

INVARIÂNCIA POR TRANSLAÇÃO EM MODELOS DEA DDF: UMA APLICAÇÃO EM ECOEFICIÊNCIA

Lidia Angulo Meza

Departamento de Engenharia de Produção (TEP) – Universidade Federal Fluminense
Rua Passo da Pátria 156, Bloco D, sala 306, São Domingos, 24210-240, Niterói – RJ
lidiaangulomeza@id.uff.br

Eliane Gonçalves Gomes

Embrapa Solos
Rua Jardim Botânico 1024, Jardim Botânico, 22460-000, Rio de Janeiro – RJ
eliane.gomes@embrapa.br

RESUMO

Na presença de variáveis negativas em modelos DEA é possível transladar eixos de modo que todas as variáveis se tornem positivas. Tal situação garante invariância em certas condições de retornos de escala e de orientação. O procedimento é análogo em casos que adotam a função de distância direcional (*directional distance function* – DDF) como estratégia de alcance da fronteira de eficiência. Neste artigo, propõe-se uma medida de ecoeficiência DEA DFF com retornos variáveis de escala e translação para medir o desempenho de sistemas de produção agropecuária, na presença de variáveis com valores positivos e negativos. Além do cálculo da medida de ecoeficiência, é de interesse a identificação de determinantes ou fatores causais da eficiência, o que foi realizado por meio do ajuste de regressão fracionária. Os resultados são comparados a estudo anterior, no qual todas as variáveis de DEA foram modeladas como *ranks* para contornar a presença de valores negativos.

PALAVRAS CHAVE. Translação. DDF. Ecoeficiência.

Tópicos: SE-DEA – Sessão Especial de Desenvolvidos Teóricos e Aplicações de Análise Envoltória de Dados

ABSTRACT

In the presence of negative variables in DEA models, it is possible to translate axes so that all variables become positive. This is only possible under certain conditions of returns to scale and model orientation. The procedure is analogous in cases where the directional distance function (DDF) is the strategy for reaching the efficiency frontier. In this paper we propose a DEA DFF eco-efficiency score under variable returns to scale and translation to measure the performance of agricultural production systems in the presence of variables that can take on negative values. In addition to calculating the eco-efficiency measures, it is our interest to identify the determinants or causal factors of efficiency, which we did by fitting a fractional regression. We compared the results to a previous study, in which we modeled all DEA variables as *ranks* to avoid the presence of negative values.

KEYWORDS. Translation. DDF. Eco-efficiency.

Paper topics: Special Session 2: Data Envelopment Analysis – DEA

1. Introdução

Em modelos de análise de envoltória de dados (DEA) é possível calcular escores de eficiência na presença de variáveis negativas, por meio de translação prévia de eixos da variável em questão. Nestes casos, são impostas restrições quanto às hipóteses de retornos de escala e de orientação permitidas. Conforme Cooper et al. (2007), o pressuposto de retornos constantes de escala não é invariante por translação e não deve ser usado nestes casos. Em modelos sob a hipótese de retornos variáveis de escala (*variable returns to scale* – VRS) em que a variável transladada seja um *input*, a orientação a *outputs* é invariante. Já os modelos VRS em que a variável transladada é *output* são invariantes na orientação a *inputs*.

Em estudos nos quais se deseja calcular uma medida de ecoeficiência não são incomuns a presença de produtos indesejáveis do processo produtivo e de variáveis com valores negativos. Por exemplo, a variável renda líquida, representativa da dimensão econômica da ecoeficiência, pode assumir valores negativos, assim como variáveis de caráter ambiental, como emissão de poluentes. O uso dessas variáveis, chamadas de *outputs* indesejáveis, tem sido amplamente estudado na literatura DEA [Halkos e Petrou, 2019]. Modelos DEA que adotam a estratégia de alcance da fronteira via funções de distância direcional (*directional distance function* – DDF) tem-se mostrado mais adequados em modelos com presença de *outputs* indesejáveis [Wang et al., 2019].

Neste artigo, são adotados os dois procedimentos de modelagem, a translação de eixos e o modelo DDF, para estudar a ecoeficiência de sistemas de produção agropecuária. Foram usados os dados do estudo de Gomes e Angulo-Meza (2024). Além da medida de ecoeficiência, são identificados seus determinantes por meio do ajuste de um modelo de regressão fracionária. Os resultados são comparados aos de Gomes e Angulo-Meza (2024), em que *ranks* foram usados para superar a presença de variáveis negativas no cálculo dos escores de ecoeficiência DEA.

2. Referencial teórico

Adotou-se o conceito de ecoeficiência para caracterizar o desempenho das unidades produtivas: maior produção, com menor uso de recursos e menor geração de poluentes. O enfoque adotado é o de Kuosmanen e Kortelainen (2005), que considera os aspectos econômico/monetário e ambiental. A dimensão econômica é caracterizada pelo lucro (ganhos - custos). A dimensão ambiental contempla pressões ou impactos ambientais gerados pelo sistema produtivo.

Em paralelo, funções de distância direcionais (DDF) são ferramentas importantes em teoria da produção. Conforme destacam Wang et al. (2019), o uso de DDF é baseado em modelagem DEA quando associado ao paradigma não-paramétrico para a estimativa de eficiência. Os modelos DDF têm-se mostrado úteis na modelagem de funções de produção na presença de *outputs* indesejáveis, pois permitem que *outputs* desejáveis possam ser expandidos simultaneamente à redução dos indesejáveis. Além disso, os modelos DDF levam em conta a relação entre as variáveis (*inputs*, *outputs* desejáveis e indesejáveis) denominada de descartabilidade [Färe e Grosskopf, 2004].

Neste artigo, o modelo DDF considerando VRS é apresentado em (1), onde β é a medida de ineficiência, g_y é a componente do vetor direcional, b é o vetor de *outputs* indesejáveis, x representa *inputs*, y *outputs* desejáveis, λ contribuições para cálculo do alvo. A escolha da direção do modelo foi feita baseada no artigo de Aparicio et al. (2016), em que foram discutidos os efeitos da translação de eixos em modelos DDF quando da presença de *inputs* ou *outputs* com valores negativos. Os autores estabeleceram a condição necessária e suficiente para caracterizar a invariância por translação em modelos DDF VRS, a qual depende, exclusivamente, do vetor direcional selecionado. Assim, dado que a variável com valor negativo é um *output* indesejável, a translação foi realizada para esta variável e o modelo DDF VRS é orientado ao *output* desejável.

Foi ajustado um modelo de regressão fracionária no segundo estágio, que assume $E(\hat{\theta}_n(x_j, y_j | z_j)) = G(\delta' z_j)$, onde z_j é o vetor de variáveis contextuais (fatores causais) para a DMU j ; $G(\cdot)$ é uma função não linear com valores em $(0,1]$; θ é o escore de ecoeficiência; x é o vetor de *inputs*; y é o vetor de *outputs*; δ é um vetor de parâmetros. Ver Ramalho et al. (2010) para detalhes.

$$\begin{aligned}
 & \text{Max } \beta \\
 & \text{sujeito a} \\
 & \sum_{k=1}^n x_{ik} \lambda_k \leq x_{io}, \forall i \\
 & \sum_{k=1}^n y_{jk} \lambda_k - \beta g_y \geq y_{jo}, \forall j \\
 & \sum_{k=1}^n b_{rk} \lambda_k \leq b_{ro}, \forall r \\
 & \sum_{k=1}^n \lambda_k = 1 \\
 & \lambda_k \geq 0, \forall k; \beta \in \Re
 \end{aligned} \tag{1}$$

3. Caso de estudo

Os dados aqui usados são os apresentados em Gomes e Angulo-Meza (2024). Referem-se a 22 sistemas modais de produção agropecuária, localizados no estado do Mato Grosso. As variáveis que caracterizam as medidas de ecoeficiência são: lucro (diferença entre valor da produção e custos; R\$), modelada como *output*; perda superficial de solo (kg) e emissões de gases de efeito estufa – GEE (ton CO₂ equivalente), *outputs* indesejáveis e modeladas como *inputs*, seguindo o enfoque de Kuosmanen e Kortelainen (2005) para ecoeficiência mencionado anteriormente.

A variável emissão de GEE apresenta valores positivos e negativos, indicativo de emissão e remoção de carbono, respectivamente. Em Gomes e Angulo-Meza (2024), a análise considerou todas as variáveis transformadas em *ranks* para evitar a translação de eixos. Neste artigo, foram usados os valores originais, com translação da variável com valores negativos. Foi somado um valor suficientemente grande à variável GEE, de forma que todos seus valores fossem maiores que zero. Os fatores causais da eficiência testados são: tipo de sistema de produção (variável categórica em três níveis: integrado, lavoura, pecuária), bioma (variável categórica em cinco níveis: Amazônia, Amazônia-Cerrado, Amazônia-Cerrado-Pantanal, Cerrado, Cerrado-Pantanal), log do percentual de área com cobertura florestal em relação ao definido pelo Código Florestal para cada bioma (floresta), log de número de anos de estudo do produtor (escolaridade) e acesso a plano de saúde pelo produtor (variável binária sim-não).

4. Resultados e discussão

Das 22 DMUs, sete foram eficientes no modelo DDF-VRS-O com translação. Foram identificadas as mesmas DMUs eficientes do modelo de Gomes e Angulo-Meza (2024), com exceção de duas. A correlação de Pearson e de Spearman entre as duas matrizes de eficiência foi de 81,0% e 92,7%, respectivamente.

A Tabela 1 traz os resultados do ajuste da regressão de segundo estágio para a identificação dos determinantes da ecoeficiência. A correlação entre valores observados e preditos deste modelo é de 75,0%, próxima à do modelo anterior com eficiência calculada com *ranks* das variáveis e modelo DEA clássico (79,9%). Na Tabela 1 observa-se diferença estatisticamente significativa entre sistemas integrados e de pecuária, como vantagem em termos de ecoeficiência para os primeiros, mas não entre os de lavoura e os de pecuária. Este último fato difere do modelo anterior, no qual foi identificada diferença significativa entre sistemas de lavoura e de pecuária, com vantagem para lavoura. Quanto aos biomas, os dois modelos mostram que sistemas de produção localizados no bioma Amazônia e Amazônia-Cerrado são os de maior ecoeficiência, com diferença para os localizados em região de transição Cerrado-Pantanal, os de menor desempenho. Quanto às variáveis de caráter socioeconômico, o modelo da Tabela 1 sugere que não há evidência

de efeito de percentual de cobertura florestal, acesso a plano de saúde e nível de escolaridade nos escores de ecoeficiência. No caso do modelo em *ranks*, não ter plano de saúde afetava de modo negativo os escores DEA.

Tabela 1: Resultados do ajuste do modelo de regressão fracionária (Probit).

	Coeficiente	Desvio padrão	z	P> z	Intervalo confiança 95%	
Sistema integrado	1,4997	0,7341	2,04	0,041	0,0608	2,9386
Sistema lavoura	0,0798	0,3704	0,22	0,829	-0,6461	0,8056
Amazônia	2,1816	0,4378	4,98	0,000	1,3235	3,0397
Amazônia-Cerrado	1,9032	0,4193	4,54	0,000	1,0814	2,7249
Amazônia-Cerrado-Pantanal	0,9233	0,2883	3,20	0,001	0,3583	1,4883
Cerrado	1,6067	0,3355	4,79	0,000	0,9491	2,2642
Cobertura florestal	0,0877	0,1336	0,66	0,512	-0,1743	0,3496
Plano de saúde=não	-0,8322	0,6571	-1,27	0,205	-2,1201	0,4557
Escolaridade	0,5404	0,5133	1,05	0,292	-0,4657	1,5465
Constante	-2,6143	1,3815	-1,89	0,058	-5,3220	0,0933

4. Considerações finais

Dada a presença de variáveis negativas que compõem os modelos de eficiência do tipo DEA, a translação de eixos é viável em estratégias DDF, desde que o vetor direcional seja adequadamente selecionado para preservar a invariância por translação.

Apesar da alta correlação entre os escores de ecoeficiência medidos pelo modelo com variáveis em *ranks* e escores DEA com o modelo com translação de eixos e uso de DDF, os resultados do segundo estágio não são exatamente os mesmos. Ambos os ajustes foram concordantes ao identificar os sistemas de produção integrados como os que despontam em ecoeficiência, assim como aqueles localizados em bioma Amazônia.

Os achados aqui apresentados sugerem que a escolha da estratégia de modelagem pode influenciar os resultados e as recomendações deles derivados. Nesse contexto, vale considerar as hipóteses de base de cada modelo para a decidir sobre a escolha da melhor medida de ecoeficiência e dos ajustes de segundo estágio decorrentes para a identificação de fatores causais.

Referências

- Aparicio, J., Pastor, J.T. e Vidal, F. (2016). The directional distance function and the translation invariance property. *Omega*, 58:1–3.
- Cooper, W.W., Seiford, L.M. e Tone, K. (2007). *Data Envelopment Analysis: A Comprehensive Text with Models, Applications, References and DEA-Solver Software*. 2nd Ed. Springer, New York.
- Gomes, E.G. e Angulo-Meza, L. (2024). Proposta de uma medida de ecoeficiência DEA para avaliar sistemas de produção agropecuária. In: *Anais do LVI Simpósio Brasileiro de Pesquisa Operacional*, p. 308176, Campinas: Galoá.
- Färe, R. e Grosskopf, S. (2004). Modeling undesirable factors inefficiency evaluation: Comment. *European Journal of Operational Research* 157:242–245.
- Halkos, G. e Petrou, K.N. (2019). Treating undesirable outputs in DEA: A critical review. *Economic Analysis and Policy*, 62:97–104.
- Kuosmanen, T. e Kortelainen, M. (2005). Measuring eco-efficiency of production with data envelopment analysis. *Journal of Industrial Ecology*, 9(4):59–72.
- Ramalho, E.A., Ramalho, J.J.S. e Henriques, P.D. (2010). Fractional regression models for second stage DEA efficiency analyses. *Journal of Productivity Analysis*, 34:239–255.
- Wang, K., Xian, Y., Lee, C., Wei, Y. e Huang, Z. (2019). On selecting directions for directional distance functions in a non-parametric framework: a review. *Annals Operations Research*, 278:43–76.