

ANÁLISE DE VARIÁVEIS AMBIENTAIS NA EFICIÊNCIA COMPARATIVA

Ricardo Antonio Lopes Maioli

Universidade Federal de Itajubá - UNIFEI
Av. BPS, 1303, bairro Pinheirinho, Itajubá - MG
maiolilopes@hotmail.com

Rafael Coradi Leme

Universidade Federal de Itajubá - UNIFEI
Av. BPS, 1303, bairro Pinheirinho, Itajubá - MG
leme@unifei.edu.br

RESUMO

Para analisar o desempenho das eficiências de Unidades Produtivas que tomam Decisões (DMU), analistas devem identificar as variáveis ambientais não gerenciáveis que mais impactam no desempenho de uma DMU. Um método muito conhecido para a análise de eficiência é a análise envoltória de dados (DEA, do inglês, Data Envelopment Analysis), onde o estudo das variáveis ambientais é geralmente realizado em dois estágios. No primeiro estágio encontram-se as eficiências das DMUs através da análise comparativa DEA, enquanto no segundo as variáveis ambientais são consideradas através do uso do método de regressão. Portanto, a seleção e a análise do impacto das variáveis ambientais no desempenho das DMUs são de extrema importância. Este trabalho final de graduação vem, portanto, estudar uma metodologia para identificar os melhores modelos de segundo estágio possíveis durante a análise comparativa das DMUs, utilizando para isso simulação e arranjos experimentais.

PALAVRAS CHAVE. DEA. DOE. 2-Stage DEA.

Área principal DEA. EST.

ABSTRACT

When analyzing efficiency performance, analysts must decide how to handle unmanageable environmental variables. Variables should be selected such that a model may be so constructed as to analyze the performance of decision-making units (DMU). In general, the efficiency score may be estimated by data envelopment analysis (DEA); the impact of environmental variables may be analyzed with the help of a regression model. This yields a two-stage DEA, a process that is gaining attention in the literature. However, little has been published on how to select a model. The motivation here is to propose a methodology that systematically analyzes the impact of some environmental variables on DMUs efficiency. This paper proposes an alternative approach, one based on the design of experiments philosophy. To gain insights into the proposed methodology, results are tested with the help of simulated models.

KEYWORDS. DEA. DOE. 2-Stage DEA.

Main area DEA. EST.

1. Introdução

Com a dificuldade em aumentar a produção ao longo dos anos, a avaliação de eficiência nas organizações passou a ter grande relevância Cook e Seiford (2009). De fato, Farrel (1957), abordou pela primeira vez a importância da medição da eficiência, tanto do ponto de vista teórico como do ponto de vista prático. A partir do trabalho de Farrel (1957), diversas propostas de análise de eficiência foram apresentadas na literatura, culminando no artigo de Charnes *et al* (1978), que apresentou uma abordagem de programação matemática, conhecido atualmente como Análise Envoltória de Dados (do inglês Data Envelopment Analysis - DEA), que tem como principal objetivo identificar através de uma análise não paramétrica, as empresas que apresentam as melhores práticas. Essas organizações ou atividades que geram produtos ou serviços são denominadas Unidades Tomadoras de Decisão (do inglês Decision Making Units – DMUs) Carvalho e Gomez (2009).

Para o cálculo das eficiências, a técnica DEA utiliza programação linear, e a região produção viável é determinada por uma região convexa definida pelas observações das DMUs. As empresas eficientes são observadas sobre a fronteira da isoquanta, enquanto as demais ficam dentro da região, sendo, desse modo, dominadas pelas mais eficientes. Dessa forma, a medida da eficiência de uma DMU é a distância dela até a isoquanta.

No entanto, existem diversas formas de se medir a eficiência como, por exemplo, a utilizada por Farrel (1957) que utiliza a medida radial. Essa medida também foi usada por Charnes *et al* (1978) para a construção do modelo inicial de DEA.

Em geral, a medida de eficiência é determinada pela distância radial observada pela DMU até a isoquanta, ou seja, a medida radial tem como objetivo medir a distância da posição da DMU até a fronteira de produção. Com isso, essa medida define quanto a DMU tem que reduzir (aumentar) em insumo (produção) para atingir a fronteira eficiente de produção, caso a DMU esteja sobre a fronteira de produção, sua eficiência é igual a 1. Em outras palavras, a medida radial define a distância das firmas ineficientes para as firmas eficientes.

Apesar de existirem diferentes variações do DEA, duas delas são consideradas principais: a primeira sendo a DEA orientada a insumos, onde a análise verifica a possibilidade da redução no consumo do insumo para a DMU ineficiente, e a segunda, DEA orientada a produtos, em que a análise verifica a possibilidade do aumento da quantidade de produtos, mantendo a quantidade de insumos constante. Os modelos de DEA também podem ser classificados em dois subtipos, o primeiro sendo o DEA de retorno de escala constante (do inglês Constant Return to Scale – CRS), proposto por Charnes *et al* (1978), em que o crescimento do insumo gera um crescimento constante do produto na mesma proporção, enquanto o segundo, que é conhecido como retorno de escala variável (do inglês Variable Return to Scale - VRS), em que a proporção do crescimento dos insumos e produtos podem ser crescentes ou decrescentes.

No entanto, quando a DEA é utilizado para avaliação de eficiência das empresas de uma forma prática, surgem novos fatores que, em geral, não podem ser controlados pelos gestores. Esses fatores são conhecidos como variáveis ambientais e podem afetar significativamente a eficiência das empresas. Exemplos práticos dessas variáveis podem ser facilmente citados, como catástrofes naturais, falta de energia ou até mesmo a falta de operadores em determinados turnos de uma fábrica.

Para analisar o impacto das variáveis ambientais na análise de eficiência através da DEA, autores como Ray (2004), Chillingirian e Sherman (2004) e Ruggiero (2004) propuseram um método denominado de Análise de Envoltória de Dados de dois estágios, também conhecido em inglês como 2-Stage DEA.

Nesta metodologia, o primeiro estágio consiste, em geral, do cálculo da eficiência comparativa através dos modelos tradicionais de DEA.

Já no segundo estágio, considera-se o impacto das variáveis ambientais na eficiência geral das empresas através de uma análise de regressão, permitindo assim a obtenção da uma eficiência calculada mais condizente com a realidade.

Contudo, a análise do segundo estágio através dos modelos de regressão tradicionais pode resultar em avaliações errôneas, uma vez que as variáveis de eficiência do modelo da

equação (1) apresentam apenas valores maiores do que 1. No entanto, para superar essa dificuldade, Simar e Wilson (2007) propuseram um método de geração de dados consistente para estimação de eficiências semi-paramétricas, viabilizando do uso de técnicas de reamostragem (*bootstrap*) para avaliação dos coeficientes da regressão.

Apesar da metodologia proposta por Simar & Wilson (2007) viabilizar a inferência sobre os coeficientes do modelo de regressão, a identificação das variáveis ambientais que realmente impactam a eficiência das DMUs é geralmente estudada através de análise qualitativa ou tentativa e erro. Dessa forma, motivado pelo propósito de sugerir uma metodologia que analise o impacto das variáveis ambientais na eficiência das empresas (DMU), esse trabalho propõe o uso de uma metodologia de análise das variáveis ambientais baseada em Planejamento de Experimentos (Design de experimentos - DOE).

O DOE foi criado a partir da idéia inicial de Fischer (1966), sendo que ele é atualmente utilizado nos mais diversos campos da ciência, tendo extrema importância para as áreas de gestão de operações, manufatura e controle da qualidade.

Apesar de ter sido inicialmente criado para atender a experimentos físicos, o DOE passou a ser utilizado em experimentos simulados devido ao grande avanço do *hardware* dos computadores. KLEIJNEN (2005). Segundo Montgomery (2009) os tipos de DOE mais comuns são os fatoriais, os de superfície de resposta e os de mistura.

Esses tipos de DOE são também conhecidos em inglês como *screening design*, pois eles visam mostrar o comportamento de uma função dependente de fatores experimentais, ou seja, realizar a pré-seleção das variáveis envolvidas no processo. Portanto, esses arranjos experimentais aparecem como uma boa alternativa para análise do impacto das variáveis ambientais na eficiência comparativa das DMUs. Por isso, este trabalho utilizará os DOE de *screening* fatorial para analisar o impacto das variáveis ambientais na eficiência das Unidades que tomam decisões.

Este artigo encontra-se estruturado em introdução, onde é apresentado o trabalho e suas motivações, objetivo, uma revisão bibliográfica, onde o estado da arte sobre DOE e DEA dois estágios é discutido, os resultados juntamente com sua análise, a conclusão e por fim as referências bibliográficas utilizadas pelo autor.

2. Objetivo

O presente trabalho teve como principal objetivo estudar e testar a aplicação de arranjos experimentais na avaliação do impacto das variáveis ambientais relacionadas em modelos de análise de eficiência e que podem ser avaliadas a partir da técnica de análise de envoltória de dados de dois estágios. Os resultados obtidos na metodologia estudada foram testados através do uso de experimentos simulados.

3. Fundamentação Teórica

3.1 Eficiência

As empresas sempre tiveram grande interesse na medição da eficiência, pois as mesmas encontram grandes problemas para aumentar sua produção. Cook e Seiford (2009)

As principais razões para esse interesse foram colocadas de uma maneira considerada clássica pelo artigo de Farrel (1957). Nele, Farrel (1957) cita que a medida da eficiência é importante tanto de ponto de vista prático quanto do ponto de vista teórico, pois, para a teoria sobre a eficiência de diferentes sistemas econômicos ser provada são necessárias medidas empíricas da eficiência. Já na parte prática, essas medidas ajudam, a saber, quanto determinada indústria pode crescer em produção somente com o aumento de sua eficiência, sem a necessidade da obtenção de mais recursos. Além disso, o autor ainda complementa o artigo observando que todas as medidas de eficiência anteriores não eram satisfatórias porque falhavam na hora de combinar processos com diferentes insumos em uma medida de eficiência única. Para resolver esse problema Farrel (1957) propôs uma análise de atividade que podia lidar melhor com os

problemas, podendo, segundo o autor, ser aplicada em qualquer organização. Simar & Wilson (2007)

No entanto, apesar de Farrel (1957) ter proposto um método mais robusto, ele confinou seus exemplos numéricos e sua discussão em situações onde havia somente um produto resultante do processo, apesar de ser possível formular casos em que múltiplos produtos eram gerados através dos insumos utilizados. Simar & Wilson (2007)

Cerca de 20 anos após o artigo de Farrel, Charnes et al (1978) propuseram um modelo baseado em programação matemática que permitia a análise de eficiência considerando retorno de escala constante, múltiplas entradas e múltiplas saídas. Essa metodologia foi chamada de DEA, em português, Análise Envoltória de Dados.

A idéia original dessa metodologia consiste em obter, com uma quantidade de empresas comparáveis, quais são as que possuem as melhores práticas de produção e que, por consequência, operem sobre uma fronteira de produção eficiente. Com isso, a metodologia permite a avaliar empresas que são consideradas ineficientes, pois estas se encontram fora da fronteira de produção. A técnica ainda permite comparar essas empresas com as empresas eficientes.

O modelo proposto por Charnes et al (1978) foi então estendido por outros pesquisadores, permitindo a análise de eficiência em diferentes hipóteses. O modelo básico utilizado para este trabalho é descrito na equação (1), e tem como característica ser um modelo de retorno de escala variável.

$$\max_{\delta_i, \lambda} \left\{ \delta_i \left| \begin{array}{l} Y \cdot \lambda \geq \delta_i y_i \\ X \cdot \lambda \geq x_i \\ \lambda \geq 0 \\ \lambda^T u = 1 \end{array} \right. \right\} \quad (1)$$

em que δ_i é a medida de eficiência para a i -ésima DMU; X é uma matriz $n \times m$ dos insumos; Y é uma matriz $n \times r$ dos produtos; λ é um vetor com os pesos do parâmetro; n , m e r são a quantidade de DMUs, a quantidade de insumos e a quantidade de produtos, respectivamente; x_i é o vetor de insumos para a DMU_i ; e y_i é o vetor de produtos para a DMU_i ;

A única diferença que separa o modelo de retornos de escala constante do modelo de retorno de escala variável é a última restrição da equação (1), que deve ser retirada caso o modelo seja CRS. Além disso, partir da equação (1), é possível verificar que se δ_i for igual a 1 para certa DMU, a quantidade produzida de produtos é a melhor possível. Porém, se δ_i for um valor menor do que 1, mais produtos poderiam ter sido produzidos com o insumo disponibilizado. É importante ressaltar, contudo, que conclusões similares podem ser obtidas através do modelo orientado a insumos.

3.2 Análise Envoltória de Dados de dois estágios

A DEA dois estágios surgiu da necessidade das empresas em observar a importância de fatores não gerenciáveis que impactam na eficiência da empresa. Também conhecidos como variáveis ambientais, esses fatores não fazem parte dos insumos e dos métodos de produção utilizados, mas sua influência na eficiência final pode não ser irrelevante. Portanto, somente o método tradicional de DEA não é capaz de resolver o problema dessas variáveis.

A fim de encontrar uma solução, autores como Ray (2004), Chilingirian e Sherman (2004) e Ruggiero (2004) propuseram dividir a análise de eficiência em dois estágios. No primeiro estágio, a eficiência é avaliada em relação aos produtos e insumos, utilizando a DEA. Já no segundo estágio, utiliza-se uma análise de regressão da eficiência obtida no primeiro estágio para avaliar o impacto das variáveis ambientais. Simar e Wilson (2007, 2010) listam uma série de publicações com caráter semelhante.

Contudo, a análise do segundo estágio através dos modelos de regressão tradicionais pode resultar em avaliações errôneas, uma vez que as variáveis de eficiência do modelo da equação (1) apresentam apenas valores maiores do que 1. No entanto, para superar essa dificuldade, Simar e Wilson (2007) propuseram um método de geração de dados consistente para

estimação de eficiências semi-paramétricas, viabilizando do uso de técnicas de reamostragem (*bootstrap*) para avaliação dos coeficientes da regressão.

Nesse modelo, se uma DMU sofrer com a influência de uma variável z_i , o modelo de regressão deve ser escrito conforme a equação (2).

$$\delta_i = \psi(z_i, \beta) + \varepsilon_i \geq 1 \quad (2)$$

em que δ é a eficiência técnica, β é vetor dos parâmetros, ψ é a função contínua, ε_i é o erro independente e identicamente normalmente distribuído;

É possível notar que na equação 2 os valores das eficiências estimados com o primeiro estágio são utilizados na regressão com as variáveis ambientais. Entretanto, como as estimações de eficiência do primeiro estágio têm grande sensibilidade e as do segundo estágio são altamente correlacionadas, métodos tradicionais para o cálculo do erro padrão dos coeficientes não são aconselhados. Por isso, devido as suas características, o *bootstrap* torna-se uma opção viável para a obtenção do erro padrão dos coeficientes da equação.

A técnica de reamostragem utilizada para resolver o problema de DEA de dois estágios é como segue:

- I. Calcular a eficiência pelo primeiro estágio dado pela equação (1);
- II. Utilizando o modelo de regressão truncada em 1.00 pela esquerda, estimar $\hat{\beta}$ de β e a variância do erro $\widehat{\sigma}_z^2$ de σ_z^2 na equação (2), considerando apenas as DMUs ineficientes.
- III. Repetir os passos (a) ate (c) L vezes para obter as estimativas reamostradas de β e σ_ε^2 ;
 - a. Para cada DMU ineficiente, gerar um número aleatório de ε_i^* de uma distribuição normal $N(0, \widehat{\sigma}_\varepsilon^2)$, truncada pela esquerda em $(1 - z * \hat{\beta})$;
 - b. Calcular as eficiências $\delta_i^* - z_i * \hat{\beta} + \varepsilon_i^*$;
 - c. Utilizando o modelo de regressão truncada em 1.00 pela esquerda, estimar $\hat{\beta}^*$ de β e a variância do erro σ_ε^{2*} de σ_ε^2 ;
- IV. Utilizar os valores reamostrados para corrigir o viés de $\hat{\beta}$ de β e da variância do erro $\widehat{\sigma}_\varepsilon^2$ e σ_ε^2 .

Uma vez estimado, o índice de eficiência de cada DMU pode ser ajustado através do impacto das variáveis ambientais. No entanto, antes de realizar o ajuste, uma análise e seleção do modelo a ser utilizado deve ser realizada de forma que as variáveis escolhidas devam, de fato, impactar no desempenho das DMUs. Um método bastante utilizado para esse fim é um índice consistente de informação, como o Critério Bayesiano de Informação (BIC). O BIC foi proposto originalmente por Schwarz (1978) e tem como principal objetivo ajudar na seleção de modelos matemáticos. Segundo Filho, Achcar e Souza (2010) o índice BIC foi modificado por Carlin e Louis (2000) de modo a considerar a densidade a posteriori para os parâmetros do modelo estudado.

O método para se calcular o índice BIC pode ser visto na equação (3).

$$BIC = E[\ln L(\theta)] - \frac{1}{2} \rho * \ln(N) \quad (3)$$

em que $E[\ln L(\theta)]$ é o valor esperado com relação à densidade a posteriori, ρ é a dimensão do vetor dos parâmetros e N é o tamanho amostral

3.3 Design of Experiments (DOE)

Um experimento pode ser definido como um teste ou uma série de testes que propõe mudanças nas variáveis de entrada para que possam ser observadas mudanças nas variáveis de saída. Por isso, o objetivo primário de um experimento é de desenvolver um processo que seja afetado o mínimo possível por fatores externos. Quanto mais fatores são envolvidos no experimento mais complicado ele tende a ficar, por isso, vários experimentos utilizam metodologias para serem realizados. O DOE estatístico é uma dessas metodologias e tem como principal característica a tentativa de coletar as informações necessárias com o menor número de experimentos possível, utilizando, para isso, conceitos estatísticos. Montgomery (2009)

A filosofia do DOE, apesar de ser largamente utilizada no campo industrial, não tem o mesmo impacto no campo da simulação. Segundo Kleijnen (2005), os analistas do campo de simulação, não convencidos pela técnica, costumam não procurar sobre o assunto, que geralmente se encontra em periódicos especializados. Além disso, a maioria dos arranjos experimentais foi desenvolvida para o mundo real ao invés da simulação. Geralmente, em simulação, as entradas e os parâmetros têm como objetivo testar os efeitos dos fatores do DOE, isso porque o DOE foi desenvolvido para capturar os efeitos desses fatores nas variáveis de resposta.

É importante notar, novamente, que o objetivo deste trabalho é usar experimentos simulados para analisar o impacto das variáveis ambientais na eficiência das DMUs. Nesse caso, a presença de cada variável deve ser analisada e, caso ela esteja presente, o valor da avaliação é (+1), caso contrário, é (-1), o que leva a um arranjo experimental de dois níveis.

O arranjo experimental fatorial completo é um tipo de arranjo que considera todas as combinações possíveis do experimento e gera estimadores independentes de todos os efeitos, porém, seu defeito reside no fato que o número de experimentos cresce exponencialmente, ou seja, 2^n , o que torna o modelo impraticável quando são utilizados vários fatores ambientais. A Tabela 1 apresenta um arranjo fatorial considerando, genericamente que F variáveis ambientais podem afetar a eficiência das DMUs.

Tabela 1 – Arranjo fatorial completo para F variáveis ambientais

r	z_1	z_2	z_3	...	z_i	BIC
1	-1	-1	-1			
F=1	2	1	-1			
	3	-1	1			
F=2	4	1	1			
	5	-1	-1	1		
	6	1	-1	1		
	7	-1	1	1		
F=3	8	1	1	1		
	
F	2^F					

Portanto, analisando os experimentos simulados, um metamodelo considerando todos os fatores principais bem como suas interações, pode ser escrito como mostrado na equação (4) para estimar o impacto de cada variável ambiental na especificação do modelo através do BIC.

$$BIC_r = \beta_0 + \sum_{i=1}^F \beta_i z_i + \sum_{i,j=1,i < j}^F \beta_{i,j} z_i z_j + \dots + \beta_{1,2,\dots,F} z_1 z_2 \dots z_F + \varepsilon \quad (4)$$

Na equação (4), β_0 , β_i , $\beta_{i,j}$, ..., $\beta_{1,2,\dots,F}$ são conhecidos como coeficientes dos efeitos (principais e interações) e z_i é a i-ésima variável ambiental observada, como na equação (2). Finalmente, r é a r-ésima corrida do arranjo experimental. Note que o arranjo experimental

da Tabela 1 é ortogonal, o que produz estimações não correlacionadas dos coeficientes da equação (4).

O metamodelo da equação (4) usualmente é obtido pelo método de mínimos quadrados, considerando o erro ε independente, identicamente e normalmente distribuídos. Uma vez estimados, testes de hipótese podem ser utilizados para avaliar o impacto de cada variável ambiental na especificação do modelo. O objetivo é verificar se a diferença na resposta média é estatisticamente significativa entre os dois níveis de cada fator, assim como nas interações entre eles

4. Resultados e análise

Para testar o método proposto foram realizados experimentos com um modelo composto de três variáveis ambientais z_i , sendo que das três variáveis presentes nesse modelo apenas duas eram significativas e afetavam a eficiência δ das DMU's. Tal modelo pode ser visto na equação (5).

$$\delta_i = 0,5 + 0,5z_1 - 0,5z_2 + 0,0z_3 \geq 1 \quad (5)$$

Nesse modelo, tais variáveis ambientais não correlacionadas seguiam distribuições normais, sendo que $z_1 \sim N(1,2)$, $z_2 \sim N(2,2)$ e $z_3 \sim N(0,5,2)$. Além disso, os dados para análise foram gerados conforme mostrado em Simar & Wilson (2007).

Visando validar o desempenho do DOE na identificação do modelo, foi utilizado o delineamento de experimentos conhecido como fatorial completo, com oito experimentos, conforme mostrado na Tabela 1. Além disso, esse arranjo experimental também foi testado em outra situação com duas réplicas.

Para verificar se tais variáveis estavam presentes ou não no modelo testado foi utilizado o índice BIC, onde o a metodologia de seleção é quanto menor melhor, ou seja, modelo com o menor BIC é o que tende a possuir melhor ajuste.

Os dois casos foram analisados da mesma forma, tanto graficamente como quantitativamente, visando assim facilitar a compreensão dos resultados obtidos.

4.1 Caso com uma réplica

Para o primeiro caso, contendo apenas uma réplica, após realizar o experimento, o resultado obtido graficamente pode ser observado nos Figuras 1 e 2:

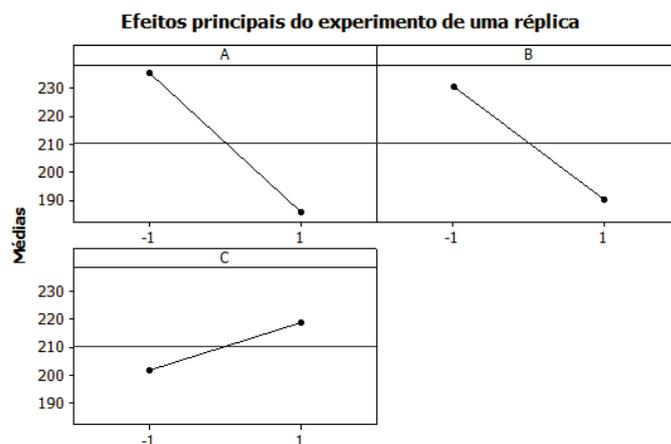


Figura 1 – Significância dos efeitos principais

Ao analisar esse gráfico, foi possível observar que as duas variáveis ambientais que afetam a eficiência são a A e a B, uma vez que a presença delas melhora a especificação do BIC. Já a variável C deve ser retirada do modelo, pois o BIC é melhor especificado com a sua

ausência. Também foi analisada a eficácia do método em transmitir se as variáveis são correlacionadas ou não, sendo que essa análise pode ser vista na Figura 2, onde é possível observar que as variáveis ambientais possuíam pouca ou nenhuma correlação entre elas.

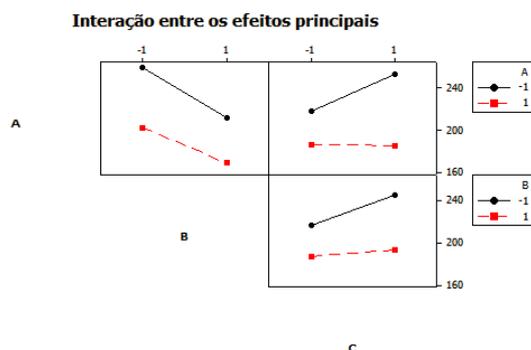


Figura 2 – Interação entre as variáveis ambientais

Porém, apesar de o método gráfico ser altamente ilustrativo, ele nem sempre permite analisar corretamente os resultados obtidos. Portanto, a utilização dos métodos quantitativos é recomendada para avaliar a robustez dos resultados. Como forma de comparação, os resultados obtidos em ambos os gráficos foram comparados com os dados obtidos de forma quantitativa. Esses resultados podem ser vistos na Tabela 2.

Tabela 2 – Valores dos efeitos e seus respectivos coeficientes para o experimento de uma réplica

Termo	Efeito	Coefficiente
Constante	*	210,35
A	-50	-25,00
B ²	-40,55	-20,27
C	17,31	8,65
A*B	6,74	3,37
A*C	-17,57	-8,79
B*C	-11,64	-5,82
A*B*C	15,62	7,81

Nota-se que, especificamente para este caso, não há a possibilidade de se realizar testes de hipótese sobre os coeficientes estimados, pois o experimento foi realizado com apenas uma réplica e analisando todas as interações.

Mesmo sem os testes de hipótese disponíveis, é possível observar que tanto as interações A*B*C, A*B e C não deveriam estar no modelo devido aos seus valores positivos, que deterioram o valor do BIC, comprometendo o modelo. Com isso, uma nova análise é realizada sem considerar as interações A*B*C, A*B.

Após a retirada dessas interações, pode-se, então testar hipóteses sobre os efeitos testados no experimento simulado. Além disso, o gráfico dos resíduos do modelo também foi observado, visando assim validar a hipótese da normalidade dos resíduos.

O valor dos efeitos das variáveis e suas interações, além do valor do teste de hipótese sobre os coeficientes são mostrados na tabela 3, enquanto os resíduos são mostrados na Figura 3.

Tabela 3 - Valores dos efeitos e seus respectivos coeficientes para o experimento de uma réplica

Termo	Efeito	Coefficiente	P-value
Constante	*	210,35	0,001
A	-50,00	-25,00	0,053
B	-40,00	-20,27	0,078
C	17,31	8,65	0,287
A*B	-17,57	-8,79	0,281
A*C	-11,54	-5,82	0,435

Pode-se concluir da Tabela 3 que apesar dos valores A*B e A*C apresentarem valores negativos em seus efeitos, o que ajudaria o BIC, eles são não estatisticamente significantes, assim como o valor de C.

No caso da variável ambiental B, apesar de ela ser considerada estatisticamente não significativa, seu valor é próximo de 0,05, o que a tornaria significativa, por isso, os termos escolhidos para serem retirados do modelo nesse caso foram A*B e A*C, eliminando assim as últimas duas interações presentes no mesmo.

Com relação aos resíduos do modelo apresentado, a Figura 3 mostra que eles estão agrupados, porém através do teste de Anderson-Darling para normalidade, não há evidência para rejeitar a hipótese de que os resíduos são normalmente distribuídos. Portanto, pode-se concluir que as variáveis ambientais que realmente impactam os índices de eficiência são as variáveis z_1 e z_2 .

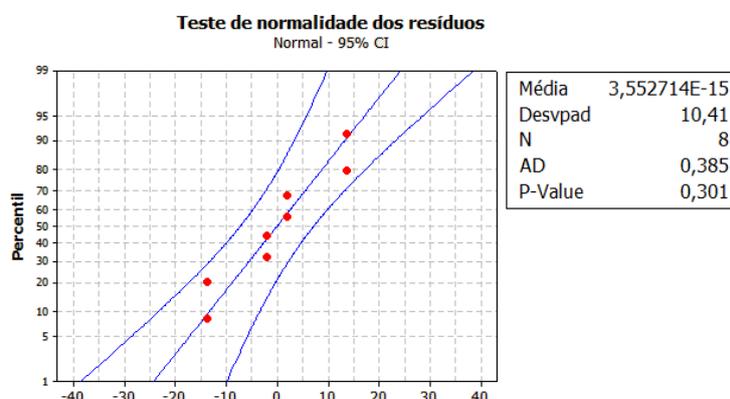


Figura 3 – Teste de normalidade dos resíduos para o experimento de uma réplica

Pode-se, ainda, considerar a análise apenas com os efeitos principais, conforme mostrado na Tabela 4. Esta análise permite concluir com facilidade que o termo C deve ser ignorado no modelo, tanto do ponto de vista de significância como do ponto de vista de efeito no BIC, pois ele possui valor positivo.

A Figura 4 demonstra que a normalidade dos resíduos, ao ser analisada e comparada com a Figura 3, mostrou uma melhora na adequação dos dados no quesito de normalidade com o novo modelo testado.

Tabela 4 - Valores dos efeitos e seus respectivos coeficientes para o experimento de uma réplica

Termo	Efeito	Coefficiente	P-value
Constante	*	210,35	0,00
A	-50,00	-25,00	0,021
B	-40,55	-20,27	0,04
C	17,31	8,65	0,27

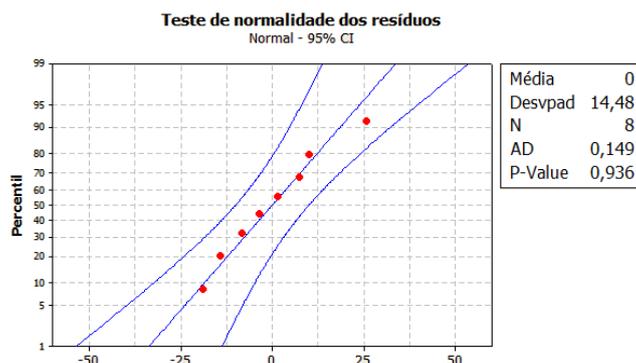


Figura 4 - Teste de normalidade dos resíduos para o experimento de uma réplica

4.2 Caso com duas réplicas

Para o segundo caso, contendo duas réplicas, após realizar o experimento, o resultado obtido graficamente pode ser observado nas Figuras 6 e 7.

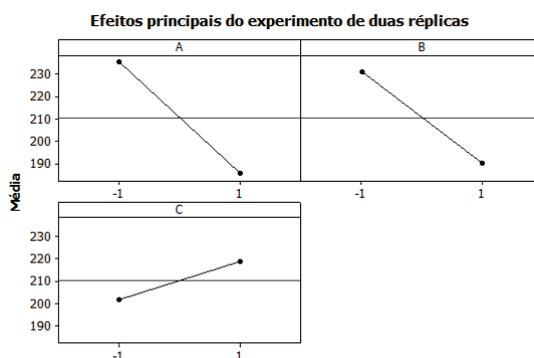


Figura 6 - Significância dos efeitos principais

Ao analisar a Figura 7 foi possível perceber que os termos principais que deveriam estar no modelo eram A e B, pois C, quando adicionado ao modelo gerava um valor maior de BIC. Já a Figura 7 permite mostrar que os fatores principais praticamente não possuem interação, ou seja, são não correlacionados. Porém, ainda segundo a figura 7, pode haver uma interação entre o termo A e o termo C.

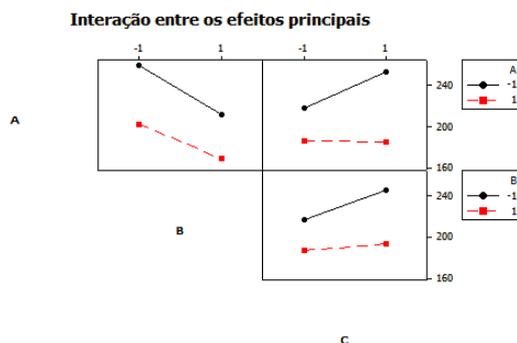


Figura 7 - Interação entre as variáveis ambientais

Da mesma forma que foi analisada o caso anterior, o experimento de duas réplicas baseou-se principalmente na parte quantitativa, pois essa gera informações mais esclarecedoras do que a ferramenta gráfica. Tal análise foi feita a partir dos dados da Tabela 5.

Tabela 5 - Valores dos efeitos e seus respectivos coeficientes para o experimento de duas réplicas

Termo	Efeito	Coefficiente	P-value
Constante	*	210,41	0,000
A	-49,95	-24,97	0,000
B	-40,79	-20,40	0,000
C	17,28	8,64	0,000
A*B	7,03	3,51	0,000
A*C	-17,94	-8,97	0,000
B*C	-11,42	-5,71	0,000
A*B*C	15,86	7,93	0,000

Com mais de uma réplica, é possível aplicar os testes de hipótese diretamente nos resultados, ao contrário do observado no caso anterior. Outro fato interessante da Tabela 5 é que todos os valores eram significantes, porém, apesar disso os termos A*B*C, A*B e C pioram a especificação do BIC do modelo. No entanto, nesta análise, a interação A*C e B*C parecem testar significantes para melhora da especificação do BIC, o que não condiz com o modelo da equação (5). É possível observar, ainda, que os termos os resíduos puderam ser considerados normalmente distribuídos, como mostra a Figura 8.

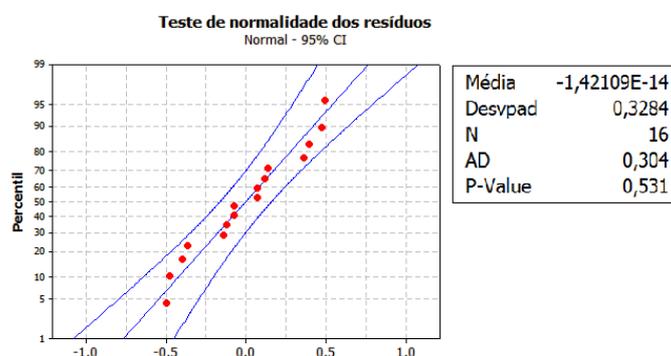


Figura 8 - Teste de normalidade dos resíduos para o experimento de duas réplicas

No entanto, observando o gráfico dos valores ajustados vs. resíduos apresentado na Figura 9, é possível observar que a variância do resíduo não é constante, conforme suposto no método de mínimos quadrados. Esse comportamento heteroscedástico do resíduo pode afetar a análise estatística do modelo. Portanto, o método de mínimos quadrados não é adequado para este tipo de análise, sendo sugerido um modelo que considere a heteroscedasticidade do resíduo.

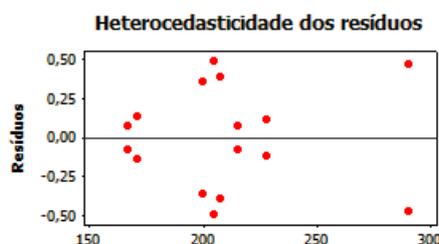


Figura 9 – Heteroscedasticidade dos resíduos

5. CONCLUSÃO

Neste artigo foi descrito um estudo simulado para determinação do modelo semi-paramétrico de dois estágios, baseado na Análise de Envoltória de Dados, para análise do impacto de variáveis ambientais. Essas variáveis, conhecidas como variáveis não controláveis,

geralmente não são consideradas em modelos de DEA tradicionais, que usam somente os insumos e produção como fatores de cálculo. No entanto, a análise das variáveis ambientais pode ser de extrema importância para gestores, visto que sua influência na eficiência final das empresas pode ser significativa e alterar consideravelmente o rendimento final.

Para determinar a importância dessas variáveis, a metodologia propôs o uso de Planejamento de Experimentos para testar os efeitos desses fatores dentro de um modelo DEA de dois estágios e, desse modo, ajudar os gestores na correta tomada de decisões. Portanto, arranjos experimentais ortogonais foram utilizados em experimentos simulados visando estudar a eficácia do método.

Os testes realizados consideraram variáveis ambientais não correlacionadas. Além disso, os resultados sugerem que o método estudado pode ser útil para determinar se as variáveis ambientais são significativas ou não no desempenho das DMUs para os casos estudados, tanto com o experimento de uma réplica, como no experimento de duas réplicas. Contudo, os estudos realizados mostraram que modelos que consideram resíduos homocedásticos podem não ser suficientes para avaliar de forma precisa o impacto das variáveis ambientais na eficiência, sendo necessários estudos com modelos de regressão que possam considerar resíduos heteroscedásticos.

6. REFERÊNCIAS BIBLIOGRÁFICAS

- Carlin, B. P. e Louis, T. A.** (2000), Bayes and Empirical Bayes methods for data analysis, *Chapman and Hall/CRC Press*.
- Carvalho, C. M. de C e Gomez, A. P.,** *Introdução à Análise Envoltória de Dados*. Editora UFV, Viçosa MG, 2009.
- Charnes, A., Cooper, W.W., e E Rhodes E.** (1978) Measuring the efficiency of decision making units. *European Journal of Operational Research*, 2, 429–44.
- Chilingerian, J.A. e Sherman, H.D.,** Health care applications: From hospitals to physicians from productive efficiency to quality frontiers, In W. W. Cooper, L.M. Seiford, J. Zhu (Eds.), *Handbook on Data Envelopment Analysis*. Kluwer, Boston, 2004
- Chung, C. A.** *Simulation modeling handbook: a practical approach*, CRC Press, Florida, 2004.
- Cook, W.D.; Seiford, L.M.** (2009), Data envelopment analysis (DEA): Thirty years on. *European Journal of Operational Research*, 192, 1–17.
- Farrel, M.J.** (1957), The measurement of productive efficiency, *Journal of the Royal Statistical Society*, Series A, 120, 253–281.
- Filho, M. B., Achcar, J. A. e Souza, R. M.** (2010), Modelos de volatilidade estocástica em series financeiras: uma aplicação para o ibovespa. *Economia Aplicada*, 14, n. 1, 25-40.
- Fisher, R.A.** *The Design of Experiments*, Hafner Publishing Company, New York, 1966.
- Kleijnen, J.P.C.** (2005), An overview of the design and analysis of simulation experiments for sensitivity analysis. *European Journal of Operational Research*, 164, 287-300.
- Montgomery, D. C.** *Design and Analysis of Experiments*, John Wiley & Sons, Inc, New York, 2009.
- Plackett, R. L e Burman, J.P.** (1946), The Design of Optimum Multifactorial Experiments. *Biometrika*, 33, 305-25.
- Ray, S.C.** *Data Envelopment Analysis: Theory and Techniques for Economics and Operations Research*. Cambridge University Press, Cambridge, 2004.
- Ruggiero, J.** Performance evaluation in education: Modeling educational production. In W. W. Cooper, L.M. Seiford, J. Zhu (Eds.), *Handbook on Data Envelopment Analysis*, Kluwer, Boston, 265-298, 2004.
- Schwarz, G.** (1978), Estimating the dimension of a model. *The annals of statistics*, 6, 461-464.
- Simar, L. e Wilson, P.W.** (2007), Estimation and inference in two-stage, semi-parametric models of production processes. *Journal of Econometrics*, 136, 31-64.
- Simar, L. e Wilson, P.W.** (2010), Two-stage DEA: Caveat emptor. *Journal of Productivity Analysis*, 36, 205-218.