

USO DE TÉCNICAS DE SELEÇÃO DE VARIÁVEIS EM *DEA* PARA ANALISAR O SETOR ELÉTRICO

Luis Felipe Aragão de Castro Senra
Programa de Engenharia de Produção
Universidade Federal Fluminense

João Carlos Correia Baptista Soares de Mello
Departamento de Engenharia de Produção
Universidade Federal Fluminense

Resumo: Uma fragilidade clássica dos modelos *DEA* é a sua fraca capacidade de ordenação. Nos últimos anos, vários métodos foram usados para resolver a questão. Um deles foi o Método Multicritério para Seleção de Variáveis em Modelos *DEA*, para o qual são propostas nesse artigo, duas alternativas para a escolha do par inicial de *Input-Output* desse método.

Com o intuito de ilustrar e demonstrar a aplicação destas duas alternativas, é analisado o caso real de avaliação da eficiência de algumas empresas distribuidoras de energia que estão no SIAGE, programa desenvolvido e coordenado pela Eletrobrás.

Palavras-Chaves: *DEA*; Distribuição de Energia; Eficiência; Seleção de Variáveis.

1. INTRODUÇÃO

A *Data Envelopment Analysis* – *DEA*, ou seja, Análise Envoltória de Dados é um método não paramétrico para calcular a eficiência comparada de diversas alternativas, *DMUs*, ou *Decision Making Units*. Proposto originalmente em 1978 por Edward Rhodes sob a supervisão de W.W. Cooper, tinha como objetivo desenvolver um método para comparar a eficiência de escolas públicas.

Ano após ano, esse método vem recebendo contribuições e sendo desenvolvido por diversos cientistas por todo o globo. Sua grande aceitação surge em razão da sua principal característica, objetividade. Como é um método que em sua essência não precisa de um decisor, ele não é afetado pela sua subjetividade, disponibilizando resultados “puros”.

A *DEA* pode ser resumida como um método que levando em consideração diversas alternativas, *DMUs*, diversos insumos, *Inputs*, e diversos produtos, *Outputs*, é capaz de através de programação linear, definir qual a eficiência de cada alternativa em relação às outras.

Portanto, a *DEA* é utilizada basicamente em três problemáticas, a saber: ordenação, alocação e escolha. Na primeira, as alternativas são ordenadas de acordo com a sua eficiência; na segunda, elas podem ser separadas em eficientes e ineficientes; a terceira é usada para selecionar uma *DMU*.

Historicamente uma fragilidade clássica da *DEA* tem sido a sua baixa capacidade de ordenar as *DMUs*, dado que, ao descrever-se um modelo com diversas variáveis em relação ao número de *DMUs*, muitas *DMUs* são classificadas como eficientes, colocando em xeque a sua capacidade de ordenação.

Com o passar dos anos vários têm sido os métodos que, de alguma forma, têm tentado superar essa fragilidade. Alguns deles consideram como norte limitar o número de variáveis, *Inputs* e *Outputs*, como uma estratégia de ao mesmo tempo, manter uma relação causal entre o modelo e a realidade e atingir uma boa ordenação. Dado que todo modelo é uma tentativa de descrever simplificada uma realidade, o seu maior desafio tem sido definir quais são as variáveis que devem ser selecionadas e quais são

dispensáveis, para que o modelo continue descrevendo de forma fidedigna essa realidade.

Dentre esses, um que se destaca é o Método Multicritério para Seleção de Variáveis em Modelos *DEA* (Soares de Mello et al., 2002), que como o próprio nome delata, utiliza um método multicritério para decidir quais são as variáveis “descartáveis”. Entretanto, tal método delega sua gene a um decisor, “poluindo” assim, o modelo com subjetividade.

Tentando minimizar os efeitos da escolha subjetiva do decisor nos resultados obtidos e manter uma boa ordenação, são propostas nesse artigo duas alternativas possíveis de tornar desnecessária a escolha por parte do decisor de como o processo iniciará.

Na tentativa de ilustrar o desenvolvimento teórico, foi feita uma análise utilizando modelo *DEA*, para avaliar um cenário da realidade brasileira, qual seja, o de avaliação de eficiência de empresas distribuidoras de energia elétrica.

2. FONTE DE DADOS

Todos os dados aqui expostos foram retirados do Sistema Integrado de Acompanhamento e Gestão Empresarial – SIAGE Esse sistema foi criado em cooperação com a COPPE-UFRJ e é mantido e atualizado pelo Departamento de Acompanhamento Empresarial da Eletrobras. É o sistema pelo qual todas as empresas federalizadas são monitoradas e serve de suporte no acompanhamento de gestão realizado pela Eletrobras nessas empresas. Foram utilizados os dados correspondentes aos meses de julho de 2002 e julho de 2003. Foram selecionadas seis empresas, a saber: Eletro Acre – EA, Manaus Energia – MESA, Companhia Elétrica do Piauí – CEPISA, Boa Vista Energia - BOVESA, Companhia Elétrica do Amazonas - CEAM e Companhia Elétrica de Rondônia – CERON.

Portanto, o cenário analisado possui doze *DMUs*, duas para cada empresa, representando um ano cada uma.

O cenário proposto analisa as empresas em relação ao seu papel social e não financeiro, ou seja, o objetivo das empresas não é obter o maior lucro, mas sim, distribuir a maior quantidade possível de energia para o maior número de habitantes possível, com a menor estrutura possível. Portanto, foram escolhidos como *Outputs* Demanda Atendida e Número de Consumidores Atendidos. Já em relação aos *Inputs* foram considerados as Despesas Operacionais mensais e o Ativo Permanente. Os dados utilizados encontram-se apresentados na Tabela 1.

Muitas outras variáveis poderiam ter sido utilizadas, como por exemplo, número de postes, número de transformadores na distribuição, tamanho da linha de distribuição entre outras. Entretanto, estes dados não estavam disponíveis.

Tabela 1 - Dados para o estudo de caso.

Base de Dados (SIAGE/Eletrobrás)				
	Output		Input	
	No de Consumidores (habitantes)	Demanda (Mw)	Despesas Operacionais (R\$)	Ativo Permanente (R\$)
EA02	118.573	31.532	7.098.308,05	88.799.688,79
EA03	122.028	32.232	10.062.040,40	94.215.610,80
MESA02	349.277	233.002	79.745.406,03	1.734.371.517,77
MESA03	368.105	240.308	110.516.903,28	1.712.296.489,36
CEPISA02	614.549	112.114	18.107.300,61	246.559.334,63
CEPISA03	639.936	107.712	22.960.740,84	237.553.008,05
BOVESA02	52.444	23.825	4.553.490,20	83.710.212,02

BOVESA03	55.143	27.977	5.974.264,79	84.344.946,64
CEAM02	147.017	27.374	19.883.069,34	209.061.961,10
CEAM03	157.660	31.910	30.425.464,35	258.832.662,12
CERON02	312.649	100.913	24.582.694,88	231.120.149,78
CERON03	331.079	106.813	32.009.565,62	237.329.367,09

O modelo *DEA* utilizado foi o modelo *Variable Returns to Scale* – VRS (Retornos Variáveis de Escala), dado que as escalas de valores da base de dados são muito dispares.

A análise foi orientada a *Output*, pois no setor elétrico analisado os investimentos são feitos em grande volume e sempre a tempo pretérito ao da demanda, objetivando disponibilizar atendimento a populações abastardas. Portanto, as empresas têm como meta atender novos públicos, ou seja, aumentar os *Outputs*, e não diminuir sua capacidade, ou seja, diminuir *Inputs*.

3. MÉTODO PRIMEVO DE ANÁLISE

A Tabela 2 apresenta os resultados do modelo *DEA* VRS, obtidos com o *software Frontier Analyst*. Nota-se que, como antes destacado, muitas *DMUs* (sete em doze) foram consideradas eficientes. Isto demonstra a fraca ordenação do método quando da existência de muitas variáveis em relação ao número de *DMUs*.

Tabela 2 - Análise Primária

DMUs	AP / DO / CO /DE
EA02	100,00
EA03	93,96
MESA02	100,00
MESA03	100,00
CEPISA02	100,00
CEPISA03	100,00
BOVESA02	100,00
BOVESA03	100,00
CEAM02	29,40
CEAM03	28,19
CERON02	96,64
CERON03	99,27

Para contornar este problema podem ser consideradas duas formas de escolha de variáveis em modelos *DEA*.

Na primeira abordagem, usada por Lins e Moreira (1999), parte-se de um par *Input-Output* inicial que tenha alta correlação. O critério de seleção de novas variáveis a serem inseridas é verificar qual das variáveis candidatas provoca maior eficiência média no modelo. O método pára ao atingir o número de variáveis considerado ideal.

Em uma segunda abordagem possível, o critério de seleção de novas variáveis a serem inseridas é provocar o menor número de *DMUs* eficientes, já que dificilmente ocorrem empates do valor da eficiência em *DMUs* ineficientes. Tal como na abordagem anterior, o método pára quando for atingido o número de variáveis considerado ideal.

Com o duplo objetivo de ao mesmo tempo atingir uma boa ordenação e de manter uma boa relação causal, o Método Multicritério para Seleção de Variáveis em Modelos *DEA* (Soares de Mello et al., 2002) tem como critérios o melhor ajuste à fronteira e a máxima discriminação.

O ajuste à fronteira é medido através da eficiência média e sua normalização cria a variável SEF, descritor de impacto do critério “melhor ajuste à fronteira”, que atinge o valor 10 na eficiência máxima e zero na eficiência mínima. A máxima discriminação é medida pelo número de *DMUs* eficientes, que se deseja minimizar. A normalização desta medida gera a variável SDIS, descritor de impacto do critério “máxima discriminação”, que atinge valor 10 para o menor número de *DMUs* na fronteira e zero para o maior número.

Para selecionar a variável que produz a melhor solução de compromisso e que deve entrar no modelo, agregam-se os dois critérios em um critério síntese, S, através da soma ponderada dos respectivos descritores de impacto, SEF e SDIS.

A variável que apresentar maior valor de S é incorporada ao modelo e, se o número de variáveis incorporadas não exceder 1/5 das *DMUs*, repete-se o processo para incorporar nova variável. Caso contrário, o procedimento está completo, ou seja todas as variáveis que devem ser incorporadas ao modelo estão selecionadas.

1. Escolher um par *Input-Output* inicial

Nessa etapa o decisor escolhe qual é o melhor par de *Input-Output* para iniciar a análise, tendo como base apenas o seu conhecimento sobre o assunto.

2. Calcular a eficiência média para cada variável acrescentada

Utilizando como base o par inicial é calculada a eficiência média de cada trio existente, ou seja, o par inicial mais cada uma das outras variáveis.

3. Calcular o número de *DMUs* na fronteira de eficiência

É contado o número de *DMUs* 100% eficientes em cada uma das alternativas calculadas na etapa 2.

4. Normalizar as escalas, calculando SEF e SDIS

SEF é a normalização da eficiência média de cada alternativa da etapa 2, é atribuído o valor de 10 a alternativa com maior eficiência média e 0 para a com menor eficiência média, as demais alternativas recebem notas entre 0 e 10 proporcionalmente as suas eficiências médias. Já o SDIS é a normalização da etapa 3, a alternativa com menos *DMUs* 100% eficientes recebe 10, a com mais *DMUs* eficientes recebe 0 e as demais como no sistema anterior.

5. Fazer a soma ponderada de SEF e SDIS

É feita a média dos valores de SEF e SDIS, de tal forma que cada alternativa tenha agora um valor S. Essa média, a princípio, deve ser aritmética; entretanto, fica disponível para o decisor a possibilidade de se usar pesos diferenciados para SEF e SDIS.

6. Escolher a variável que tenha maior valor de S

A alternativa que tiver o maior valor de S é considerada a alternativa que melhor concilia uma boa ordenação (alto valor de SDIS) e uma boa relação causal (alto valor de SEF).

7. Verificar se o número de *DMUs* excede o quántuplo¹ do número de variáveis

Se sim, deve-se reiniciar o processo na etapa 2 caso contrário, deve-se encerrar a escolha de variáveis.

4. PROPOSTAS

São propostas aqui duas alternativas para retirar o poder de escolha do decisor, tornando o Método Multicritério para Seleção de Variáveis em Modelos *DEA* ainda mais objetivo e imune a más interpretações (culposas ou dolosas) do decisor. Ambas propõem uma forma lógico-matemática de se definir na etapa 1 o par inicial que deve ser escolhido como origem do método.

Ao invés de delegar-se ao decisor o poder de escolher, propõe-se que seja feita uma média da contribuição *Input / Output* de cada variável no cenário que todas elas são utilizadas (vide Tabela 3).

Tabela 3 - Análise de Contribuição

Contribuição <i>Input/Output</i>				
	No de Consumidores	Demanda	Despesas Operacionais	Ativo Permanente
EA02	100,00	0,00	0,00	100,00
EA03	0,00	100,00	0,00	100,00
MESA02	0,00	100,00	100,00	0,00
MESA03	43,00	57,00	50,00	50,00
CEPISA02	0,00	100,00	100,00	0,00
CEPISA03	100,00	0,00	0,00	100,00
BOVESA02	0,00	100,00	0,00	100,00
BOVESA03	0,00	100,00	0,00	100,00
CEAM02	0,00	100,00	1,00	99,00
CEAM03	0,00	100,00	0,00	100,00
CERON02	0,00	100,00	0,00	100,00
CERON03	0,00	100,00	0,00	100,00
Media	20,25	79,75	20,92	79,08

Proposta 1 – Devem ser utilizados como par inicial o *Input* com maior média e o *Output* com maior média.

Tabela 4 - Análise Proposta 1

1.1.1. Par Inicial Demanda/Ativo Permanente		
<i>DMUs</i>	No Consumidores	Despesas Operacionais
EA02	100,00	100,00
EA03	93,96	93,96
MESA02	96,96	100,00
MESA03	100,00	100,00
CEPISA02	100,00	100,00
CEPISA03	100,00	100,00
BOVESA02	100,00	100,00
BOVESA03	100,00	100,00
CEAM02	29,40	29,40

¹ No método publicado originalmente, sugeria-se o triplo, entretanto, a mais recente tem-se sugerido o uso do quántuplo.

CEAM03	28,19	28,19
CERON02	96,64	96,64
CERON03	99,27	99,27
Eff Media	87,04	87,29
Eficientes	6	7
S	6	4

Note-se que a média da normalização S é igual, ou seja, 5 em cada caso, entretanto o primeiro trio foi o selecionado, pois foi atribuído a SDIS um maior peso, posto que o objetivo do estudo de caso é ordenar as *DMUs*.

Proposta 2 – Devem ser utilizados como par inicial o *Input* com menor média e o *Output* com menor média.

Tabela 5 - Análise Proposta 2

Par Inicial No Consumidores/Despesas Operacionais		
DMUs	Demanda	Ativo Permanente
EA02	78,05	100,00
EA03	53,98	88,71
MESA02	100,00	54,58
MESA03	100,00	57,52
CEPISA02	100,00	100,00
CEPISA03	100,00	100,00
BOVESA02	100,00	100,00
BOVESA03	84,57	90,86
CEAM02	24,01	27,24
CEAM03	26,01	24,64
CERON02	80,85	50,64
CERON03	76,63	51,80
Eff Media	77,01	70,50
Eficientes	5	4
S	4	6

Cabe aqui o mesmo comentário da proposta anterior.

5. CONCLUSÕES

A alteração do método mostrou-se muito boa, pois minimiza nos resultados obtidos a subjetividade do decisor, que pode vir de forma intencional ou não a influir no melhor resultado possível.

As variáveis escolhidas mostraram-se muito boas, pois o resultado final coincidiu com a opinião dos técnicos responsáveis pelo SIAGE.

Quanto às alternativas propostas, conclui-se que a escolha do *Input* e *Output* com menores médias leva a uma melhor ordenação das *DMUs*.

Entretanto, essa é apenas uma das muitas possibilidades existentes para melhorar a ordenação em modelos *DEA*. Recomenda-se, portanto, outras análises que venham a confirmar a coerência dessa alternativa.

6. AGRADECIMENTOS

Este trabalho teve o apoio do Fundo Setorial de Energia (CT-Energ), por intermédio do CNPq (CT-Energ/CNPq 01/2003), processo 400646/2003-0.

7. BIBLIOGRAFIA

- [1] ELETROBRAS – DPAC, Empresas Federais de Distribuição de Energia Elétrica – SIAGE, Setembro/2003, DPAC, Rio de Janeiro, 2003.
- [2] LINS, M. P. E. e MEZA, L. A., Análise Envoltória de Dados: e perspectivas de integração no ambiente do Apoio à Decisão, Editora da COOPE/UFRJ, Rio de Janeiro, 2000.
- [3] LINS, M.P.E., MOREIRA, M.C.B., Método I-O Stepwise para Seleção de Variáveis em Modelos de Análise Envoltória de Dados, Pesquisa Operacional, 19 (1), p. 39-50, 1999.
- [4] SOARES DE MELLO, J. C. C. B., GOMES, E. G., SOARES DE MELLO, M. H. C., ESTELLITA LINS, M. P., Método Multicritério para Seleção de Variáveis em Modelos DEA, Revista de Pesquisa Naval, Rio de Janeiro, v.15, p. 55-66, 2002.