

## **AVALIAÇÃO DE UM CURSO DE MATEMÁTICA À DISTÂNCIA COM MODELOS DEA E SELEÇÃO DE VARIÁVEIS**

**Fernando do Valle Silva Andrade**

*Graduação em Engenharia de Produção*

*Universidade Federal Fluminense (UFF)*

*Niterói – RJ*

fernando\_do\_valle@hotmail.com

**Luana Carneiro Brandão**

*Graduação em Engenharia de Produção*

*Universidade Federal Fluminense (UFF)*

*Niterói – RJ*

brandao\_luana@hotmail.com

**João Carlos C. B. Soares de Mello**

*Departamento de Engenharia de Produção*

*Universidade Federal Fluminense (UFF)*

*Niterói – RJ*

jcsmello@producao.uff.br

### **Resumo**

O artigo consiste em avaliar de forma quantitativa o curso de Matemática nos diversos pólos do Centro de Educação Superior à Distância do Estado do Rio de Janeiro (CEDERJ), consórcio entre a o Governo de Estado, prefeituras e algumas universidades. A ferramenta utilizada para isso foi a Análise Envoltória de Dados (DEA), que mede a eficiência relativa entre Unidades Decisoras (DMUs), combinada com um método de seleção de variáveis.

**Palavras-chave:** Educação à distância; DEA; seleção de variáveis.

### **Abstract**

This article evaluates, in quantitative terms, Mathematical courses offered by the Undergraduate Distance Education Center of Rio de Janeiro (CEDERJ), consortium between the state government, city halls and several universities. For this purpose, we used the Envelopment Data Analysis (DEA) tool, which measures the relative efficiency between different Decision Making Units (DMUs), combined with a method of variables selection.

Keywords: Distance education; DEA; selection of variables.

### **1. Introdução**

O Centro de Educação Superior à Distância do Estado do Rio de Janeiro (CEDERJ) é um consórcio elaborado em 1999 e conta com a parceria do Governo do Estado do Rio de Janeiro, das prefeituras das cidades que sediam os pólos regionais e das seis universidades públicas do Estado (UFF, UERJ,

UFRJ, UFRRJ, UENF e UNIRIO). O consórcio pretende contribuir para a interiorização do ensino superior público no Estado, bem como facilitar o acesso a cursos superiores de pessoas com indisponibilidade de estudar em horários normais.

O ensino superior à distância é alvo de várias discussões por parte de profissionais da área de educação, porém baseadas em análises qualitativas. Sendo assim, este artigo busca contribuir para essas análises com dados quantitativos, através da ferramenta Análise Envoltória de Dados (*Data Envelopment Analysis - DEA*).

O método consiste na relação comparada de entradas (*inputs*) e saídas (*outputs*) das diferentes DMUs. A eficiência relativa é calculada obtendo a razão entre a soma ponderada de seus produtos pela soma ponderada de seus insumos. Tais pesos são calculados através de programação linear de modo a maximizar a eficiência de cada DMU com relação ao conjunto de DMUs em análise.

Limitamos a aplicação da ferramenta ao curso de Matemática oferecido pelo CEDERJ, buscando iniciar o estudo em uma escala micro, possibilitando a continuidade da avaliação estendida a todos os cursos em trabalhos subsequentes.

## 2. Seleção de Variáveis

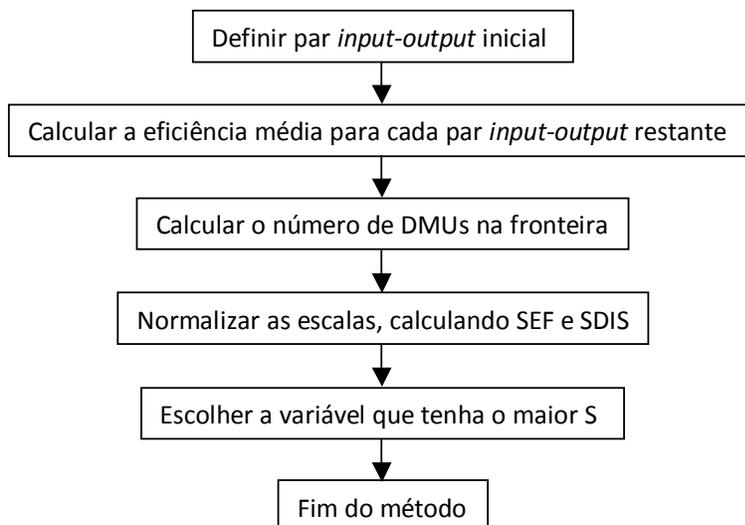
As variáveis de entradas e saída em DEA são fundamentais no cálculo das eficiências e por isso vários estudos sobre métodos de seleção de variáveis são conduzidos no meio acadêmico. Senra *et al* (2007) faz um estudo sobre os métodos *I-O Stepwise Exhaustivo*, *Multicritério*, *Multicritério Combinatório Inicial* e *Multicritério Combinatório Parcial* (por cenários). Os métodos, nesta seqüência, buscam a melhor eficiência média possível entre as DMUs e a melhor capacidade de ordenação. Isto significa uma maior explicabilidade das variáveis de saída pelas variáveis de entrada (maior eficiência média), mas sem causar um empate de DMUs eficientes possibilitando uma melhor ordenação das melhores unidades.

A utilização de seleção de variáveis neste estudo de caso se mostra interessante para que possamos encontrar a melhor relação causal possível entre os dados de entrada e os dados de saída. Além disso, levando em consideração que nosso principal objetivo é a avaliação (e para isso necessitamos de uma ordenação) dos pólos, a seleção de variáveis também é bastante útil para reduzir o número de DMUs eficientes.

Existe um leque farto de opções de métodos de seleção de variáveis que buscam os dois objetivos supracitados com importâncias diferentes dadas a cada um deles. A seleção de variáveis em DEA possibilita a obtenção de diferentes resultados para um mesmo conjunto de DMUs. Além disso, dependendo do método escolhido, pode-se necessitar de mais ou menos opiniões dos decisores, de acordo com as necessidades e a disponibilidade de informações.

No caso em estudo iremos utilizar um método com semelhanças com relação aos métodos Multicritério e Multicritério Combinatório Inicial para Seleção de Variáveis. Os dados disponíveis nos trazem 4 variáveis de entrada e apenas uma variável de saída. Entretanto, 3 das 4 variáveis de entrada são mutuamente excludentes como será mostrado posteriormente. Sendo assim, o par inicial já está definido e iremos apenas acrescentar mais uma variável que melhor represente a relação causal entre os dados de entrada e saída.

Podemos representar o método utilizado neste estudo de caso através do fluxograma da Figura 1.



**Figura 1** – Método de seleção de variáveis utilizado no estudo de caso.

Descrevendo os passos temos:

- 1) Definir par *input-output* inicial. Neste passo, conforme já descrito anteriormente, selecionamos o par *input-output* inicial como *default* definido pelos decisores.
- 2) Calcular a eficiência média para cada par *input-output* restante. Este passo consiste na geração de 3 modelos DEA, um para cada variável diferente restante. Para cada resultado calcula-se a eficiência média de todas as DMUs.
- 3) Calcular o número de DMUs na fronteira. Conta-se o número de DMUs eficientes encontrados para cada um dos 3 modelos DEA gerados.
- 4) Normalizar as escalas, calculando SEF e SDIS. Os valores obtidos para as eficiências médias são normalizados de modo a atribuir valor 1 à maior eficiência média e 0 à menor. A esse valor damos o nome de SEF. De maneira semelhante, normaliza-se o número de DMUs na fronteira, atribuindo valor 1 ao menor número de DMUs na fronteira obtido e 0 ao maior. A esse valor denominamos SDIS.
- 5) Escolher a variável que tenha o maior S. S é uma variável calculada pela média ponderada de SEF e SDIS. Pode-se escolher o peso dado a cada variável no cálculo de S de modo que a soma dos pesos seja igual a 1. Neste caso definimos pesos iguais de 0,5 para SEF e SDIS.

O método aqui desenvolvido é, na verdade, uma adaptação dos métodos expostos por Senra *et al* (2007) de acordo com as nossas necessidades e disponibilidade de dados.

### 3. Estudo de Caso

O CEDERJ oferece cursos de Licenciatura em Matemática atualmente em 24 pólos. Entretanto, apenas 13 pólos formaram alunos no primeiro semestre de 2009. Além disso, dispomos da informação do número de tutores por pólo. Neste caso totalizamos 28 pólos. Isto ocorre porque além de oferecer

apoio de tutores em todos os pólos que oferecem cursos de Matemática, alguns outros pólos também oferecem apoio de tutores mesmo sem haver a disponibilidade do curso naquele local.

Com base nisso, definimos as 3 principais variáveis do modelo:

**AI** – Número de alunos ingressantes por pólo em um determinado semestre (*input*);

**NT** – Número de tutores por pólo no primeiro semestre de 2009 (*input*);

**AF** – Número de alunos formados por cada pólo no primeiro semestre de 2009 (*output*).

De uma maneira simplificada, imaginamos que tais variáveis são as mais importantes variáveis básicas para medirmos a capacidade de converter alunos ingressantes em alunos formados. Dados os números de alunos ingressantes, o número de tutores disponíveis que auxiliam os alunos em sua formação e o número de alunos formados, o modelo estará bem descrito, de uma maneira bem simplificada.

A seleção de variáveis é importante justamente na escolha da variável **AI**. Dada a dificuldade de estimar com precisão o semestre principal de entrada dos alunos que formaram no primeiro semestre de 2009, a seleção de variáveis será capaz de nos fornecer a melhor relação causal possível da variável **AI** e da variável **NT** com relação à variável **AF**. Além disso, teremos uma melhor ordenação da eficiência dos pólos.

Vale ressaltar também que iremos considerar apenas os pólos que tiveram alunos ingressantes nos 3 semestres avaliados, além de alunos formados e tutores no primeiro semestre de 2009. Entradas ou saídas de valor 0 podem gerar distorções nos resultados do modelo.

No modelo utilizado, utilizaremos o modelo BCC, uma vez que temos pólos de tamanhos diferentes com diferentes quantidades de alunos e tutores. Sendo assim, para que possamos obter retornos variáveis de escala, utilizaremos o método supracitado.

Outra definição diz respeito à orientação do modelo a *output*. Isto significa maximizar os produtos obtidos, mantendo constantes os recursos definidos como *inputs*.

#### 4. Aplicação do Método

Sabendo que o curso de Licenciatura em Matemática oferecido pelo CEDERJ tem duração média de 8 períodos, sabemos que em média os ingressantes do segundo semestre de 2005 estariam formando no primeiro semestre de 2009. Colocando uma margem de um período de atraso e um período de antecipação, definimos os seguintes períodos para alunos ingressantes: 1/2005, 2/2005 e 1/2006.

A lista de pólos, suas abreviações e os valores dos dados de *inputs* e *output* são apresentados na Tabela 1.

**Tabela 1** – Valores de *inputs* e *outputs* de cada pólo com sua respectiva sigla.

| Sigla | Pólo           | AI - 1/2005 | AI - 2/2005 | AI - 1/2006 | NT | AF |
|-------|----------------|-------------|-------------|-------------|----|----|
| ARE   | Angra dos Reis | 57          | 60          | 48          | 6  | 8  |
| CAN   | Cantagalo      | 38          | 40          | 40          | 7  | 2  |
| CGR   | Campo Grande   | 11          | 62          | 61          | 6  | 1  |

|     |                     |    |    |    |    |    |
|-----|---------------------|----|----|----|----|----|
| ITA | Itaperuna           | 40 | 36 | 28 | 7  | 4  |
| MAC | Macaé               | 29 | 29 | 35 | 6  | 3  |
| PAR | Paracambi           | 73 | 72 | 49 | 7  | 9  |
| PET | Petrópolis          | 79 | 79 | 58 | 8  | 1  |
| PIR | Piraí               | 26 | 23 | 22 | 6  | 6  |
| SAQ | São Fidélis         | 59 | 61 | 49 | 6  | 2  |
| SFR | São Pedro da Aldeia | 32 | 20 | 21 | 5  | 1  |
| SPE | Saquarema           | 61 | 62 | 51 | 6  | 4  |
| TRI | Três Rios           | 61 | 60 | 50 | 8  | 3  |
| VRE | Volta Redonda       | 99 | 99 | 81 | 10 | 10 |

Para a resolução do modelo DEA, utilizamos o *software* SIAD (ANGULO MEZA *et al*, 2004). O Sistema Integrado de Apoio à Decisão nos fornece “todos os resultados dos modelos DEA clássicos (eficiência, pesos, alvos, benchmarks e folgas)” (SOARES DE MELLO *et al*, 2005).

As eficiências obtidas para cada variável **AI** analisada são apresentadas na Tabela 2.

**Tabela 2** – Escolha da variável **AI** para a análise do modelo.

| <i>Inputs</i>        | AI (1/2005),<br>NT | AI (2/2005),<br>NT | AI (1/2006),<br>NT |
|----------------------|--------------------|--------------------|--------------------|
| <i>Output</i>        | AF                 | AF                 | AF                 |
| DMU 1 - ARE          | <b>1,0000</b>      | <b>1,0000</b>      | <b>1,0000</b>      |
| DMU 2 - CAN          | 0,2952             | 0,2841             | 0,2500             |
| DMU 3 - CGR          | <b>1,0000</b>      | 0,1250             | 0,1250             |
| DMU 4 - ITA          | 0,5794             | 0,5886             | 0,6000             |
| DMU 5 - MAC          | 0,4844             | 0,4744             | 0,4286             |
| DMU 6 - PAR          | <b>1,0000</b>      | <b>1,0000</b>      | <b>1,0000</b>      |
| DMU 7 - PET          | 0,1083             | 0,1080             | 0,1077             |
| DMU 8 - PIR          | <b>1,0000</b>      | <b>1,0000</b>      | <b>1,0000</b>      |
| DMU 9 - SAQ          | 0,2500             | 0,2500             | 0,2500             |
| DMU 10 - SFR         | <b>1,0000</b>      | <b>1,0000</b>      | <b>1,0000</b>      |
| DMU 11 - SPE         | 0,5000             | 0,5000             | 0,5000             |
| DMU 12 - TRI         | 0,3636             | 0,3630             | 0,3322             |
| DMU 13 - VRE         | <b>1,0000</b>      | <b>1,0000</b>      | <b>1,0000</b>      |
| Eficiência Média     | 0,6601             | <b>0,5918</b>      | 0,5841             |
| DMUs na<br>Fronteira | 6                  | <b>5</b>           | 5                  |
| SEF                  | 1,0000             | <b>0,1007</b>      | 0,0000             |
| SDIS                 | 0                  | <b>1</b>           | 1                  |
| S                    | 0,5000             | <b>0,5504</b>      | 0,5000             |

Observamos que há um empate no número de DMUs na fronteira entre as variáveis **AI** (2/2005) e **AI** (1/2006). Sendo assim, como **AI** (2/2005) tem uma eficiência média maior, a variável **AI** (1/2006)

torna-se dominada. A decisão então fica entre os alunos que ingressaram no primeiro semestre de 2005 e no segundo semestre de 2005.

Como demos pesos iguais para SEF e SDIS na equação do cálculo de S, o simples fato do modelo com a variável **AI** (2/2005) não ter a pior eficiência média torna-o o de maior S. Vale observar que esta foi uma decisão arbitrária e que se fosse dado um peso maior para a eficiência média na escolha das variáveis, possivelmente teríamos escolhido a variável **AI** (1/2005) para entrar no modelo.

Com base nisso, podemos então realizar a análise dos resultados.

## 5. Análise dos Resultados

Conforme resultado obtido na seleção de variáveis realizada anteriormente, obtivemos resultados de eficiência para cada pólo, conforme Tabela 3.

**Tabela 3** – Eficiências ordenadas de cada pólo.

| <i>Inputs</i> | AI (2/2005), NT |
|---------------|-----------------|
| <i>Output</i> | AF              |
| DMU 1 - ARE   | <b>1,0000</b>   |
| DMU 6 - PAR   | <b>1,0000</b>   |
| DMU 8 - PIR   | <b>1,0000</b>   |
| DMU 10 - SFR  | <b>1,0000</b>   |
| DMU 13 - VRE  | <b>1,0000</b>   |
| DMU 4 - ITA   | 0,5886          |
| DMU 11 - SPE  | 0,5000          |
| DMU 5 - MAC   | 0,4744          |
| DMU 12 - TRI  | 0,3630          |
| DMU 2 - CAN   | 0,2841          |
| DMU 9 - SAQ   | 0,2500          |
| DMU 3 - CGR   | 0,1250          |
| DMU 7 - PET   | 0,1080          |

Com base nesses resultados, podemos observar que 5 pólos são considerados eficientes. São eles:

- Angra dos Reis;
- Paracambi;
- Pirai;
- São Pedro da Aldeia;
- Volta Redonda.

Além disso, para os pólos não eficientes, o modelo DEA traz *benchmarks* a serem observados. Tais *benchmarks* mostram os alvos das DMUs não eficientes, indicando em que aspectos melhorar para que cada pólo possa tornar-se eficiente.

A Tabela 4 traz os níveis de influência para *benchmarks* de cada DMU eficiente com relação a cada DMU. Vale ressaltar que para as DMUs eficientes, seus alvos são elas mesmas.

**Tabela 4** – Níveis de influência para *benchmarks*.

| DMU          | DMU 1 - ARE | DMU 6 - PAR | DMU 8 - PIR | DMU 10 - SFR | DMU 13 - VRE |
|--------------|-------------|-------------|-------------|--------------|--------------|
| DMU 1 - ARE  | 1           | 0           | 0           | 0            | 0            |
| DMU 2 - CAN  | 0           | 0,34693878  | 0,65306122  | 0            | 0            |
| DMU 3 - CGR  | 1           | 0           | 0           | 0            | 0            |
| DMU 4 - ITA  | 0           | 0,26530612  | 0,73469388  | 0            | 0            |
| DMU 5 - MAC  | 0,16216216  | 0           | 0,83783784  | 0            | 0            |
| DMU 6 - PAR  | 0           | 1           | 0           | 0            | 0            |
| DMU 7 - PET  | 0           | 0,74074074  | 0           | 0            | 0,25925926   |
| DMU 8 - PIR  | 0           | 0           | 1           | 0            | 0            |
| DMU 9 - SAQ  | 1           | 0           | 0           | 0            | 0            |
| DMU 10 - SFR | 0           | 0           | 0           | 1            | 0            |
| DMU 11 - SPE | 1           | 0           | 0           | 0            | 0            |
| DMU 12 - TRI | 0           | 0,75510204  | 0,24489796  | 0            | 0            |
| DMU 13 - VRE | 0           | 0           | 0           | 0            | 1            |

É interessante observar que São Pedro da Aldeia não é *benchmark* para nenhuma outra DMU. Isto ocorre porque, apesar de ser eficiente, o número de formandos é muito baixo. Sua eficiência se dá pelo fato de conseguir formar alunos mesmo tendo poucos alunos ingressantes e tutores disponíveis. Entretanto, como a orientação do modelo é a *outputs* (maximizar os *outputs* para a mesma quantidade de recursos), não é interessante para nenhum outro pólo conseguir apenas 1 aluno formando mantendo a mesma base de alunos ingressantes.

Outro ponto interessante é que o pólo de Angra dos Reis é *benchmark* único para outros 3 pólos (Campo Grande, São Fidélis e Saquarema). Isto é bem ilustrado pelo fato destes 3 pólos possuírem o mesmo número de tutores e praticamente o mesmo número de alunos ingressantes, apesar de formarem número bem menor de alunos. Isso mostra pólos com praticamente o mesmo tamanho, mas não conseguindo desenvolver seus alunos da mesma maneira.

Vale ressaltar ainda a baixíssima eficiência do pólo de Petrópolis. Mesmo sendo o segundo maior pólo em alunos ingressantes e número de tutores, além de estar localizado em cidade referência na região, formou apenas 1 aluno no primeiro semestre de 2009. Uma das possíveis justificativas para o resultado ruim talvez seja a boa disponibilidade de cursos superiores (incluindo cursos presenciais na região).

Petrópolis possui como *benchmark* principal o pólo de Paracambi, de tamanho semelhante. Além disso, caso o pólo deseje crescer ainda mais, pode espelhar-se no pólo de Volta Redonda: o maior pólo em valor absoluto de todas as variáveis analisadas.

Ainda com relação ao pólo de Volta Redonda, por ser um pólo bem maior que os demais, considerando os valores das variáveis analisadas, só poderia servir de *benchmark* para Petrópolis, o segundo maior. Entretanto, obviamente o quadro pode ser alterado caso haja maior oferta de vagas em outros pólos.

Outro ponto a ser observado são os alvos do pólo de Três Rios. Segundo os resultados fornecidos pelo SIAD, o alvo do número de tutores para este pólo é de 6,75 aproximadamente. Isto significa uma redução de quase 16%. Isto é reflexo do desempenho de outros pólos de tamanho semelhante que conseguem obter resultados melhores mesmo com número menor de tutores. Apesar de indicar uma redução de input, medidas importantes podem ser tomadas neste caso, como melhor capacitação e treinamento para os tutores deste pólo, bem como uma pesquisa entre os alunos para buscar suas principais carências.

## 6. Considerações Finais

É importante ressaltar que a abordagem deste artigo limita-se aos dados quantitativos disponíveis. O resultado do modelo DEA conduzido indica um caminho a ser observado e ressalta pontos de atenção. Entretanto, para avaliações mais profundas é muito importante analisar a estrutura dos pólos, a dimensão e importância da cidade dentro de sua respectiva região e ainda a disponibilidade de bons cursos superiores que poderiam gerar uma espécie de concorrência com os cursos do CEDERJ.

Vale reiterar a importância da discussão e da avaliação do ensino superior à distância no Brasil. Como sugestão para futuros trabalhos sobre este tema, é interessante observar o CEDERJ como um todo, incluindo todos os seus pólos e cursos. Além disso, mais variáveis podem ser levadas em consideração na formação dos alunos, também sendo boa sugestão para outras pesquisas.

Com relação à ferramenta técnica utilizada, consideramos o estudo de caso muito interessante para observar o poder de alcance da Análise Envoltória de Dados. Pudemos confirmar observações prévias mais destacadas e ainda alcançar bons resultados com análises mais profundas.

Consideramos como válidas as aplicações de DEA a casos de análise de eficiência educacional. Mais trabalhos semelhantes podem ser vistos em Abbott & Doucouliagos (2003) ou Abramo *et al* (2008).

## 7. Agradecimentos

À FAPERJ pelo apoio financeiro e ao CEDERJ por ter fornecido os dados.

## Referências Bibliográficas

Senra, L. F. A. D. C.; Nanci, L. C.; Soares de Mello, J. C. C. B.; Meza, L. A. *Estudo sobre métodos de seleção de variáveis em DEA*. **Pesquisa Operacional**, 27 (2): 2007. 191-207.

Angulo Meza, L.; Biondi Neto, L.; Soares de Mello, J.C.C.B.; Gomes, E.G.; Coelho, P.H.G. *FSDA – Free Software for Decision Analysis (SLAD – Software Livre de Apoio à Decisão): A Software Package for Data Envelopment Analysis Models*. In: XII Congresso Latino-iberoamericano de Investigación Operativa **Memorias...** La Habana: La Habana, 2004.

Soares de Mello, J.C.C.B.; Angulo Meza, L.; Gomes, E.G.; Biondi Neto, L. *Curso de Análise Envoltória de Dados*. Anais do XXXVII SBPO. Gramado: 2005.

Abbott, M.; Doucouliagos, C. *The efficiency of australian universities: A data envelopment analysis*. **Economics of Education Review**, 22 (1): 2003. 89-97.

Abramo, G.; D'Angelo, C. A.; Pugini, F. *The measurement of italian universities' research productivity by a non parametric-bibliometric methodology*. **Scientometrics** 76 (2): 2008. 225-244.