences in efficiency, being the Southeast and South the two most efficient regions, followed by the Central-West region. In all regions the technical efficiency changed between 1995/96 and 2006. However, the biggest gain in efficiency can be observed in the Central-West region.

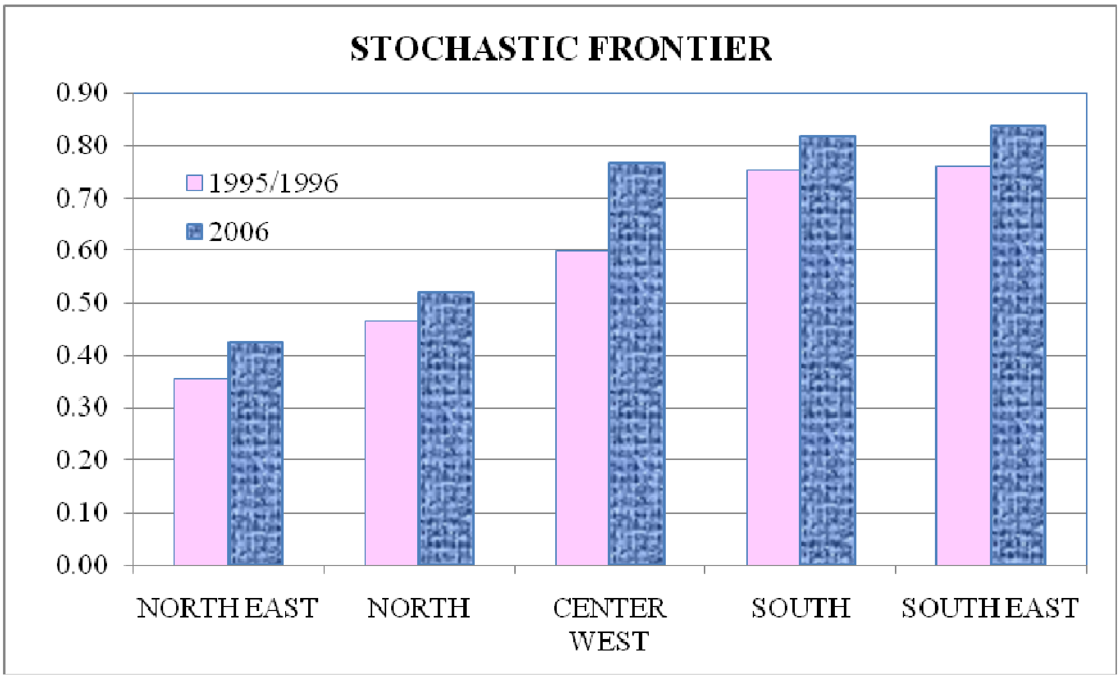

Figure 1. Results from the stochastic frontier model. Years: 1995/96 and 2006.

Table 2 shows statistics (stochastic efficiency estimates) computed as a function of model parameter estimates. The Pearson correlation between observed and predicted values is about 99%, indicating a good fit for the frontier model. The 95% confidence interval for the parameter γ suggests a technical components model. The mean stochastic technical efficiencies for each state are shown in this table. Considering an average of both analyzed years, the most efficient state is São Paulo (0.93) and the least efficient is Piauí (0.25). The Southeast dominates, followed by South, Central-West, North and Northeast. The dominance of the Southeast and South over the other regions is strong. These results are somewhat expected and serves the purpose to further validation of our model.

Table 2. Rank of stochastic technical efficiencies.

States	Region	1995/96	States	Region	2006
São Paulo	Southeast	0.9040	Distrito Federal	Central-West	0.9529
Distrito Federal	Central-West	0.8898	São Paulo	Southeast	0.9467
Amapá	North	0.8142	Rio de Janeiro	Southeast	0.8773
Santa Catarina	South	0.8125	Espírito Santo	Southeast	0.8571
Rio de Janeiro	Southeast	0.8018	Santa Catarina	South	0.8426
Rio Grande do Sul	South	0.7538	Paraná	South	0.8113
Espírito Santo	Southeast	0.7387	Rio Grande do Sul	South	0.7971
Paraná	South	0.6857	Mato Grosso	Central-West	0.7937
Amazonas	North	0.6814	Minas Gerais	Southeast	0.6720
Minas Gerais	Southeast	0.5947	Goiás	Central-West	0.6632
Mato Grosso do Sul	Central-West	0.5900	Mato Grosso do Sul	Central-West	0.6599
Pernambuco	Northeast	0.4811	Amazonas	North	0.6004
Goiás	Central-West	0.4716	Rondônia	North	0.5953
Acre	North	0.4538	Amapá	North	0.5454
Mato Grosso	Central-West	0.4445	Pernambuco	Northeast	0.5075
Alagoas	Northeast	0.4261	Roraima	North	0.5021
Pará	North	0.3821	Acre	North	0.4929
Rondônia	North	0.3804	Rio Grande do Norte	Northeast	0.4877
Rio Grande do Norte	Northeast	0.3748	Bahia	Northeast	0.4869
Ceará	Northeast	0.3743	Pará	North	0.4732
Paraíba	Northeast	0.3634	Sergipe	Northeast	0.4661
Bahia	Northeast	0.3412	Alagoas	Northeast	0.4408
Sergipe	Northeast	0.3351	Ceará	Northeast	0.4373
Roraima	North	0.3206	Paraíba	Northeast	0.4282
Piauí	Northeast	0.2443	Tocantins	North	0.4247
Maranhão	Northeast	0.2441	Maranhão	Northeast	0.3098
Tocantins	North	0.2215	Piauí	Northeast	0.2619

The chi-square test for constant returns to scale ($l+k+t+c=1$) has a *p*-value of 90%, non significant. Although the confidence intervals for all input variables do intercept, the pair wise Wald test of equality indicates that the labor (*l*) elasticity is stronger than the capital (*k*) elasticity (*p*-value 0.02), and the capital elasticity is weaker then the land elasticity (*t*) (*p*-value 0.03). The difference between labor and land elasticities is marginal (*p*-value 0.10). No other pair wise comparison was significant.

As shown in Tables 3 and 4, the average income *per capita* elasticity over all states and years is 0.67 with a standard error of 0.17. The minimum *per capita* income elasticity is 0.05, and the maximum 0.77. States highly efficient have smaller income elasticities. This means, that a 1% increase in *per capita* income will increase the agricultural production in wealthier states like Distrito Federal by only 0.14%, but can achieve up to 0.77% in those states with lower technical efficiency.

Table 3. Average elasticities by state, in descending order.

States	Average elasticity
Distrito Federal	0.1383
São Paulo	0.2701
Rio de Janeiro	0.4020
Rio Grande do Sul	0.5043
Santa Catarina	0.5253
Espírito Santo	0.5589
Paraná	0.6479
Amazonas	0.6689
Mato Grosso	0.7069
Minas Gerais	0.7359
Mato Grosso do Sul	0.7470
Goiás	0.7601
Amapá	0.7618
Roraima	0.7668
Rondônia	0.7689
Sergipe	0.7697
Tocantins	0.7698
Acre	0.7699
Alagoas	0.7699
Bahia	0.7699
Ceará	0.7699
Maranhão	0.7699
Pará	0.7699
Paraíba	0.7699
Pernambuco	0.7699
Piauí	0.7699
Rio Grande do Norte	0.7699

Table 4. Average elasticities per region, in descending order.

Region	Frequency	Average elasticity
Southeast	8	0.4917
South	6	0.5592
Central-West	8	0.5881
North	14	0.7537
Northeast	18	0.7699

6. CONCLUSIONS

We fitted a stochastic frontier model to state agricultural production data in Brazil. The fit was very good as measured by a correlation of about 99% between observed and predicted values. The technology seems to show constant returns to scale.

The model also includes a statistically significant contextual inefficiency effect defined by *per capita* income. The average *per capita* income elasticity is 0.67 with a standard error of 0.17. The income variable is used as a proxy for infra structure and technology assessment. We find stronger elasticities results for labor, other inputs (running costs) and land.

Southeast and South states are significantly more efficient than other states on average. São Paulo, Distrito Federal, Rio de Janeiro, Santa Catarina, Espírito Santo, Rio Grande do Sul and Paraná are the most efficient states with product oriented technical efficiencies over 70% well above the other states.

These empirical results suggest one important finding. There are significant possibilities to increase efficiency levels in Brazil agriculture production, especially in the Northeast and North Region.

Finally the results indicate the diversity of the scores of efficiency among regions. This suggests that the considerable variability of regions in climate, natural resources, irrigation, etc. (infrastructure, agro industries), can have different impacts on efficiency in Brazil agricultural production in different regions.

7. REFERENCES

BANCO CENTRAL DO BRASIL (BACEN). **Anuário Estatístico do Crédito Rural – 1995**. Brasília: Departamento de Cadastro e Informação. 403 p.

BANCO CENTRAL DO BRASIL (BACEN). **Anuário Estatístico do Crédito Rural - 2006**. Disponível em: <http://www.bcb.gov.br/default.asp?id=relrural&ano=2006>. Acesso em: Out. 2008.

BATTESE, G.E.; BROCA, S.S. Functional forms of stochastic frontier production functions and models for technical inefficiency effects: a comparative study for wheat farmers in Pakistan. **Journal of Productivity Analysis**. v. 8. p. 395-414. 1997.

BATTESE, G.E.; COELLI, T.J. A model for technical inefficiency effects in a stochastic frontier production function for panel data. **Empirical Economics**. v. 20. p. 325-332. 1995.

BHATTACHARAYYA, A.; PARKER, E. Labor productivity and migration in Chinese agricultures: a stochastic frontier approach. **China Economic Review**. v. 10. p. 59-74. 1999.

CHEN, Z.; SONG, S. Efficiency and technology gap in China's agriculture: A regional meta-frontier analysis. **China Economic Review**. v. 19. n. 2. p. 287-296.

COELLI, T.J.; PRASADA RAO, D.S.; O'DONNELL, C.J.; BATTESE, G.E. **An Introduction to Efficiency and Productivity Analysis**. 2nd Edition. New York: Springer. 2005. 350p.

GALLANT, A.R. Unbiased determination of production technologies. **Journal of Econometrics**. v. 20. p. 285-323. 1982.

HOFLER, R.A.; PAYNE, J.E. Regional efficiency differences and development policy of agriculture in the Yugoslav republics: estimates from panel data. **Review of Regional Studies**. v. 25. n. 3. p. 287-300. 1995.

ÍNDICE DE DESENVOLVIMENTO HUMANO (IDH). Disponível em <http://pt.wikipedia.org/wiki/Lista_de_estados_do_Brasil_por_IDH >. Acesso em: 18 mar. 2009.

INSTITUTO DE PESQUISA ECONÔMICA APLICADA (IPEA). **IpeaData**. 2008. Disponível em: <http://www.ipeadata.gov.br>. Acesso em: Out. 2008.

KANEKO, S.; TANAKA, K.; TOYOTA, T.; MANAGI, S. Water efficiency of agricultural production in China: regional comparison from 1999 to 2002. **International Journal of Agricultural Resources**. Governance and Ecology. v. 3. n. 3-4. p. 231-251. 2004.

KUMBHAKAR, S.; LOVELL, C.A.K. **Stochastic Frontier Analysis**. New York: Cambrigde University Press. 2000. 344p.

ONISHI, A.; MORISUGI, M.; IMURA, H.; SHI, F.; WATANABE, T.; FUKUSHIMA, Y. Study on the efficiency of agricultural water use in the Yellow River Basin. **Journal of Global Environment Engineering**. v. 13. p. 51-67. 2008.