



# Custos operacionais eficientes das distribuidoras de energia elétrica: um estudo comparativo dos modelos DEA e SFA

*The cost efficiency of the Brazilian electricity distribution utilities: a comparative study with DEA and SFA models*

Marcus Vinicius Pereira de Souza<sup>1</sup>  
Reinaldo Castro Souza<sup>2</sup>  
José Francisco Moreira Pessanha<sup>3</sup>

**Resumo:** Este artigo apresenta medidas de eficiência para 40 distribuidoras de energia elétrica que operam no setor elétrico brasileiro. As medidas foram obtidas por modelos de análise envoltória de dados (DEA) e modelos de fronteira estocástica (SFA), duas técnicas que podem mitigar a assimetria de informação e aprimorar a habilidade do agente regulador comparar os desempenhos das distribuidoras, requisitos fundamentais em esquemas de regulação incentivada. As duas abordagens são apresentadas e os resultados obtidos pelos diferentes modelos são comparados.

**Palavras-chave:** Análise envoltória de dados. Fronteira estocástica. Regulação econômica.

**Abstract:** This paper shows the efficiency measurements of 40 Brazilian electricity distribution companies. The efficiency scores are obtained using the data envelopment analysis (DEA) and stochastic frontier analysis (SFA) models, techniques that can reduce the information asymmetry and improve the regulator skills to compare the performance of the electricity companies, which are fundamental aspects of regulatory regimes. The two approaches are described, and the main results obtained from the different models are compared.

**Keywords:** Data envelopment analysis. Stochastic frontier. Economic regulation.

## 1 Introdução

No Setor Elétrico Brasileiro (SEB) as tarifas de fornecimento são revisadas periodicamente a cada quatro ou cinco anos, dependendo do contrato de concessão das empresas distribuidoras. No ano de realização da revisão periódica as tarifas são reposicionadas em um nível compatível com a cobertura dos custos operacionais eficientes e com a remuneração adequada dos investimentos realizados prudentemente pelas distribuidoras, garantindo a manutenção do equilíbrio econômico-financeiro (EEF) da concessão.

No período entre duas revisões tarifárias, as tarifas são reajustadas anualmente pelo IRT (Índice de Reajuste Tarifário) com a finalidade de assegurar que o equilíbrio econômico-financeiro da concessão não sofrerá a corrosão do processo inflacionário. Na Figura 1 são ilustrados os processos de revisão tarifária

e reajuste tarifário. Para simplificar, admite-se que as variações do IGPM sejam nulas.

A tarifa inicial  $T_0$ , fixada no contrato de concessão, permanece constante (em termos reais) até a primeira revisão tarifária (2003). Esta tarifa garante o EEF da concessão no primeiro período tarifário (1998-2003), cobrindo os custos não gerenciáveis pela distribuidora (compras de energia e encargos setoriais) e os custos gerenciáveis pela distribuidora (mão de obra, serviços de terceiros, material, depreciação e remuneração adequada dos ativos investidos e do capital de giro). Neste período inicial, os ganhos de produtividade decorrentes do crescimento do mercado e da redução dos custos operacionais alcançados pela gestão eficiente da concessionária, conforme indicado pelo triângulo cinza na Figura 1,

<sup>1</sup> Departamento de Engenharia Industrial, Pontifícia Universidade Católica – PUC, Rua Marquês de São Vicente, 225, Gávea, CEP 22451-041, Rio de Janeiro – RJ, E-mail: mvinic@engenharia.ufjf.br

<sup>2</sup> Departamento de Engenharia Elétrica, Pontifícia Universidade Católica – PUC, Rua Marquês de São Vicente, 225, Gávea, CEP 22451-041, Rio de Janeiro – RJ, E-mail: reinaldo@ele.puc-rio.br

<sup>3</sup> Instituto de Matemática e Estatística, Universidade do Estado do Rio de Janeiro – UERJ, Rua São Francisco Xavier, 524, Maracanã, CEP 20550-900, Rio de Janeiro – RJ, E-mail: francisc@cepel.br

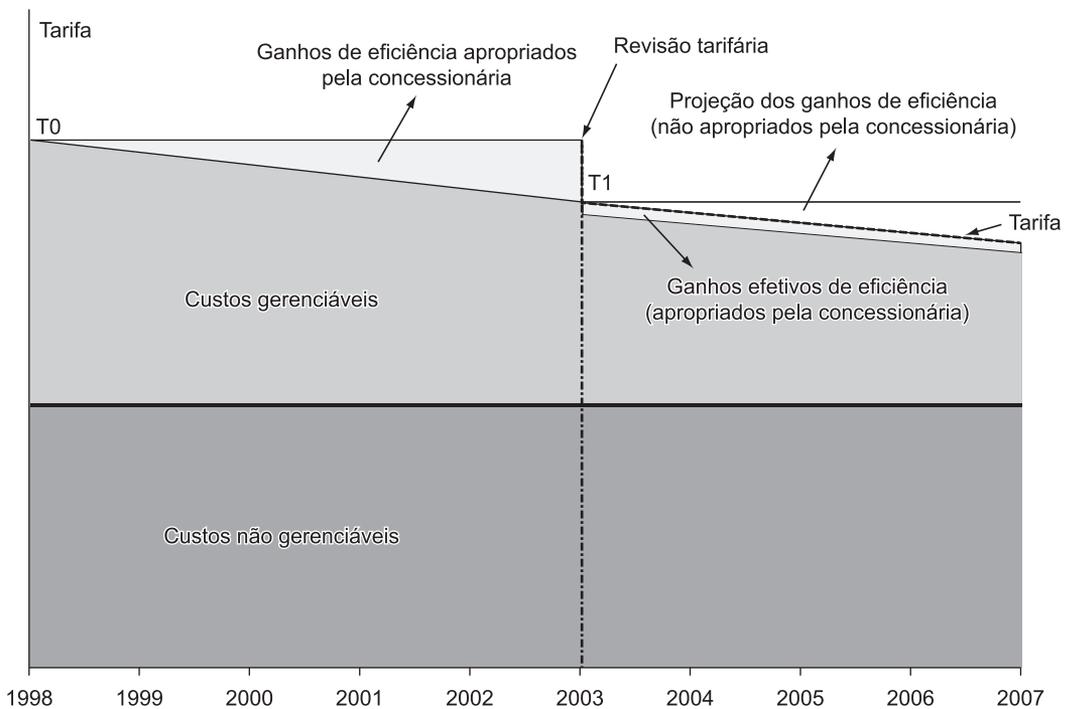


Figura 1. Processo de revisão tarifária (BRASIL, 2003).

são integralmente apropriados pela concessionária, aumentando sua remuneração.

No final do primeiro período tarifário acontece a primeira revisão tarifária periódica (2003), a qual transcorre em duas etapas. Primeiro, o regulador reposiciona a tarifa em um nível (T1) compatível com a receita necessária para atender às condições eficientes de manutenção e operação da atividade de distribuição, bem como uma taxa de retorno considerada adequada ao risco do negócio e que remunere o capital investido. Em seguida, a partir dos ganhos de eficiência projetados ou metas de produtividade para o segundo período tarifário (2004-2007), o regulador define o Fator X, um componente do índice de reajuste tarifário (IRT) a ser aplicado anualmente durante o próximo período tarifário, Equação 1:

$$IRT = \left[ VPA1 + VPB0 \cdot (IGPM - X) \right] / RAO \quad (1)$$

em que: VPA1 é a parcela referente aos custos não gerenciáveis pela distribuidora na data do reajuste, RAO é a receita anual calculada considerando-se as tarifas vigentes (sem ICMS) na data da revisão tarifária e o respectivo mercado de referência, e VPB0 é a parcela relativa aos custos gerenciáveis pela distribuidora na data da revisão ( $VPB0 = RAO - VPA0$ ).

Conforme indicado na Equação 1, os custos não gerenciáveis (VPA) são repassados integralmente às tarifas finais, enquanto a parcela relativa aos custos gerenciáveis (VPB) é atualizada pelo IGPM,

descontado do Fator X. Este fator incide apenas sobre os custos gerenciáveis e constitui o mecanismo que permite compartilhar com os consumidores os ganhos de produtividade decorrentes do crescimento da demanda atendida pela distribuidora.

Na distribuição de energia elétrica, os ganhos de produtividade decorrem principalmente do crescimento do mercado atendida pela concessionária, tanto pelo maior consumo das unidades consumidoras existentes como pela ligação de novas unidades. Em função da presença de economias de escala, a expansão do mercado é atendida com custos incrementais decrescentes, resultando em um ganho de produtividade para a distribuidora que não decorre de uma gestão mais eficiente da empresa e que, portanto, deve ser repassado para os consumidores com a finalidade de promover a modicidade tarifária.

Conforme indicado na Figura 1, no segundo período tarifário (2004-2007), a tarifa segue a trajetória decrescente representada pela linha tracejada, cobrindo o repasse integral dos custos não gerenciáveis e uma parte dos custos gerenciáveis que se reduz anualmente em função do seu reajuste pelo IGPM descontado do Fator X.

O resultado deste mecanismo é que a projeção dos ganhos de produtividade decorrentes do crescimento do mercado no período 2004/2007, conforme indicado pelo triângulo em branco na Figura 1, não são apropriados integralmente pela concessionária, mas compartilhados com os consumidores. No segundo período tarifário, a aplicação efetiva do Fator X

reduz a parcela dos ganhos de eficiência que podem ser apropriados pela concessionária, limitando os ganhos que ultrapassem as metas de produtividade definidas na data da revisão tarifária, conforme indicado pela estreita faixa em cinza na Figura 1. Assim, a concessionária é incentivada a explorar seu potencial de eficiência, e quanto maior for a eficiência da concessionária maior será seu benefício, dado pela diferença entre a meta de produtividade e o custo efetivamente realizado. No entanto, a concessionária fica sujeita ao risco de perdas, caso os custos não sigam uma trajetória mais eficiente que metas de produtividade.

Em resumo, permite-se um ganho extraordinário temporário, a fim de minimizar custos, e posteriormente há o compartilhamento desse ganho com os consumidores.

Em concordância com o exposto, a Resolução da Agência Nacional de Energia Elétrica (ANEEL) Nº 55/2004 define o Fator X como sendo a combinação de três componentes ( $X_e$ ,  $X_a$  e  $X_c$ ), conforme a Equação 2:

$$X = (X_e + X_c) \cdot (IGPM - X_a) + X_a \quad (2)$$

A componente  $X_a$  reflete o efeito da aplicação do IPCA na componente mão de obra do VPB da concessionária. A componente  $X_c$  relaciona-se com a qualidade percebida pelos consumidores atendidos pela concessionária. Por fim, a componente  $X_e$  reflete os ganhos esperados de produtividade da concessionária em virtude do crescimento do mercado atendido pela concessionária. Esta última componente é a mais importante, sendo definida com base no método do fluxo de caixa descontado do tipo *forward looking* de maneira a igualar o valor presente líquido do fluxo de caixa da concessionária no período tarifário, acrescido do valor residual, com o valor dos ativos da concessionária no início do período tarifário, Equação 3:

$$A_0 = \sum_{t=1}^N \left[ \frac{\left( \begin{array}{l} RO_t \cdot (1 - X_e)^{t-1} - \\ - T_t - OM_t - d_t \end{array} \right) \cdot (1 - g) + d_t - I_t}{(1 + r_{WACC})^t} \right] + \frac{A_N}{(1 + r_{WACC})^N} \quad (3)$$

em que:  $N$  é o período em anos entre duas revisões tarifárias,  $A_0$  é o valor dos ativos da concessionária na data da revisão tarifária,  $A_N$  é o valor dos ativos da concessionária no final do período tarifário,  $g$  são

as alíquotas do imposto de renda e da Contribuição Social sobre o Lucro Líquido (CSLL),  $r_{WACC}$  é o custo médio ponderado do capital,  $RO_t$  é a receita operacional da concessionária,  $T_t$  são os valores dos tributos (PIS/PASEP, COFINS e P&D),  $OM_t$  são os custos de operação e manutenção da concessionária,  $I_t$  é o montante de investimentos realizados e  $d_t$  é a depreciação, todos referentes ao ano  $t$ .

Vale esclarecer que as parcelas que compõem o fluxo de caixa na Equação 3 são projetadas segundo os critérios descritos na Resolução ANEEL Nº 55/2004. Por exemplo, a projeção da receita operacional é determinada pelo produto entre o mercado previsto e a tarifa média do reposicionamento tarifário, enquanto, os custos operacionais (custos de operação, manutenção, administração e gestão comercial) são projetados com base nos custos da “Empresa de Referência”, referenciados à data do reposicionamento tarifário.

A empresa de referência é uma distribuidora virtual na qual se simula a prestação do serviço de distribuição de energia elétrica nas mesmas condições que opera uma distribuidora real. Para um dado nível de qualidade exigido pelo regulador, a empresa de referência é orientada para a utilização eficiente dos recursos, estabelecendo assim um padrão de desempenho (*benchmark*) a ser perseguido pela empresa real. Esta estratégia segue as linhas gerais da regulação por comparação (*yardstick competition*), pois emula a competição entre uma distribuidora real e a respectiva empresa de referência.

A metodologia da empresa de referência adotada no primeiro ciclo de revisões tarifárias (BRASIL, 2006) segue uma abordagem tipo *bottom-up* que se inicia com a identificação de todos os processos inerentes à atividade de distribuição de energia elétrica nas áreas comercial e técnica, passa pelo estabelecimento dos custos eficientes de cada um dos processos identificados e se encerra com uma estimativa para o total dos custos operacionais eficientes. Portanto, trata-se de uma metodologia complexa e que abre a possibilidade de envolver o regulador em uma espécie de microgestão da empresa, justamente o contrário do que recomenda a boa prática da ação reguladora. Para evitar a complexidade da atual metodologia da empresa de referência e tornar mais objetivo o processo de estabelecimento dos custos operacionais eficientes, a ANEEL não descarta a possibilidade de utilizar técnicas de *benchmarking*, entre as quais as fronteiras de eficiência, conforme o modelo atual adotado na quantificação dos custos operacionais eficientes das empresas de transmissão (BRASIL, 2007).

De posse dessas informações, a seguir, na seção 2, é feito um breve histórico acerca da regulação do setor elétrico. A seção 3 aborda os modelos de fronteira estocástica para dados tipo *cross-section*, enquanto,

a análise envoltória de dados é descrita na seção 4. O mapa de Kohonen, uma rede neural não supervisionada utilizada na identificação dos grupos de distribuidoras similares, é apresentado na seção 5. As especificações dos modelos DEA e SFA são discutidas na seção 6 e os principais resultados obtidos são reportados na seção 7. Finalmente, na seção 8, são resumidas as principais conclusões do trabalho.

## 2 Regulação econômica do setor elétrico

A pesquisa realizada por Jasmab e Pollit (2000) mostra que os principais métodos de *benchmarking* usados na regulação dos serviços de distribuição de eletricidade são: a) Análise Envoltória de Dados (DEA – *Data Envelopment Analysis*), uma técnica baseada em programação linear (COOPER; SEIFORD; TONE, 2000); b) modelos econométricos, entre os quais destacam-se os modelos de fronteira estocástica (SFA – *Stochastic Frontier Analysis*) (KUMBHAKAR; LOVELL, 2000). Isto posto, vale enfatizar que os primeiros trabalhos desenvolvidos à luz da técnica DEA, para avaliar a eficiência das empresas de eletricidade, são creditados a Färe, Grosskopf e Logan (1983) e Charnes et al. (1989). Adicionalmente, é importante citar os estudos elaborados por Jasmab e Pollit (2003), que comparam 63 distribuidoras em 6 países europeus, a partir de modelos DEA, SFA e COLS (*Corrected Ordinary Least Square*) e Estache, Rossi e Ruzzier (2004) que realizam análise similar para as empresas de eletricidade da América do Sul.

No âmbito do SEB, Resende (2001), Vidal e Távora Junior (2003), Pessanha, Souza e Laurencel (2004) e Sollero e Lins (2004) divulgam resultados estimados por distintos modelos DEA na avaliação da eficiência das distribuidoras de energia elétrica. Índices de eficiência obtidos por modelos SFA são apresentados por Zanini (2004) e Arcoverde, Tannuri-Pianto e Sousa (2005). Souza, Souza e Pessanha (2007) preconizam uma comparação dos modelos DEA e SFA.

Além dessas referências, Souza (2008) discute diversos modelos DEA e SFA (clássico e Bayesiano). Deve ser ressaltado, contudo, que os modelos econométricos com enfoque Bayesiano, registrados em Souza (2008) e Souza et al. (2010), são mais complexos e, conforme se sabe, utilizam simulação estocástica (*Markov chain Monte Carlo* – MCMC).

Por fim, os métodos DEA e SFA partem de pressupostos diferentes e, de acordo com a aplicação em que são empregados, exibem vantagens e desvantagens. Por esta razão, é interessante investigar os índices de eficiências obtidos pelas duas abordagens aplicados em um mesmo conjunto de dados. Nesse sentido, aqui, os dois métodos são aplicados na avaliação dos custos operacionais eficientes de 40 distribuidoras que operam no SEB.

## 3 Modelos de fronteira estocástica

### 3.1 Fronteira de produção

Considere uma amostra tipo *cross-section* com  $I$  DMUs (*decision making units*), neste caso cada DMU é uma concessionária de distribuição que utiliza  $n$  tipos de insumos representados pelo vetor de quantidades de insumos  $x_i \in R^n_+$ ,  $i = 1, I$ , na produção de uma quantidade  $y_i$  de um único tipo de produto. A medida de eficiência técnica para a  $i$ -ésima DMU ( $0 \leq TE_i \leq 1$ ) é definida pela razão entre a sua produção ( $y_i$ ) e a máxima produção possível, especificada pela fronteira estocástica de eficiência [ $f(x_i, \beta)e^{v_i}$ ], em que  $\beta$  é o vetor de parâmetros da função de produção a ser estimada e  $v_i$  é uma variável aleatória irrestrita em sinal, Equação 4:

$$TE_i = \frac{y_i}{f(x_i, \beta)e^{v_i}} \quad (4)$$

Arranjando os termos da igualdade (Equação 4), tem-se a equação da fronteira estocástica, Equação 5:

$$y_i = TE_i \cdot f(x_i, \beta)e^{v_i} \quad (5)$$

A fronteira estocástica é composta de duas partes: a fronteira determinística  $f(x_i, \beta)$ , comum a todos os produtores, e o termo  $e^{v_i}$ , cuja finalidade é capturar o efeito de choques aleatórios que afetam especificamente a  $i$ -ésima DMU. Fazendo  $TE_i = e^{-u_i}$ , em que  $u_i \geq 0$  é uma variável aleatória, e substituindo-se na Equação 5, tem-se que, Equação 6:

$$y_i = f(x_i, \beta)e^{v_i - u_i} \quad (6)$$

Admitindo-se que  $f(x_i, \beta)$  seja linear nos logaritmos, tem-se o seguinte modelo, após a aplicação de uma transformação logarítmica na Equação 6, em que  $u_i = -\ln TE_i \geq 0$ , Equação 7:

$$\ln y_i = \beta_0 + \beta_1 \ln x_i + \dots + \beta_n \ln x_n + v_i - u_i \quad (7)$$

Na Equação 7, o desvio entre o nível de produção observado e a parte determinística da fronteira de produção é dado pela combinação de duas componentes:  $u_i$ , um erro que assume apenas valores não negativos e que captura o efeito da ineficiência técnica, e  $v_i$ , um erro simétrico que captura qualquer choque aleatório não gerenciável. A hipótese de simetria da distribuição de  $v_i$  é suportada pelo fato de que condições ambientais desfavoráveis e favoráveis são igualmente prováveis. Fazendo-se  $\varepsilon_i = v_i - u_i$ , tem-se a seguinte equação de regressão, Equação 8:

$$\ln y_i = \beta_0 + \beta_1 \ln x_i + \dots + \beta_n \ln x_n + \varepsilon_i \quad (8)$$

Admitindo-se que  $v_i$  seja *i.i.d.* com distribuição simétrica, usualmente uma distribuição normal, e que seja independente de  $u_i$ , o qual sempre assume um valor não negativo, a distribuição do erro composto  $\varepsilon_i$  é assimétrica. Portanto, métodos de estimação que admitem erros simetricamente distribuídos e com média zero não são apropriados para estimar a função fronteira estocástica na Equação 8.

Na formulação básica dos modelos de fronteira estocástica (modelo Normal/Half-Normal), são admitidas as seguintes hipóteses acerca dos termos aleatórios  $u_i$  e  $v_i$ :

$$(H1) \quad v_i \sim iid N(0, \sigma_v^2) \Rightarrow f(v) = \frac{1}{\sqrt{2\pi}\sigma_v} e^{-v^2/2\sigma_v^2}$$

$$u_i \sim iid N^+(0, \sigma_u^2)$$

$$(H2) \quad (Half - Normal \text{ não negativa}) \Rightarrow f(u) = \frac{2}{\sqrt{2\pi}\sigma_u} e^{-u^2/2\sigma_u^2}$$

(H3)  $u_i$  é independente de  $v_i$

(H4)  $v_i$  e  $u_i$  são independentes de  $x_i$

As hipóteses acima permitem estimar os parâmetros da Equação 8 por máxima verossimilhança. A partir da hipótese (H3), tem-se que a distribuição conjunta de  $u_i$  e  $v_i$  é o produto das densidades marginais em (H1) e (H2), Equação 9:

$$f(u, v) = \frac{2}{2\pi\sigma_u\sigma_v} e^{-\left(\frac{u^2}{2\sigma_u^2} + \frac{v^2}{2\sigma_v^2}\right)} \quad (9)$$

Lembrando que  $\varepsilon_i = v_i - u_i$ , obtém-se a distribuição conjunta de  $u_i$  e  $\varepsilon_i$ , Equação 10:

$$f(u, \varepsilon) = \frac{2}{2\pi\sigma_u\sigma_v} e^{-\left(\frac{u^2}{2\sigma_u^2} + \frac{(\varepsilon+u)^2}{2\sigma_v^2}\right)} \quad (10)$$

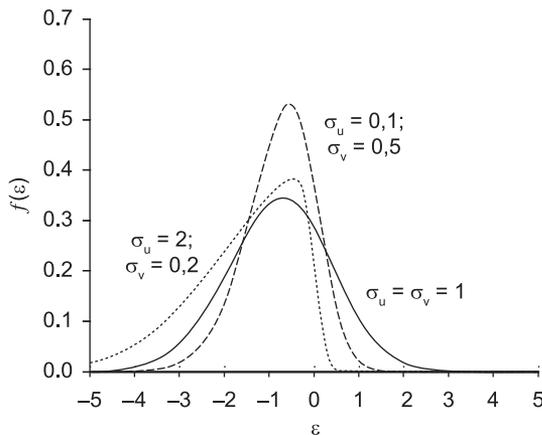


Figura 2. Função densidade do erro composto no modelo Normal/Half-Normal.

A distribuição marginal de  $\varepsilon_i$  é obtida pela integração da densidade da Equação 10 em relação à  $u_i$ , Equação 11:

$$f(\varepsilon) = \int_0^\infty f(u, \varepsilon) du = \frac{2}{\sqrt{2\pi}\sigma} \left[ 1 - \Phi\left(\frac{\varepsilon\lambda}{\sigma}\right) \right] \cdot e^{-\frac{\varepsilon^2}{2\sigma^2}} = \frac{2}{\sigma} \phi\left(\frac{\varepsilon}{\sigma}\right) \Phi\left(-\frac{\varepsilon\lambda}{\sigma}\right) \quad (11)$$

em que:  $\lambda = \sigma_u/\sigma_v$ ,  $\sigma = \sqrt{\sigma_u^2 + \sigma_v^2}$ ,  $\phi$  é a função densidade de uma  $N(0,1)$  e  $\Phi$  é a respectiva função distribuição de probabilidade acumulada.

Conforme ilustrado na Figura 2 para diferentes valores de  $\sigma_u^2$  e  $\sigma_v^2$ , a distribuição de  $\varepsilon_i$  é negativamente assimétrica com média e variância dadas pelas Equações 12 e 13:

$$E(\varepsilon) = -E(u) = -\sigma_u\sqrt{2/\pi} \quad (12)$$

$$V(\varepsilon) = \sigma_u^2(\pi - 2)/\pi + \sigma_v^2 \quad (13)$$

Lembrando que  $\varepsilon_i = \ln y_i - (\beta_0 + \beta_1 \ln x_{i1} + \dots + \beta_n \ln x_{in})$  e usando-se a função densidade na Equação 11, tem-se o logaritmo da função de verossimilhança para uma amostra com  $I$  produtores, Equação 14:

$$\ln L = \text{constante} - I \ln \sigma + \sum_i \ln \Phi\left(\frac{\varepsilon_i \lambda}{\sigma}\right) - \frac{1}{2\sigma^2} \sum_i \varepsilon_i^2 \quad (14)$$

As estimativas de máxima verossimilhança, são os valores de  $\sigma_u$ ,  $\sigma_v$  e  $\beta_j, j = 0, n$ , que maximizam a Equação 14. Estas estimativas são assintoticamente consistentes.

A parametrização do logaritmo da verossimilhança em termos de  $\sigma^2 = \sigma_u^2 + \sigma_v^2$  e  $\gamma = \sigma_u^2 / (\sigma_u^2 + \sigma_v^2) \in [0,1]$  ajuda na maximização da função de verossimilhança e facilita a interpretação do modelo. Quando  $\gamma = 0$ , não há ineficiência, pois  $\sigma_v^2$  domina a variância do erro composto e os desvios em relação à fronteira devem-se unicamente a ruídos aleatórios, por outro lado, quando  $\gamma = 1$ ,  $\sigma_u^2$  domina a variância do erro composto e o desvio em relação à fronteira deve-se inteiramente à ineficiência. Portanto, pode-se utilizar o teste da razão da verossimilhança (*LR test*) para testar a hipótese nula  $H0: \gamma = 0$ . Porém, como  $\gamma$  pode cair nos limites do intervalo  $[0,1]$ , a distribuição da razão de verossimilhança é aproximadamente uma  $\chi^2_1$ .

Para estimar a eficiência técnica de cada produtor deve-se ter uma estimativa de  $u_i$ , a componente do erro que captura o efeito da ineficiência técnica. Esta estimativa pode ser obtida a partir dos resíduos  $\hat{\varepsilon}_i$ . Com as funções densidade nas Equações 10 e 11, pode-se estimar  $u_i$  como sendo a média/moda da seguinte

distribuição condicionada, em que  $\mu_* = -\varepsilon\sigma_u^2 / \sigma^2$  e  $\sigma_*^2 = \sigma_u^2 \sigma_v^2 / \sigma^2$ , Equação 15:

$$f(u|\varepsilon) = \frac{f(u, \varepsilon)}{f(\varepsilon)} = \frac{1}{\sqrt{2\pi}\sigma_*} e^{-\frac{(u-\mu_*)^2}{2\sigma_*^2}} \Big/ \left[ 1 - \Phi\left(-\frac{\mu_*}{\sigma_*}\right) \right] \Rightarrow u_i | \varepsilon_i \sim N^+(\mu_*, \sigma_*^2)$$

A estimativa pontual de  $u_i$  é definida pela média de  $f(u|\varepsilon)$ , Equação 16:

$$\hat{u}_i = E(u_i | \varepsilon_i) = \mu_{*i} + \sigma_* \left[ \frac{\phi(-\mu_{*i}/\sigma_*)}{1 - \Phi(-\mu_{*i}/\sigma_*)} \right] = \sigma_* \left[ \frac{\phi(\varepsilon_i \lambda / \sigma)}{1 - \Phi(\varepsilon_i \lambda / \sigma)} - \frac{\varepsilon_i \lambda}{\sigma} \right]$$

ou pela moda de  $f(u|\varepsilon)$ , Equação 17:

$$\hat{u}_i = M(u_i | \varepsilon_i) = -\varepsilon_i \cdot \sigma_u^2 / \sigma^2, \text{ se } \varepsilon_i \leq 0 \quad (17)$$

$\hat{u}_i = 0$ , caso contrário

Lembrando que  $u_i = -\ln TE_i$ , a estimativa da eficiência técnica do  $i$ -ésimo produtor é, Equação 18:

$$TE_i = e^{-\hat{u}_i} \quad (18)$$

### 3.2 Fronteira de custo

Como na fronteira de produção,  $v_i$  é irrestrita em sinal e captura choques aleatórios não gerenciáveis e  $u_i$  é uma variável não negativa que reflete o grau de ineficiência do produtor. No entanto, na fronteira de custo o erro  $u_i$  captura o efeito da ineficiência econômica, i.e., reflete as parcelas alocativa e técnica da ineficiência. A medida de eficiência econômica da  $i$ -ésima DMU é a razão entre o custo da  $i$ -ésima DMU ( $E_i$ ) e o mínimo custo possível especificado pela função fronteira estocástica  $C(y_i, w_i, \beta)$ , Equação 19:

$$CE_i = E_i / \left[ C(y_i, w_i, \beta) e^{v_i} \right] = e^{u_i} \quad (19)$$

em que:  $w_i = (w_{i1}, \dots, w_{in})$  é o vetor de preços dos *inputs*,  $\beta$  é o vetor de parâmetros da tecnologia de produção,  $y_i$  é o vetor de *outputs* produzidos,  $x_i$  é o vetor de *inputs*,  $E_i = w_i^T x_i$  é o custo total incorrido pela  $i$ -ésima DMU,  $C(y_i, w_i, \beta) e^{v_i}$  é o custo estabelecido pela fronteira estocástica.

Arranjando-se os termos da igualdade na Equação 19, encontra-se a seguinte Equação 20:

$$E_i = C(y_i, w_i, \beta) e^{v_i + u_i}, i = 1, I \quad (20)$$

Neste trabalho são apresentados modelos de fronteira com equação única. Tais modelos baseiam-se no custo total, nas quantidades dos produtos e nos preços dos insumos e não permitem decompor a medida de eficiência econômica nas componentes alocativa e técnica. Se a eficiência alocativa é assumida, o erro  $u_i$  é relacionado com a ineficiência técnica e os modelos apresentados estimam medidas de eficiência técnica, segundo uma orientação ao insumo (*input*).

Assumindo que a parcela determinística da fronteira estocástica de custo,  $C(y_i, w_i, \beta)$ , tenha apenas um produto (*output*), pode-se adotar uma forma funcional Cobb-Douglas que após a transformação logarítmica resulta na seguinte equação, em que  $u_i = -\ln CE_i \geq 0$ , Equação 21:

$$\ln E_i = \beta_0 + \beta_y \ln y_i + \beta_1 \ln w_{i1} \dots + \beta_n \ln w_{in} + v_i + u_i \quad (21)$$

Comparando-se as Equações 7 e 21, pode-se perceber que na fronteira de custo o termo relacionado à ineficiência,  $u_i$ , é adicionado à fronteira. A razão para tal consiste no fato de que a função custo representa o custo mínimo enquanto a função de produção representa o produto máximo. Apesar destas diferenças, a estimação dos parâmetros da Equação 21 e dos índices de eficiência das DMUs é realizada da mesma forma que na fronteira de produção.

### 4 Análise envoltória de dados

Uma tecnologia de produção transforma um vetor de insumos  $X = \{x_1, \dots, x_n\} \in R_+^s$  em um vetor de produtos  $Y = \{y_1, \dots, y_m\} \in R_+^m$ . A união de todas as maneiras possíveis de transformar  $X$  em  $Y$  forma o Conjunto de Possibilidades de Produção (CPP), definido como, Equação 22:

$$T(X, Y) = \{ (X, Y) / \text{é viável produzir } Y \text{ a partir de } X \} \quad (22)$$

Sob o enfoque da conservação de recursos (orientação ao insumo), define-se a eficiência técnica para uma DMU  $(X, Y)$  como sendo a máxima contração radial do vetor de insumos que permite produzir as mesmas quantidades de produtos, ou seja, Equação 23:

$$\text{Eficiência técnica} = \text{Min} \{ \theta / (\theta X, Y) \in T(X, Y) \} \quad (23)$$

A variável  $\theta$  assume um valor entre 0 e 1. Um valor unitário indica que não é possível reduzir a quantidade de insumos e manter a mesma produção. Neste caso a DMU é classificada como sendo tecnicamente eficiente. Caso contrário, há um excesso de insumos que deve ser reduzido ( $\theta < 1$ ) e considera-se a DMU tecnicamente ineficiente.

Com base nestes resultados e admitindo as hipóteses de rendimentos constantes de escala e tecnologia convexa, Charnes, Cooper e Rhodes (1978) propuseram

o modelo DEA-CRS (24). Neste modelo o cálculo da eficiência é formulado como um problema de programação linear (PPL), cuja função objetivo é a máxima contração dos insumos (orientação ao insumo) e as restrições representam o CPP. Posteriormente, Banker, Charnes e Cooper (1984) adicionaram uma restrição de combinação convexa [ $\lambda_1 + \lambda_2 + \dots + \lambda_N = 1$  no PPL em (25)] ao CRS, criando um modelo que contempla a hipótese de rendimentos variáveis de escala, conhecido como DEA-VRS (25).

Na Tabela 1 são apresentados modelos DEA orientados ao insumo nas versões CRS e VRS, em que  $N$  é o total de DMUs analisadas, o par  $(X_j, Y_j)$  representa os vetores de insumos e produtos da  $j$ -ésima DMU,  $j=1, N$  e  $j_0$  denota a DMU avaliada.

Seja  $\theta^*, \lambda_1^*, \dots, \lambda_N^*$  a solução ótima dos modelos (24) ou (25), a DMU $_{j_0}$  é eficiente, se, e somente se,  $\theta^* = 1$  e todas as folgas nas restrições são nulas. Caso contrário, quando  $\theta^* < 1$  ou  $\theta^* = 1$ , porém com alguma folga positiva, a DMU $_{j_0}$  é ineficiente. Neste caso, o conjunto de referência (*benchmarks*) da DMU $_{j_0}$  é formado pelas DMUs associadas aos coeficientes  $\lambda_j^* > 0$ .

### 5 Mapa de Kohonen

O mapa de Kohonen ou rede SOM (*Self Organizing Map*) é uma rede neural com treinamento competitivo e não supervisionado (CARVALHO; BRAGA; LUDERMIR, 1998) aplicada em problemas de classificação. Nesta classe de redes neurais, os neurônios da camada de saída são organizados em uma grade, geralmente uni ou bidimensional, conforme ilustrado na Figura 3.

Esta grade forma um sistema de coordenadas sobre o qual são projetados padrões de entrada com  $m$  atributos e que permite mapear o espaço de entrada, cuja dimensão é  $m$ , em um espaço de saída com dimensão reduzida, neste caso, bidimensional. Cada neurônio recebe todas as variáveis de entrada e

funciona como um discriminador de características, cujos parâmetros (pesos sinápticos) são ajustados por um processo de aprendizagem competitivo.

Quando um padrão de entrada é apresentado, a rede SOM classifica-o no neurônio com o maior valor da função de ativação, ou seja, no neurônio com o vetor de pesos mais parecido com o padrão de entrada. Durante a fase de aprendizado, os neurônios se especializam na detecção de um conjunto de padrões de entrada e se organizam topologicamente, fazendo com que os padrões detectados por um dado neurônio estejam relacionados com a sua posição na grade de saída. Assim, o Mapa de Kohonen obtém um mapeamento entrada-saída, onde a topologia do espaço de entrada é preservada no espaço de saída, de tal forma que padrões de entrada semelhantes ativem neurônios fisicamente próximos na grade de saída.

O treinamento da rede SOM ocorre em duas fases (SOUZA, 2002): fase de ordenação e fase de convergência. Na fase de ordenação, os vetores de pesos, inicialmente orientados de forma aleatória, são topologicamente ordenados de forma a agrupar os neurônios do mapa em *clusters* que reflitam a distribuição espacial dos padrões de entrada. Ao final desta fase a rede descobre quantos *clusters* deve identificar e quais as suas posições relativas no mapa. Durante o treinamento, a taxa de aprendizado, é decrescente. O mesmo acontece com a região de vizinhança que se reduz gradualmente, conforme mostra a Figura 4. A seguir tem-se o algoritmo de treinamento da rede SOM:

1. Inicialize os pesos ( $w_{ij}$ ) e os parâmetros da rede SOM : taxa de aprendizado  $\eta$  e vizinhança. O peso  $w_{ij}$  é o peso da conexão entre o  $i$ -ésimo elemento do padrão de entrada ( $x_i$ ) e o neurônio  $j$ .
2. Para cada padrão de treinamento  $x$ ;
  - 2.1. identifique o neurônio vencedor;
  - 2.2. atualize os pesos deste neurônio e de seus vizinhos. Se o neurônio  $j \in \Lambda(t)$ , i.e., pertencer à vizinhança do neurônio vencedor no instante de tempo  $t$ , faça  $w_{ji}(t+1) = w_{ji}(t) + \eta(t) [x_i(t) - w_{ji}(t)]$ , caso contrário, faça  $w_{ji}(t+1) = w_{ji}(t)$ ;
  - 2.3. Se o número de ciclos de treinamento for múltiplo de  $N$ , reduza a taxa de aprendizado  $\eta$  e a área de vizinhança; e
3. Repetir o passo 2 até que o mapa de características não mude.

### 6 Especificação dos análise envoltória de dados e modelos de fronteira estocástica

Na aplicação dos modelos DEA e SFA, é importante garantir a comparabilidade entre as distribuidoras analisadas. Para atender a este requisito, os dados das

**Tabela 1.** Modelos DEA na versão envelope e orientados ao insumo.

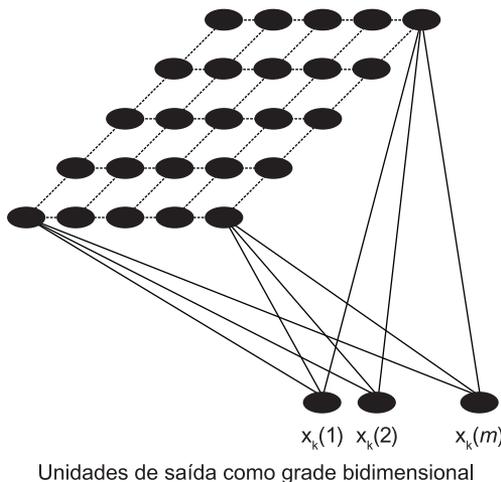
Modelo CRS	Modelo VRS
$Min_{\theta, \lambda} \theta$ (24)	$Min_{\theta, \lambda} \theta$ (25)
s.a.	s.a.
$\theta X_{j_0} \geq \sum_{j=1}^N \lambda_j X_j$	$\theta X_{j_0} \geq \sum_{j=1}^N \lambda_j X_j$
$Y_{j_0} \leq \sum_{j=1}^N \lambda_j Y_j$	$Y_{j_0} \leq \sum_{j=1}^N \lambda_j Y_j$
$\lambda_j \geq 0 \forall j = 1, N$	$\sum_{j=1}^N \lambda_j = 1$
	$\lambda_j \geq 0 \forall j = 1, N$

40 distribuidoras analisadas foram processados por uma rede neural auto-organizável com 25 unidades de saída arranjadas em uma grade quadrada  $5 \times 5$ , conforme apresentado em Pessanha, Souza e Laurencel (2005). A escolha do mapa de Kohonen deve-se ao fato desta rede neural fornecer um mapa topologicamente ordenado das distribuidoras, o que facilita a visualização e a interpretação dos agrupamentos.

Na construção do mapa, caracterizou-se cada concessionária por oito atributos referentes ao tamanho, concentração e composição do mercado, conforme apresentado na Tabela 2. Os dados utilizados referem-se ao ano de 2001 e no treinamento da rede neural tomaram-se os valores padronizados dos seus logaritmos com o objetivo de evitar que a segmentação das distribuidoras se concentrasse apenas nas variáveis mais heterogêneas.

O treinamento da rede neural foi realizado por rotinas da *Toolbox* de redes neurais que acompanha o Matlab. As 40 distribuidoras analisadas foram ordenadas topologicamente no mapa  $5 \times 5$  (25 neurônios) da Figura 5, onde são apresentados os perfis das distribuidoras (cada curva representa uma empresa). A identificação das distribuidoras em cada neurônio pode ser visualizada na Figura 6.

Por ser um mapa topologicamente ordenado, as distribuidoras com perfis semelhantes ocupam o mesmo neurônio (Light e a Eletropaulo) ou ocupam neurônios vizinhos (Light e as concessionárias Bandeirante, CPFL e Piratininga), conforme ilustrado na Figura 6. Com base em uma análise pictórica do mapa de perfis (Figura 6) buscou-se agrupar os neurônios que apresentassem concessionárias com perfis semelhantes. Desta maneira as distribuidoras foram agrupadas em três agrupamentos maiores A, B e C, conforme indicado na Figura 6.



No agrupamento A foram classificadas as distribuidoras que atendem a mercados grandes, concentrados e com expressiva participação da classe industrial. Este grupo é composto pelas principais distribuidoras das regiões Sul e Sudeste. O agrupamento B é formado pelas principais concessionárias da região Nordeste e pelas concessionárias das regiões Centro-Oeste (Celg, Cemat e Enersul). Por fim, o agrupamento C é formado por concessionárias de menor porte.

Com relação à avaliação das eficiências das distribuidoras, esta foi caracterizada por quatro atributos, conforme indicados na Tabela 3. Nesse aspecto, o único *input* é o OPEX e, portanto, as medidas de eficiência obtidas pelos modelos DEA e SFA devem expressar o potencial de redução dos custos operacionais de cada concessionária, i.e., trata-se de medidas orientadas ao insumo. Por outro lado, os *outputs* são os *cost-drivers* do OPEX; assim, foi selecionado o número de consumidores (NC) como uma *proxy* da quantidade de serviço e o montante de energia fornecida (MWh) como uma *proxy* do produto total. A consideração da extensão da rede (KM) como um *output* é interessante, pois esta variável reflete a dispersão dos consumidores na área de concessão, um fator determinante dos custos da concessionária.

No que se refere ao rendimento de escala do modelo DEA, duas opções foram avaliadas: rendimentos constantes (CRS) e rendimentos variáveis (VRS). O modelo DEA/CRS foi aplicado separadamente em cada agrupamento de DMUs, pois, desta forma, pode-se garantir que as DMUs avaliadas sejam comparáveis, i.e., homogêneas com relação ao tamanho. Já no modelo DEA/VRS, a restrição de convexidade  $\lambda_1 + \lambda_2 + \dots + \lambda_N = 1$  garante que as DMUs ineficientes sejam comparadas somente com as DMUs com mesmo tamanho ou nível de atividade (COELLI et al., 2005).

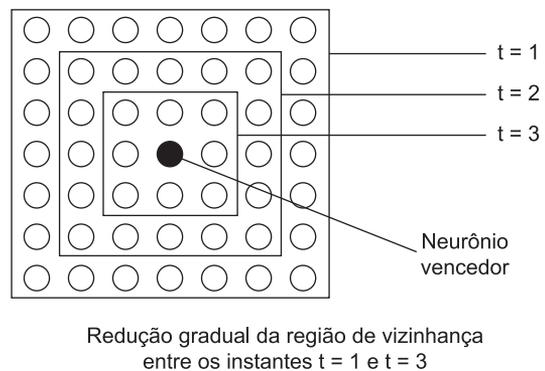


Figura 3. Mapa de Kohonen.

Desta forma, o modelo DEA/VRS pode ser aplicado ao conjunto formado por todas as distribuidoras sem a necessidade de classificá-las previamente. No entanto, a aplicação do modelo VRS ao conjunto de todas as distribuidoras não evita que as DMUs ineficientes em um determinado agrupamento sejam comparadas com as distribuidoras dos demais agrupamentos.

Uma propriedade indesejável do modelo VRS é que as DMUs que apresentam os menores níveis de insumos ou os maiores níveis de produto, em pelo menos uma das variáveis, são classificadas como eficientes. Assim uma distribuidora pode ser considerada eficiente pelo fato de atender ao maior mercado ou por ter a maior rede de distribuição.

Em função da hipótese de convexidade, a fronteira de eficiência estimada pelo modelo VRS é sempre caracterizada por rendimentos crescentes de escala nos menores níveis de atividade e por rendimentos decrescentes de escala nos maiores níveis de atividade. Ressalta-se que esta propriedade do modelo VRS independe dos dados contidos na amostra de DMUs. Em resumo, na fronteira gerada pelo modelo VRS, os rendimentos não são crescentes ao longo de toda fronteira e a produtividade diminui à medida que o nível de atividade aumenta para além da escala ótima. Esta característica da fronteira VRS não atende aos objetivos da regulação econômica, em particular o estímulo aos ganhos de produtividade. Para fins de um esquema de regulação por incentivo, a alternativa ao modelo VRS consiste em estimar a fronteira pelo modelo NDRS (*non-decreasing return to scale*), uma variante do modelo VRS que não admite rendimentos decrescentes de escala, conforme indicado na Figura 7 para o caso de um insumo (X) e um produto (Y) e pelo modelo em 26, em que a restrição  $\lambda_1 + \lambda_2 + \dots + \lambda_N \geq 1$  substitui a restrição convexa  $\lambda_1 + \lambda_2 + \dots + \lambda_N = 1$  no modelo DEA/VRS.

Quanto aos modelos SFA, foram testadas quatro especificações distintas para a fronteira estocástica de custos, resultantes das combinações de duas formas funcionais (Cobb-Douglas e Translog) com duas distribuições para o termo aleatório  $u_i$  (*Half-Normal* e *Normal truncada*) Equações 26 e 27:

Cobb-Douglas

$$\ln OPEX_i = \beta_0 + \beta_1 \ln MWh_i + \beta_2 \ln NC_i + \beta_3 \ln KM_i + v_i + u_i \tag{26}$$

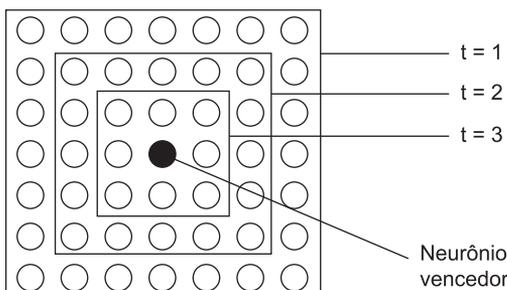
Translog

$$\ln OPEX_i = \beta_0 + \beta_1 \ln MWh_i + \beta_2 \ln NC_i + \beta_3 \ln KM_i + 0,5 \cdot \left( \begin{matrix} \beta_{12} MWh_i NC_i + \\ \beta_{13} MWh_i KM_i + \\ \beta_{23} NC_i KM_i \end{matrix} \right) + v_i + u_i \tag{27}$$

Na distribuição Normal-Truncada,  $u_i \sim iid N(m_i, \sigma_u^2)$  considerou-se a média definida como sendo  $m_i = \delta_0 + \delta_1 \ln(OPEX_i/MWh_i)$ , em que a razão  $OPEX/MWh$  é um índice tradicionalmente considerado na avaliação da produtividade das concessionárias de distribuição de energia elétrica. A aplicação do método DEA foi realizada com o auxílio do *software* DEAP® ([www.uq.edu.au/economics/cepa/software.htm](http://www.uq.edu.au/economics/cepa/software.htm)). Nas distintas versões para o modelo SFA, estimadas com o *FRONTIER 4.1*® ([www.uq.edu.au/economics/cepa/software.htm](http://www.uq.edu.au/economics/cepa/software.htm)), foram adicionadas duas variáveis *dummies* na forma aditiva com o objetivo de informar em que *cluster* se encontra cada distribuidora. Vale informar que, sob a perspectiva da metodologia SFA,

**Tabela 2.** Variáveis consideradas na classificação das distribuidoras.

Dimensões	Variáveis
Estrutura	Participação da classe residencial no consumo (%R)
	Participação da classe comercial no consumo (%C)
	Participação da classe industrial no consumo (%I)
Concentração	Tamanho da rede de distribuição (km)
	Carregamento da rede (MWh/km)
	Consumo médio residencial (CMR)
Tamanho	Nº de unidades consumidoras (NC)
	Energia elétrica distribuída (MWh)



**Figura 4.** Redução gradual da região de vizinhança entre os instantes  $t = 1$  e  $t = 3$ .

**Tabela 3.** Variáveis consideradas na avaliação da eficiência.

Tipo da variável	Variável	Descrição
Insumo (DEA) ou dependente (SFA)	OPEX	Custo operacional (R\$ 1.000)
Produtos (DEA) ou Independentes (SFA)	MWh	Quantidade de energia distribuída (MWh)
	NC	Total de unidades consumidoras atendidas
	KM	Extensão da rede de distribuição (km)

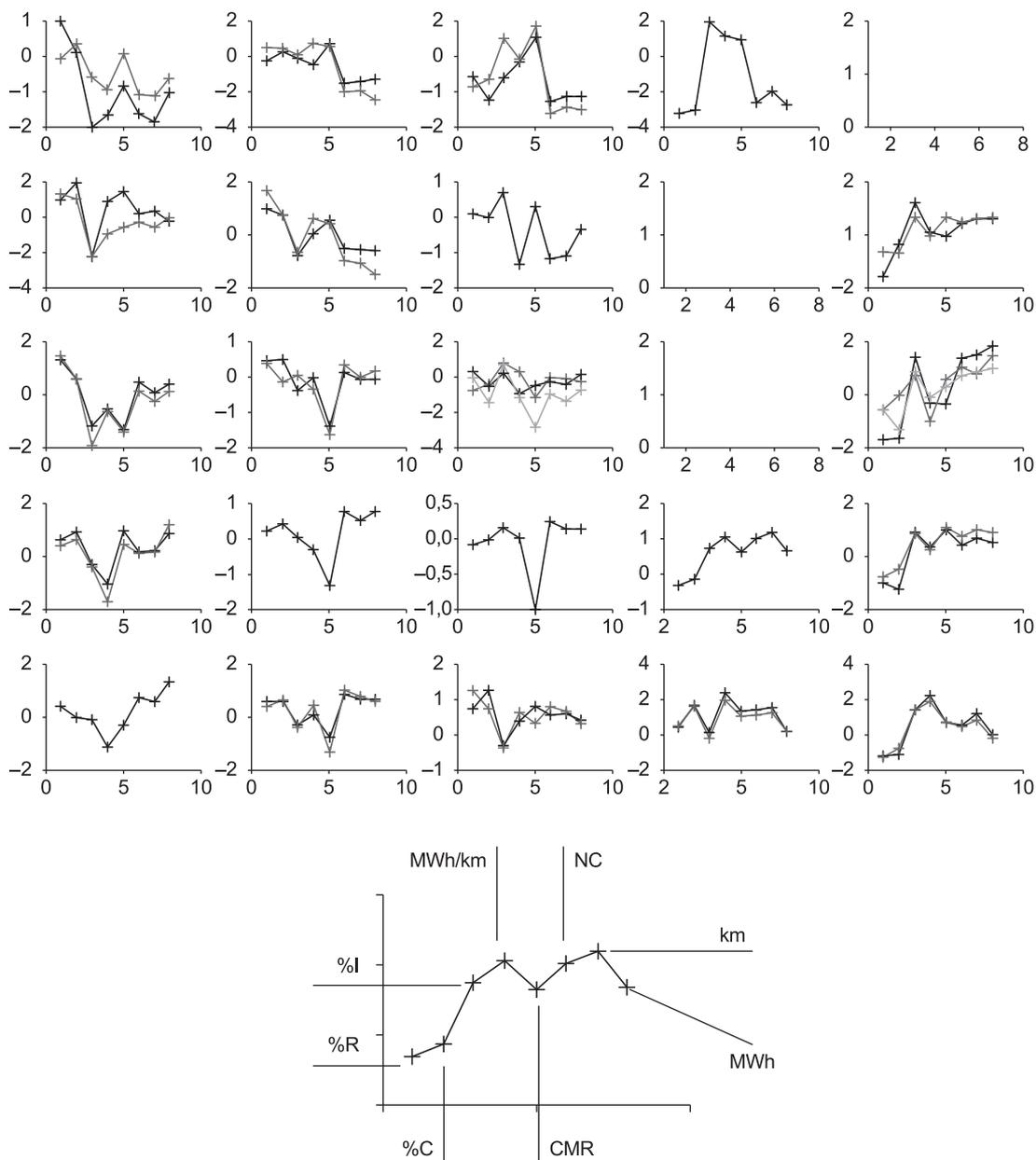


Figura 5. Mapa de Kohonen com os perfis das concessionárias.

Chesp Santa Maria	Eletrocar Panambi	C Cpée Xanxerê	Urussanga	
CEB Celtins	Caiuá Cenf	Poços de Caldas		Escelsa RGE
Cemar Cepisa	Ceal Saelpa	Cat-Leo Energipe Sulgipe		Cemig Copel Elektro
B Cemat Enersul	Coelce	Consern	CPFL	AES-Sil Celesc
Cemat Enersul	Celpe Coelba	CEEE Ampla	A Eletropaulo Light	Bandeirantes Piratinga

Figura 6. Agrupamentos de concessionárias.

**Tabela 4.** Índices de eficiência ( $\theta_i$ ) para as distribuidoras avaliadas.

DMUs ( <i>cluster</i> )	SFA <i>Half Normal</i>		SFA Normal Truncada		DEA	
	Cobb Douglas	Translog	Cobb-Douglas	Translog	CRS <sup>a</sup>	NDRS <sup>b</sup>
	SFA HN CD	SFA HN TL	SFA NT CD	SFA NT TL		
AES-SUL (A)	1,000	0,990	1,000	0,980	1,000	1,000
AMPLA (A)	0,629	0,962	0,382	0,847	0,764	0,744
BANDEIRANTES(A)	0,552	0,645	0,800	0,662	0,799	0,799
CEEE (A)	0,251	0,264	0,190	0,254	0,278	0,274
CELESC (A)	0,424	0,469	0,410	0,461	0,439	0,439
CEMIG (A)	0,532	0,943	0,508	0,962	0,731	0,654
COPEL (A)	0,741	0,971	0,427	1,000	1,000	1,000
CPFL (A)	0,775	0,926	0,694	0,877	0,816	0,810
ELEKTRO (A)	0,758	0,808	0,649	0,813	0,837	0,834
ELETROPAULO (A)	0,595	0,781	0,495	0,794	0,622	0,615
ESCELSA (A)	0,654	0,606	0,585	0,610	0,681	0,682
LIGHT (A)	0,613	0,763	0,521	0,758	0,637	0,631
PIRATININGA (A)	0,909	0,971	0,917	1,000	0,913	0,913
RGE (A)	0,935	0,862	0,787	0,870	0,999	0,999
<b>Média (A)</b>	<b>0,669</b>	<b>0,783</b>	<b>0,598</b>	<b>0,778</b>	<b>0,751</b>	<b>0,742</b>
CAT-LEO (B)	0,602	0,602	0,277	0,602	0,701	0,626
CEAL (B)	0,617	0,617	0,265	0,621	0,606	0,614
CELG (B)	0,431	0,549	0,227	0,606	0,558	0,507
CELPE (B)	1,000	0,943	0,455	1,000	1,000	1,000
CEMAR (B)	0,595	0,617	0,214	0,633	0,691	0,679
CEMAT (B)	0,385	0,562	0,243	0,552	0,590	0,459
CEPISA (B)	0,581	0,617	0,205	0,621	0,683	0,668
COELBA (B)	0,741	0,658	0,331	0,709	0,735	0,735
COELCE (B)	0,775	0,758	0,340	0,800	0,824	0,796
COSERN (B)	0,901	0,870	0,444	0,877	0,977	0,842
ENERGIPE (B)	0,800	0,834	0,420	0,820	0,921	0,717
ENERSUL (B)	0,602	0,926	0,366	1,000	1,000	1,000
SAELPA (B)	0,806	0,833	0,298	0,847	0,890	0,889
SULGIPE (B)	0,735	0,990	0,246	0,926	0,913	0,943
<b>Média (B)</b>	<b>0,684</b>	<b>0,741</b>	<b>0,309</b>	<b>0,758</b>	<b>0,792</b>	<b>0,748</b>
CAUIÁ (C)	0,763	0,917	0,262	0,870	0,811	0,481
CEB (C)	0,518	0,704	0,217	0,680	0,563	0,291
CELTINS (C)	0,541	0,855	0,150	0,794	0,606	0,394
CENF (C)	0,870	0,943	0,262	0,980	0,887	0,599
CHESP (C)	0,935	0,787	0,237	0,800	1,000	1,000
CPEE (C)	0,962	1,000	0,388	0,935	1,000	0,683
ELETROCAR (C)	0,847	0,826	0,310	0,775	0,870	0,722
P. DE CALDAS (C)	0,763	0,943	0,303	0,885	1,000	0,750
PANAMBI (C)	0,741	0,862	0,262	0,935	0,715	0,979
SANTA MARIA (C)	0,840	0,901	0,291	0,855	0,907	0,637
URUSSANGA (C)	0,412	0,926	0,267	0,813	0,690	1,000
XANXERÊ (C)	0,606	0,625	0,253	0,581	0,651	0,547
<b>Média (C)</b>	<b>0,729</b>	<b>0,848</b>	<b>0,270</b>	<b>0,820</b>	<b>0,807</b>	<b>0,674</b>

<sup>a</sup>O modelo DEA/CRS é aplicado ao conjunto de empresas em um mesmo agrupamento. <sup>b</sup>O modelo DEA/NDRS foi aplicado ao conjunto forado pelas 40 distribuidoras.

$$\begin{aligned}
 & \text{Min}_{\theta, \lambda} \theta & (26) \\
 & \text{s.a.} \\
 & \theta X_{j0} \geq \sum_{j=1}^N \lambda_j X_j \\
 & Y_{j0} \leq \sum_{j=1}^N \lambda_j Y_j \\
 & \sum_{j=1}^N \lambda_j \geq 1 \\
 & \lambda_j \geq 0 \forall j = 1, N
 \end{aligned}$$

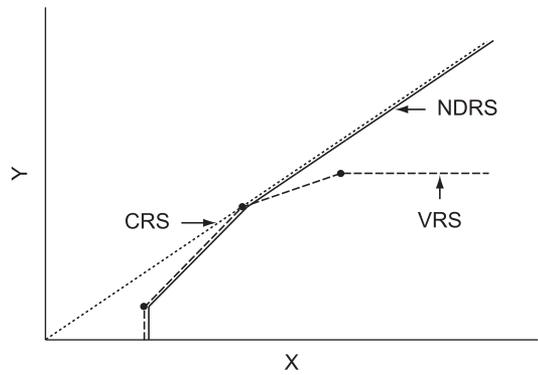


Figura 7. Modelo VRS e fronteiras de eficiência estimadas por diferentes modelos DEA.

as eficiências são calculadas usando o preditor baseado no valor esperado de (18), o qual é programado no software FRONTIER 4.1<sup>®</sup>. Matematicamente, tem-se, Equação 28:

$$TE_i = E(e^{-\hat{u}_i}) \quad (28)$$

### 7 Resultados

Na Tabela 4, são apresentadas as medidas de eficiência para as 40 distribuidoras analisadas. Conforme revelam as médias dos índices de eficiência na Tabela 4 e os *boxplots* na Figura 8, os índices de eficiência obtidos pelo modelo SFA, com especificação Cobb-Douglas e Normal truncada (SFA NT CD), são bastante discrepantes dos índices encontrados pelos demais modelos. Contrastando com resultados dos modelos SFA/Cobb-Douglas, os índices de eficiência obtidos pelos modelos SFA/Translog são mais aderentes aos índices obtidos pelos modelos DEA CRS e NDRS, em especial nos *clusters* A e B conforme evidenciado pelos *boxplots* na Figura 8 e pelos elevados coeficientes Pearson e de Spearman (SPIEGEL, 1993) na Tabela 5. Vale observar que em todos os modelos SFA a estatística  $\gamma$  é significativa e assume um valor igual a 0,999, indicando a presença da ineficiência técnica, e que nos modelos SFA com normal truncada os coeficientes  $\delta_0$  e  $\delta_1$  são significativos.

Os modelos DEA/CRS e DEA/NDRS são aplicados em diferentes conjuntos de concessionárias de distribuição, o primeiro compara apenas as empresas classificadas em um mesmo agrupamento, enquanto o segundo compara as 40 empresas analisadas. Esta diferença explica os desvios entre os índices obtidos pelos dois modelos e também o fato de que na maioria dos casos os índices de eficiência obtidos pelo modelo CRS são maiores que os encontrados pelo modelo NDRS. Esta última observação deve-se à perda do poder de discriminação dos modelos DEA provocada pela redução do número de DMUs avaliadas. Note

que na implementação efetuada o modelo NDRS é aplicado ao conjunto das 40 distribuidoras, enquanto o modelo CRS é aplicado em conjuntos com 12 e 14 distribuidoras.

Os resultados na Tabela 4, juntamente com os *boxplots* na Figura 8 e as correlações na Tabela 5, mostram que as maiores discrepâncias entre os índices de eficiência obtidos pelos modelos CRS e NDRS residem nas empresas classificadas no *cluster* C, distribuidoras de menor porte e que por isso operam em uma região no início da fronteira eficiente VRS, caracterizada por rendimentos crescentes de escala (Figura 7). Para estas empresas a fronteira NDRS confunde-se com a fronteira VRS, a qual inclui as DMUs com os menores valores nos insumos, neste caso a distribuidora Urussanga (eficiência igual 1 no modelo NDRS), pelo fato de apresentar o menor custo operacional, a única variável insumo.

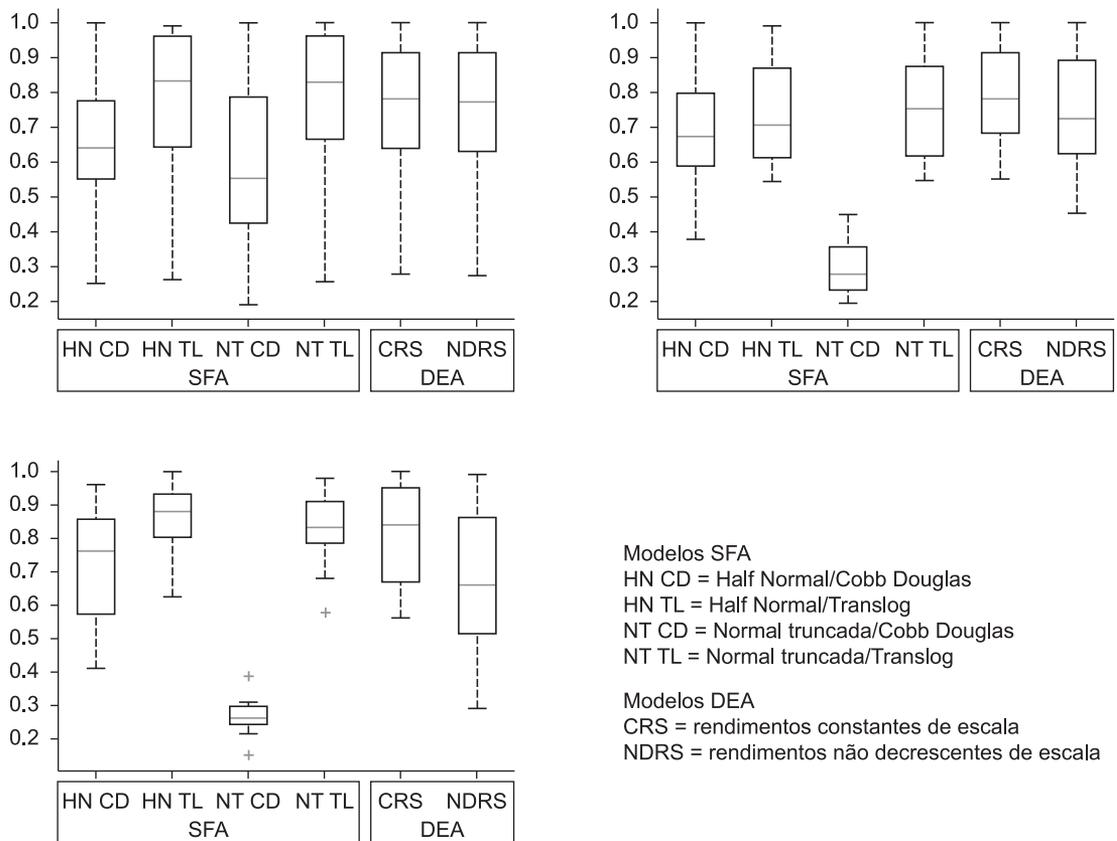
Por sua vez, as distribuidoras do agrupamento A operam na região de rendimentos decrescentes de escala, no final da fronteira VRS, pois atendem aos maiores mercados. Para estas empresas a fronteira NDRS confunde-se com uma fronteira de rendimentos constantes de escala, conforme ilustrado na Figura 7. Por esta razão, os índices obtidos pelos modelos CRS e NDRS (Tabela 4) são fortemente correlacionados nas empresas classificadas no agrupamento A (Tabela 5), inclusive os dois modelos identificaram as mesmas empresas eficientes: Aes-Sul e Copel.

As empresas do agrupamento B ocupam uma posição intermediária na fronteira VRS com algumas distribuidoras na região de rendimentos decrescentes de escala e outras na região com rendimentos crescentes de escala. Para empresas na região da fronteira VRS com rendimentos decrescentes de escala, o modelo NDRS calcula os índices de eficiência tomando como referência uma fronteira com rendimentos constantes de escala, já para as empresas com rendimentos crescentes de escala, estas provavelmente ocupam uma região em que a fronteira VRS é próxima da fronteira com rendimentos constantes de escala. Assim,

**Tabela 5.** Coeficientes de correlação (Pearson; Spearman) entre os índices de eficiência.

	Modelos	SFA HN CD	SFA HN TL	SFA NT CD	SFA NT TL	DEA CRS	DEA NDRS
Cluster A	sfa hn tl	0,78 <sup>a</sup> ; 0,69 <sup>a</sup>	1				
	sfa nt cD	0,82 <sup>a</sup> ; 0,71 <sup>a</sup>	0,53; 0,39	1			
	sfa nt tl	0,80 <sup>a</sup> ; 0,68 <sup>a</sup>	0,99 <sup>a</sup> ; 0,96 <sup>a</sup>	0,57 <sup>b</sup> ; 0,43	1		
	dea-CRS	0,91 <sup>a</sup> ; 0,85 <sup>a</sup>	0,85 <sup>a</sup> ; 0,80 <sup>a</sup>	0,73 <sup>a</sup> ; 0,64 <sup>b</sup>	0,88 <sup>a</sup> ; 0,82 <sup>a</sup>	1	
	DEA-NDRS	0,92 <sup>a</sup> ; 0,87 <sup>a</sup>	0,82 <sup>a</sup> ; 0,77 <sup>a</sup>	0,73 <sup>a</sup> ; 0,65 <sup>b</sup>	0,84 <sup>a</sup> ; 0,78 <sup>a</sup>	0,99 <sup>a</sup> ; 0,99 <sup>a</sup>	1
Cluster B	sfa hn tl	0,74 <sup>a</sup> ; 0,74 <sup>a</sup>	1				
	sfa nt cD	0,78 <sup>a</sup> ; 0,83 <sup>a</sup>	0,65 <sup>b</sup> ; 0,67 <sup>a</sup>	1			
	sfa nt tl	0,74 <sup>a</sup> ; 0,72 <sup>a</sup>	0,97 <sup>a</sup> ; 0,96 <sup>a</sup>	0,71 <sup>a</sup> ; 0,68 <sup>a</sup>	1		
	dea-CRS	0,80 <sup>a</sup> ; 0,77 <sup>a</sup>	0,95 <sup>a</sup> ; 0,92 <sup>a</sup>	0,78 <sup>a</sup> ; 0,83 <sup>a</sup>	0,96 <sup>a</sup> ; 0,92 <sup>a</sup>	1	
	DEA-NDRS	0,75 <sup>a</sup> ; 0,70 <sup>a</sup>	0,93 <sup>a</sup> ; 0,93 <sup>a</sup>	0,59 <sup>b</sup> ; 0,65 <sup>b</sup>	0,95 <sup>a</sup> ; 0,97 <sup>a</sup>	0,92 <sup>a</sup> ; 0,92 <sup>a</sup>	1
Cluster C	sfa hn tl	0,37; 0,35	1				
	sfa nt cD	0,58 <sup>a</sup> ; 0,47	0,48; 0,64 <sup>b</sup>	1			
	sfa nt tl	0,53; 0,46	0,90 <sup>a</sup> ; 0,85 <sup>a</sup>	0,39; 0,44	1		
	dea-CRS	0,86 <sup>a</sup> ; 0,85 <sup>a</sup>	0,56; 0,55	0,68 <sup>b</sup> ; 0,62 <sup>b</sup>	0,56; 0,56	1	
	DEA-NDRS	0,23; 0,26	0,28; 0,25	0,33; 0,45	0,34; 0,30	0,43; 0,54	1

<sup>a</sup>Significativo ao nível de 1%. <sup>b</sup>Significativo ao nível de 5%.



**Figura 8.** Boxplots dos índices de eficiência segundo diferentes modelos.

no agrupamento B, os índices de eficiência obtidos pelos modelos NDRS e CRS são bem aderentes, conforme indicado pelas correlações na Tabela 5, inclusive as empresas eficientes são as mesmas:

Celpe e Enersul. Note que esta aderência é menor que a observada no agrupamento A, porém muito superior à observada no agrupamento C, no qual o pequeno porte das empresas permite inferir que elas

ocupam uma região no início da fronteira VRS com rendimentos crescentes de escala e bem distante da fronteira de rendimentos constantes de escala do modelo NDRS.

Um importante resultado dos modelos DEA é a identificação do conjunto de referência (*peer set*) para cada distribuidora ineficiente. Este conjunto fornece um *benchmark* a ser perseguido pela empresa ineficiente, conferindo uma transparência ao processo de avaliação dos custos operacionais eficiente.

No modelo CRS, a segmentação por *cluster* garante que o *peer set* de uma distribuidora ineficiente é formado apenas por empresas eficientes classificadas no mesmo *cluster* onde se encontra a distribuidora. Por sua vez, na implementação do modelo NDRS não há a segmentação por *cluster* e o *peer set* de uma empresa ineficiente pode ser constituído por empresas classificadas em outros agrupamentos e com pouca similaridade com a distribuidora avaliada, portanto, o *benchmark* pode não ser factível. Por esta razão, a aplicação do modelo DEA/CRS segmentada por *cluster* parece ser uma estratégia mais atraente.

## 8 Conclusões

Nos próximos ciclos de revisão tarifária a ANEEL sinaliza com a possibilidade de utilizar modelos DEA e SFA na determinação dos custos operacionais eficientes, um elemento fundamental no cálculo do Fator X das distribuidoras. As duas abordagens partem de pressupostos diferentes; o método DEA é determinístico e os desvios em relação à fronteira de eficiência são tomados como sendo unicamente devido às ineficiências das distribuidoras. Além disso, conforme explicado em (SOUZA et al. 2010), os modelos DEA são sensíveis a *outliers* e pontos discrepantes, afetando, portanto, sensivelmente na mensuração das eficiências. Já o método SFA tem uma natureza estocástica e estima uma medida de eficiência livre dos impactos dos fatores aleatórios não controláveis pelas DMUs. No entanto, para este caso, o problema diz respeito às especificações da forma funcional e distribuição de probabilidade para as eficiências.

De qualquer forma, independentemente do método a ser escolhido pelo regulador, é interessante utilizar as duas abordagens, pois uma complementa a outra e permite avaliar a robustez dos resultados. Apesar das significativas diferenças metodológicas entre os modelos DEA e SFA/Translog, as respectivas medidas de eficiências para as 40 distribuidoras analisadas são aderentes.

Por fim, na passagem dos índices de eficiência para os custos operacionais eficientes, a ANEEL pode considerar um critério semelhante ao adotado para as transmissoras, no qual são reconhecidos pelo

menos 80% dos custos operacionais. Isto implica na seguinte normalização (BRASIL, 2007) dos índices de eficiência ( $\theta_i$ ) obtidos pelos modelos DEA e SFA, antes de serem aplicados no cálculo dos custos operacionais eficientes  $OPEX_i (1 - \theta_i)$  Equação 28:

$$\theta_i^* = 0,20 \cdot [\theta_i - \min(\theta_i)] / [\max(\theta_i) - \min(\theta_i)] + 0,80 \quad (28)$$

## Referências

- ARCOVERDE, F. D.; TANNURI-PIANTO, M. E.; SOUSA, M. C. S. Mensuração das eficiências das distribuidoras do setor energético brasileiro usando fronteiras estocásticas. In: ENCONTRO NACIONAL DE ECONOMIA, 33., 2005, Natal. **Anais...**
- BANKER, R. D.; CHARNES, A.; COOPER, W. W. Some models for estimating technical and scale inefficiencies. **Management Science**, v. 39, 1984.
- BRASIL. Agência Nacional de Energia Elétrica – ANEEL. **Nota Técnica nº 52/2003**. Brasília, DF, 2003.
- BRASIL. Agência Nacional de Energia Elétrica – ANEEL. **Nota Técnica nº 262/2006**. Brasília, DF, 2006.
- BRASIL. Agência Nacional de Energia Elétrica – ANEEL. **Nota Técnica nº 125/2007**. Brasília, DF, 2007.
- CARVALHO, A. C. P. L. F.; BRAGA, A. P.; LUDERMIR, T. B. **Fundamentos de redes neurais artificiais**. 11. ed. Rio de Janeiro: Escola de Computação, 1998.
- CHARNES, A. et al. Comparisons of DEA and existing ratio and regression systems for effecting efficiency evaluation of regulated electric cooperatives in Texas. **Research in Governmental and Nonprofit Accounting**, v. 5, p. 187-210, 1989.
- CHARNES, A.; COOPER, W. W.; RHODES, E. Measuring the efficiency of decision making units. **European Journal of Operational Research**, v. 2, 1978.
- COELLI, T. J. et al. **An introduction to efficiency and productivity analysis**. 2<sup>nd</sup> ed. Springer, 2005.
- COOPER, W. W.; SEIFORD, L. M.; TONE, K. **Data envelopment analysis: a comprehensive text with models applications, reference and dea-solver software**. Kluwer Academic Publishers, 2000.
- ESTACHE, A.; ROSSI, M. A.; RUZZIER, C. A. The case for international coordination of electricity regulation: evidence from the measurement of efficiency in South America. **Journal of Regulatory Economics**, v. 25, n. 3, p. 271-295, 2004.
- FÄRE, R.; GROSSKOPF, S.; LOGAN, J. The relative efficiency of Illinois electric utilities. **Resources and Energy Economics**, 1983.
- JASMAB, T.; POLLIT, M. Benchmarking and regulation: international electricity experience. **Utilities Policy**, v. 9, n. 3, p. 107-130, 2000.
- JASMAB, T.; POLLIT, M. International benchmarking and regulation: an application to European electricity distribution utilities. **Energy Policy**, v. 31, p. 121-130, 2003.
- KUMBHAKAR, S. C.; LOVELL, C. A. K. **Stochastic frontier analysis**. Cambridge, 2000.

- PESSANHA, J. F. M.; SOUZA, R. C.; LAURENCEL, L. C. Usando DEA na avaliação da eficiência operacional das distribuidoras do setor elétrico brasileiro. In: CONGRESO LATINO-IBEROAMERICANO DE INVESTIGACIÓN DE OPERACIONES Y SISTEMAS, 12., 2004, Havana, Cuba. **Anais...**
- PESSANHA, J. F. M.; SOUZA, R. C.; LAURENCEL, L. C. Utilizando a análise envoltória de dados na regulação da continuidade do fornecimento de energia elétrica. In: SIMPÓSIO BRASILEIRO DE PESQUISA OPERACIONAL, 37., 2005, Gramado. **Anais...**
- RESENDE, M. Relative efficiency measurement and prospects for yardstick competition in Brazilian electricity distribution. **Energy Economics**, 2001.
- SOLLERO, M. K. V.; LINS, M. P. E. Avaliação de eficiência de distribuidoras de energia elétrica através da análise envoltória de dados com restrições aos pesos. In: SIMPÓSIO BRASILEIRO DE PESQUISA OPERACIONAL, 36., 2004, São João del Rei. **Anais...**
- SOUZA, M. V. P. et al. The cost efficiency of the Brazilian electricity distribution utilities: a comparison of Bayesian SFA and DEA models. **Mathematical Problems in Engineering**, p. 1-20, 2010.
- SOUZA, M. V. P. **Identificação da eficiência de empresas de telecomunicações empregando análise de envoltória de dados e redes neurais de Kohonen**, Dissertação de Mestrado em Engenharia Mecânica-Aeronáutica, Instituto Tecnológico de Aeronáutica, São José dos Campos, 2002.
- SOUZA, M. V. P. **Uma abordagem Bayesiana para o cálculo dos custos operacionais eficientes das distribuidoras de energia elétrica**. 2008. Tese (Doutorado em Engenharia Elétrica)-PUC-Rio, Rio de Janeiro, 2008.
- SOUZA, M. V. P.; SOUZA, R. C.; PESSANHA, J. F. M. Custos operacionais eficientes das distribuidoras de energia elétrica: comparando modelos DEA e SFA. In: SIMPÓSIO BRASILEIRO DE PESQUISA OPERACIONAL, 39., 2007, Fortaleza. **Anais...**
- SPIEGEL, M. R. **Estatística**. São Paulo: Makron Books, 1993.
- VIDAL, D. N. A.; TÁVORA JUNIOR, J. L. Avaliação da eficiência técnica das empresas de distribuição de energia elétrica brasileiras utilizando a metodologia DEA. In: SIMPÓSIO BRASILEIRO DE PESQUISA OPERACIONAL, 35., 2003, Natal.
- ZANINI, A. **Regulação econômica no setor elétrico brasileiro: uma metodologia para definição de fronteiras de eficiência e cálculo do fator X para empresas distribuidoras de energia elétrica**. Tese (Doutorado em Engenharia Elétrica)-PUC-Rio, Rio de Janeiro, 2004.