Avaliação da Eficiência de Cooperativas de Crédito através de DEA e PCA

Luís Guilherme Ribeiro Martins[1] (lgrmlgm@hotmail.com)

Maria Teresinha Arns Steiner[1] (maria.steiner@pucpr.br)

Amanda Trojan Fenerich[1] (amanda.fenerich@gmail.com)

Bruno Samways dos Santos[1] (brunosantos@utfpr.edu.br)

Luis Henrique de Lara Ramos[1] (ramos.luishenrique@gmail.com)

[1]PUC-PR: Rua Imaculada Conceição, 1155. Curitiba – PR.

RESUMO

O uso de ferramentas que auxiliem na avaliação de unidades produtivas pode se tornar um fator-chave para o gerenciamento empresarial, resultando em grande vantagem competitiva. Este trabalho apresenta a aplicação de *Data Envelopment Analysis* (DEA) e *Principal Components Analysis* (PCA) para avaliar a eficiência de cooperativas de crédito paranaenses. Esta avaliação foi realizada considerando dados trimestrais de cooperativas de crédito, entre Janeiro de 2009 e Julho de 2015 (27 períodos) e a escolha das variáveis foi auxiliada pela técnica PCA. A análise considerou 45 unidades (cada uma das cooperativas de crédito pesquisadas), sendo que de cada uma delas 10 variáveis (8 *inputs* e 2 *outputs*) em cada período. Os resultados da pesquisa são apresentados, concluindo-se que DEA e PCA são ferramentas eficientes e de fácil implementação e que, além disso, a sua aplicabilidade é mais indicada a ambientes com muitas unidades e/ou poucas variáveis.

Palavras-chaves: DEA, PCA, Cooperativas de Crédito.

Tópicos: DEA - Análise Envoltória de Dados

ABSTRACT

The use of tools which can aid in the productive units evaluation may become a key factor for companies management, resulting in an important competitive advantage. This paper presents the use of Data Envelopment Analysis (DEA) and Principal Components Analysis (PCA) in order to evaluate Paraná's credit union's efficiency. This evaluation has taken into account quarterly credit union's data, from January 2009 to July 2015 (27 periods). The variables selection has been taken by the PCA technique. The analysis considered 45 units (each credit union researched), each of one with 10 variables in each period (8 inputs and 2 outputs). The research results are presented, concluding that DEA and PCA are both efficient and easy to implement, and also more suitable to environments with several units and/or few variables.

Keywords: DEA, PCA, Credit Unions.

Topics: DEA –Data Envelopment Analysis

1. Introdução

Conforme o número de organizações aumenta, bem como a complexidade de suas tarefas, cresce a necessidade de ferramentas para auxiliar as suas avaliações de desempenho. Conhecer o desempenho individual dos componentes de uma cadeia produtiva é fundamental para traçar metas e planos para atingi-las.

Neste trabalho a Análise por Envoltória de Dados, também conhecida como Análise por Envelopamento de Dados (*Data Envelopment Analysis*; DEA), na versão BCC orientada ao *input*, é aplicada à realidade das cooperativas de crédito paranaenses. A seleção de dados é auxiliada por *Principal Component Analysis* (*PCA*), após uma exclusão inicial de dados irrelevantes.

A principal característica de DEA é ser não paramétrica, isto é, possibilita que a utilização de "pesos" não seja obrigatória para cada variável, como citado por [Boussofiane et al. 1991]. Além disso, a técnica dispensa a função de produção, ou seja, não há necessidade de se investigar a fundo as relações entre *inputs* e *outputs* da cadeia produtiva. Apesar disso, pode incluir pesos, se assim for desejado. Em outras palavras, tem a capacidade de analisar a eficiência de unidades produtivas com diversas variáveis sem escala comum [Mahdiloo et al. 2011].

As cooperativas de crédito podem ser intituladas como "maior", "melhor", dentre outras, dependendo dos critérios de classificação. Desta forma, concluir qual cooperativa é a "melhor" em aspectos que englobem duas ou mais variáveis escolhidas sem algum critério tem se mostrado inócuo e infrutífero, pela dificuldade em conceituar eficiência, como observado através do trabalho de [Fernandez et al. 2005]. A ferramenta DEA tem a capacidade de trabalhar simultaneamente com diversos *inputs* e outputs. Entretanto, a quantidade de variáveis incluídas no modelo é inversamente proporcional à qualidade da análise. Por essa razão, um método independente de seleção de variáveis é importante. Neste estudo é empregada a técnica PCA, largamente utilizada para obter informações de grandes bases de dados [Slottje 1991], [Zhu 1998], [Premachandra 2001], [Bruce Ho e Dash Wu 2009], [Adler 2010], [Dong et al 2015a], [Liang et al.2009], [Dong et al 2015b], [Chen et al 2015]. O objetivo dessa técnica é "resumir" as variáveis presentes em um sistema, com a mínima perda de informação possível. Neste estudo, o PCA é utilizado para determinar quais das características tabuladas são mais significativas para integrar a análise de eficiência com DEA.

A metodologia DEA tem sido utilizada para estudos na área de análise e otimização financeira por [Bruce Ho, Dash Wu 2009], [Donthu et al. 2005], [Eken e Kale 2011], [Holod e Lewis 2011], [Kuosmanen e Kortelainen, 2007], [Luo 2003], [Vilela et al. 2007], entre outros.

Este trabalho está organizado da seguinte forma: na seção 2 é apresentada a metodologia; na seção 3, os critérios e a seleção dos dados utilizados. Ao final das aplicações das técnicas são discutidos os resultados na seção 4. E, por fim, na seção 5, são apresentadas as conclusões.

2. Metodologia

Esta seção apresenta as técnicas DEA, na versão BCC e o PCA.

2.1. DEA

DEA é uma técnica não paramétrica de Programação Linear que trabalha com fronteiras de eficiência. Essa técnica é alimentada com variáveis de entrada e de saída. As variáveis de entrada são também chamadas de insumo ou *inputs*, ou seja, os recursos a partir dos quais a empresa produz. Alguns exemplos são: capitais à sua disposição; capital investido; matéria prima; escolaridade de seus colaboradores; entre outros. As variáveis de saída são também chamadas de produto ou *outputs*. São os resultados da produtividade da empresa, como por exemplo: produção bruta; lucro; incremento de *marketshare*; vendas; *turnover*; entre outros.

A técnica DEA tem sido bastante estudada e ganhando destaque nos últimos anos. Em rápidas pesquisas por motores de busca online pode-se encontrar publicações recentes da ordem de centenas nos últimos anos. Alguns dos trabalhos que embasam o presente trabalho são

[Helfand e Levine 2004], [Anderson et al.2011], [Haugland et al. 2007], [Luo 2003], [Vilela et al. 2007], [Ceretta e Niederauder2001], [Macedo et al. 2010], [Souza et al. 2011], [Gonçalves et al. 2007], [Marinho 2003], [Marinho et al. 2004], [Souza 2006] e [Vaninsky2006].

Através de DEA podem ser comparadas apenas unidades de produção semelhantes, também chamadas de unidades tomadoras de decisão (*decision making units*; DMU), e classificá-las em termos de eficiência. Os modelos DEA calculam uma fronteira ótima de produção, ou seja, uma curva de máxima produtividade. Essa fronteira representa as diversas combinações consideradas eficientes entre *inputs* e *outputs* para as DMUs avaliadas. Caso o nível de produção, de uma DMU estivesse situado sobre essa fronteira, ela seria considerada 100% eficiente; caso contrário, a classificação de cada unidade avaliada seria obtida sob um "percentual de eficiência", conforme a distância que a sua produtividade atual estivesse em relação à fronteira de eficiência.

A diferença entre essa técnica e outras que calculam a tendência média de eficiência é que as DMUs são comparadas levando em consideração as melhores práticas produtivas.

A Figura 1 representa a diferença entre uma função de produção média (obtida através de Regressão Linear, por exemplo) e uma função de produção de fronteira (obtida através de DEA, por exemplo):

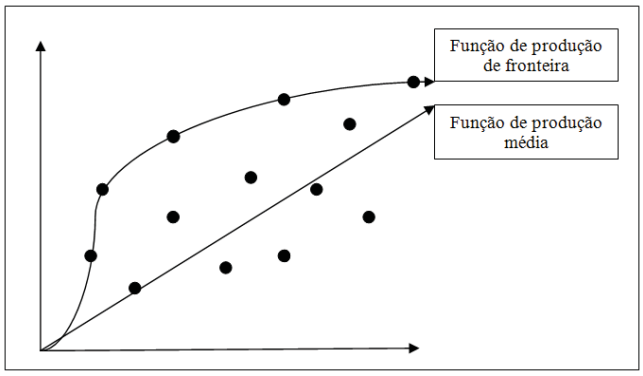

Figura 1 – Função de Produção de Fronteira e Função de Produção Média

A metodologia DEA possui características que podem incentivar ou não sua aplicação, como descrito no trabalho pioneiro de [Banker et al.1984] e de vários outros pesquisadores como [Charnes et al. 1997], [Helfand e Levine 2004], [Anderson et al. 2011], [Haugland et al. 2007], [Kassai 2002], [Luo 2003], [Vilela et al. 2007], [Ceretta e Niederauder 2001], [Macedo et al. 2010], [Souza et al. 2011], [Gonçalves et al. 2007], [Marinho 2003], [Marinho et al. 2004], [Souza 2006], [Weyman-Jones et al.2012] e [Vaninsky 2006], elencadas abaixo:

- produz uma medida agregada individual para cada DMU, de acordo com sua utilização de insumos para produzir os produtos almejados;
- tem foco em observações individuais em contraste com as da população;
- pode utilizar múltiplos produtos e insumos em diferentes unidades de medida;
- pode ser ajustada para variáveis exógenas, ou seja, integrar variáveis novas se necessário;
- pode incorporar variáveis categóricas, ou seja, que assumem valores "1" ou "0";
- é livre de valores e não requer especificação de pesos ou preços de insumos, a princípio;
- não coloca restrições à forma funcional da função de produção;
- pode incorporar julgamentos quando desejado;
- produz estimativas específicas das mudanças almejadas nos insumos e produtos para projeção das DMUs localizadas abaixo da fronteira de eficiência sobre a fronteira;
- é Pareto eficiente;
- enfatiza as melhores práticas identificadas, ao invés das medidas de tendência central, como ocorre em técnicas baseadas em Regressão Linear;
- satisfaz ao critério de equidade estrita na avaliação relativa de cada DMU;

- é sensível à presença de *outliers*;
- analisa efetivamente a eficiência relativa de uma DMU na presença de diversas variáveis;
- utiliza o conceito de fronteira de eficiência como um padrão de excelência, porém com múltiplos formatos considerados eficientes. Pode estabelecer inúmeros conjuntos de *inputs*-*outputs* considerados simultaneamente 100% eficientes;
- pode sobrepujar a falta de medida comum entre as variáveis sem a necessidade de atribuição de pesos arbitrariamente;
- não necessita, a princípio, de uma função de produção explícita, dispensando estimativas;
- pode ser facilmente computadorizado;
- funciona melhor quanto maior a razão entre o número de DMUs e o número de variáveis, ou seja, $\frac{n^o \ de\ DMUs}{n^o\, de\ variáveis}$;
- permite identificar pontos fracos e fortes das DMUs.

2.1.1. Modelo DEA BCC

A metodologia DEA possui diversas versões. Neste estudo foi utilizado o modelo BCC, voltado para o *input*, inicialmente proposto por Banker, Charnes e Cooper [Banker et al. 1984], cujas iniciais dos pesquisadores deram origem ao nome do modelo. Este modelo leva em conta o retorno de escala, que é o ganho não proporcional ao crescimento dos *inputs*, podendo tanto ser mais que proporcional ou menos que proporcional, inclusive mudando de comportamento conforme o nível de produção se altera.

O modelo mostrado pelas expressões de (1) a (4), a seguir, apresenta o BCC orientado para o *input*. A função objetivo (1) procura estabelecer qual seria o nível de produção máximo para a DMU *k***.** A primeira restrição, em (2), garante que seja usado o mesmo nível de *inputs*, visto que se deseja produzir mais utilizando os mesmos recursos. A segunda restrição, em (3), coloca novamente a variável u'_k considerando o retorno de escala na relação entre os *inputs* e *outputs*, ou seja, no resultado global da unidade. Isto porque cada uma das DMUs pode estar trabalhando em escalas diferentes de produtividade, mas que ao serem comparadas sob a mesma ótica podem ainda ser consideradas igualmente eficientes. A restrição (4) é a de não-negatividade. Após a implementação, o modelo retorna valores da função objetivo entre "0" e "1", ou seja, entre 0% e 100% de eficiência para a DMU *k* analisada. O modelo deve ser resolvido N vezes, ou seja, uma para cada DMU. Em outras palavras, esta versão do modelo indica a distância do nível de produção atual da DMU *k* ao máximo nível de eficiência possível no ambiente em que se encontra (fronteira de produção). Se o nível de produção da DMU *k* estiver situado sobre a fronteira, a DMU *k* é considerada 100% eficiente. Caso contrário, o modelo retorna o valor correspondente à distância entre o nível atual de produção da DMU *k* e a fronteira de produção.

$$Maximizar \sum_{r=1}^{m} u_r\, y_{rk} - u'_k \qquad (1)$$

Sujeito a:

$$\sum_{i=1}^{n} v_i\, x_{ik} = 1 \qquad (2)$$

$$\sum_{r=1}^{m} u_r\, y_{rj} - \sum_{i=1}^{n} v_i x_{ij} - u'_k \leq 0 \qquad (3)$$

$$u_{r,} v_i \geq 0 \qquad (4)$$

onde:

- u_r: *peso do output r* (*r* = 1, ..., *m*; *m* = quantidade de *outputs*, neste caso 8);
- y_{rk}: *output r* da DMU *k* (*k* é a DMU que está sendo analisada);
- u'_k: retorno de escala da DMU *k*;

- v_i: peso do *input i* (i = 1, ..., n; n = quantidade de *inputs*, neste caso 2)
- x_{ik}: *input i* da DMU k;
- y_{rj}: *output r* da DMU j (j = 1, ..., N; N = quantidade de DMUs, neste caso 45)
- x_{ij}: *input i* da DMU j.

A Figura 2 ilustra a alteração nos valores de *inputs* que poderiam tornar uma unidade 100% eficiente. Neste caso particular, os pontos A e B correspondem a unidades eficientes, pois estão situadas sobre a fronteira de eficiência. Os pontos C e D são unidades ineficientes. O ponto K representa também uma unidade ineficiente, com nível de *input i*. Se os recursos da unidade K fossem reduzidos a i', sem alteração do *output*, então o novo índice de eficiência seria K', situado sobre a fronteira, ou seja, seria considerado 100% eficiente.

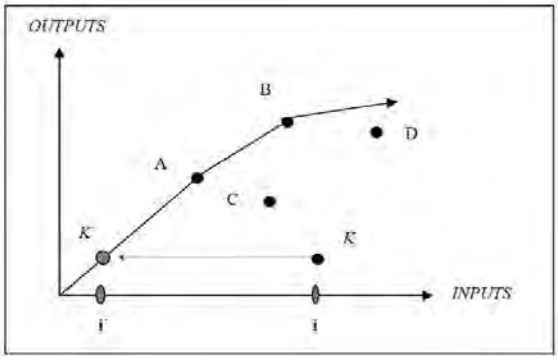

Figura 2: Alteração do nível de *input* para tornar uma unidade eficiente.
Fonte: adaptado de [Gonçalves et al. 2007]

Vale destacar que a quantidade de variáveis a serem incluídas na aplicação de DEA, seguindo [Fernandez et al. 2005], [Bruce Ho e Dash Wu 2009] e [Unsal e Orkcu 2015] pode alterar o resultado dos índices de eficiência obtidos.Têm-se a seguir, sugestões com relação a esta quantidade (máxima ou mínima de DMUs):

- Boussofiane e Dyson, 1991: nº de DMUs $\geq$ nº de *inputs* * nº de *outputs*
- Golany e Roll, 1989: nº de DMUs $\geq 2 *$ (nº de *inputs* + nº de *outputs*)
- Bowlin, 1998: nº de DMUs $\geq 3 *$ (nº de *inputs* + nº de *outputs*)
- Dyson et al., 2001: nº de DMUs $\geq 2 *$ (nº de *inputs* *nº de *outputs*).

Já a escolha das variáveis é feita através da técnica PCA, discutida na seção 2.2 a seguir.

2.2. PCA

Existem diversas formas de se "filtrar" os dados úteis de uma base de dados, de forma a evidenciar apenas as características mais importantes de um sistema. Alguns exemplos são: PLS (*Partial Least Squares*), MLR (*Multiple Linear Regression*), FA (*Factor Analysis*), dentre outros. Neste estudo foi aplicada a técnica PCA, largamente utilizada para obter informações de grandes bases de dados, segundo indicado por vários pesquisadores, dentre os quais têm-se: [Slottje 1991], [Zhu 1998], [Premachandra 2001], [Bruce Ho e Dash Wu 2009], [Adler 2010], [Dong et al. 2015a], [Liang et al. 2009], [Dong et al. 2015], [Chen et al. 2015].

O objetivo da técnica PCA é reduzir o tamanho do conjunto de dados, através de combinações lineares entre as variáveis originais, mantendo o máximo de informações possíveis [Pavanelli et al. 2001]. Esta técnica traça um vetor em um hiperplano onde as distâncias para os valores observados de cada variável seja a menor possível. Esse vetor é a primeira componente principal. A segunda componente principal é um vetor ortogonal ao primeiro, com as distâncias para aos valores observados, novamente a menor possível. Naturalmente, as distâncias à primeira

componente principal serão menores, pois o segundo vetor (e os seguintes) são inseridos em sequência. Observa-se que as variáveis com maior representatividade nos principais componentes são as mais importantes no sistema como um todo. Pode-se então substituir o conjunto original pelos vetores resultantes, ou utilizar a técnica para identificar os fatores principais de um sistema e ignorar os demais [Camacho et al. 2010].

Neste trabalho, o PCA contribui para a seleção de dados a fim de integrar a análise com DEA, conforme descrito na seção 3, a seguir.

3. Seleção dos Dados

Os dados para análise são características apresentadas por cooperativas de crédito, fornecidas pelo Sistema OCEPAR, formado por três sociedades distintas, a saber: Ocepar, Sescoop PR e Fecoopar. Estas três sociedades são compostas por cooperativas que, por sua vez, se subdividem em cooperativas singulares, objetos da presente análise.

O conjunto de dados possui variáveis observadas trimestralmente no período de Janeiro de 2009 a Julho de 2015, totalizando 27 observações, em três sistemas de cooperativas de crédito paranaenses: Sicoob-Unicoob, Sicredi e Uniprime, sendo que cada sistema é um conjunto de cooperativas singulares. Os dados utilizados compreendem um total de 45 cooperativas singulares, sendo 16 do Sistema Sicoob-Unicoob, 24 do Sistema Sicredi e 5 do Sistema Uniprime. Cada uma dessas cooperativas singulares foi considerada uma unidade tomadora de decisão (DMU) autônoma.

A Figura 3 a seguir representa o fluxo da seleção de dados realizada neste estudo.

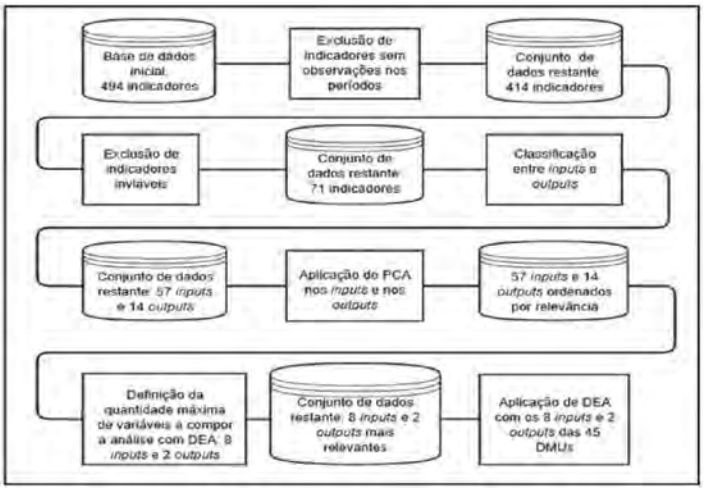

Figura 3: Fluxo de seleção de dados

A base de dados utilizada possui 494 indicadores, os quais foram reduzidos, inicialmente para 414 indicadores (414 na Sicoob-Unicoob, 334 na Sicredi e 373 na Uniprime), já que ao menos 80 (80 na Sicoob-Unicoob, 160 na Sicredi e 121 na Uniprime) não possuíam nenhuma observação válida em todo período compreendido nesta análise**.** Depois disso, excluiu-se os indicadores:

- que continham informações econômicas (IPCA, CUB, variação câmbio, etc.);
- que eram repetidos (um em valores absolutos e outro percentual, por exemplo);
- que não foram oriundos das DMUs (dados setorizados, por exemplo);
- que eram muito específicos e de difícil generalização (medida de eficiência de um projeto social específico de uma singular, por exemplo);
- com valores médios inferiores a 0,5, denotando muitas lacunas nas observações;

- com menos de 20 observações durante o período analisado.

Após este procedimento restaram 71 indicadores, divididos entre 57 *inputs* e 14 *outputs*. Considerou-se para a escolha da quantidade de variáveis neste trabalho, o critério apontado por [Dysonet al. 2001], citado na seção 2, onde a maior quantidade de variáveis, restrita às 45 DMUs disponíveis, é de 8 *inputs* e 2 *outputs*.

Ao classificar os *inputs*, os 8 prioritários foram:

1. Passivo operacional;
2. Fontes de recursos;
3. Depósitos a prazo;
4. Recursos administrados;
5. Patrimônio líquido;
6. Ativo operacional;
7. Reservas e fundos;
8. Capital social integralizado.

Quanto aos *outputs*, os 2 prioritários foram:

1. Ingressos e receitas líquidas;
2. Resultado à disposição da Assembléia Geral Ordinária (AGO)

A lista das cooperativas singulares está a seguir, apresentada por meio dos códigos utilizados pelas próprias cooperativas, sem propriedades quantitativas:

- Sicoob-Unicoob: 416, 431, 452, 453, 462, 469, 473, 478, 479, 484, 485, 489, 490, 493, 507, 546.
- Sicredi: 210, 233, 237, 250, 251, 252, 257, 260, 289, 290, 299, 301, 304, 315, 318, 319, 346, 357, 358, 406, 464, 492, 498, 517.
- Uniprime: 419, 448, 449, 450, 465.

As variáveis com valores negativos, quando observadas,foram substituídas por valores positivos muito pequenos ("1", neste caso), seguindo a proposta de [Bowling 1998], já que esta versão de DEA não trabalha tais grandezas. A base de dados inicial compunha 45 tabelas de 494 linhas e 27 colunas, inviabilizando sua representação gráfica neste trabalho. Ao invés disso, está presente como exemplo uma das 45 tabelas, apresentada no Quadro 1 a seguir, com a observação de um período (Janeiro de 2009), de uma singular (a 484), já com as variáveis selecionadas através de PCA, com os dados em reais.

Indicadores	01/2009 (R$)
ATIVO OPERACIONAL	14.435.904,49
PASSIVO OPERACIONAL	10.865.398,90
PATRIMÔNIO LÍQUIDO	3.790.624,34
INGRESSOS E RECEITAS LÍQUIDAS	202.009,62
RECURSOS ADMINISTRADOS	14.656.023,24
DEPÓSITOS A PRAZO	7.883.683,10
FONTES DE RECURSOS	14.617.936,43
RESERVAS E FUNDOS	1.472.665,02
CAPITAL SOCIAL INTEGRALIZADO	2.311.514,82
RESULTADO À DISPOSIÇÃO DA AGO	109.618,47

Quadro 1: Exemplo de dados de uma dasCooperativas de Crédito Singulares.
Fonte: OCEPAR, 2015.

4. Obtenção e Discussão dos Resultados

A aplicação da metodologia DEA foi feita com a versão BCC, a princípio voltada para o *input*, descrita na seção 2, por meio do modelo matemático (1) a (4). Assim, obtém-se um conjunto de resultados para cada um dos 27 trimestres disponíveis para as 45 DMUs, considerando as 10 variáveis (8 *inputs* e 2 *outputs*) indicadas na seção 3.

Apenas as cooperativas singulares 448, 450, 465 e 546 não foram consideradas nos períodos entre 01/2009-10/2009, 01/2009-10/2009, 01/2009-10/2010 e 01/2009-01/2010, respectivamente, pois não continham dados observados suficientes para possibilitar a análise. Nestes casos "não avaliáveis", as células foram preenchidas com NA. No entanto, as referidas cooperativas continuam a ser avaliadas e fazer parte do panorama nos demais períodos.

Para a operacionalização do PCA foi utilizado o suplemento para o *MS Excel* chamado *Multibase*, desenvolvido pela empresa *Numerical Dynamics Japan*.

Os índices de eficiência foram calculados implementando-se a versão BCC de DEA através do aplicativo *MaxDea Basic*, desenvolvido pela empresa *Beijing Realworld Software Company Ltd*, versão 6.6, orientado para o *input*, considerando retornos de escala variáveis. Os conjuntos de eficiência são apresentados resumidamente no Quadro 2, a seguir, demonstrando um trimestre por ano analisado.

Cooperativas	Índices de Eficiência DEA BCC Orientado ao *Input*							Cooperativas	Índices de Eficiência DEA BCC Orientado ao *Input*						
	jul/09	jul/10	jul/11	jul/12	jul/13	jul/14	jul/15		jul/09	jul/10	jul/11	jul/12	jul/13	jul/14	jul/15
210	0.764	0.834	1.000	0.996	1.000	0.976	0.868	**431**	1.000	0.673	1.000	0.768	1.000	0.87	0.745
233	1.000	1.000	1.000	1.000	1.000	1.000	1.000	**449**	1.000	0.884	1.000	1.000	1.000	1.000	1.000
237	1.000	1.000	0.973	1.000	0.903	1.000	1.000	**452**	0.954	0.777	0.903	1.000	1.000	1.000	1.000
250	0.762	0.832	0.908	0.973	1.000	1.000	1.000	**453**	1.000	1.000	1.000	1.000	1.000	1.000	1.000
251	1.000	0.852	1.000	1.000	1.000	0.99	0.922	**462**	0.552	0.563	0.695	0.692	0.719	0.705	0.575
252	0.87	0.695	0.769	0.714	1.000	0.859	0.962	**464**	0.538	0.474	1.000	0.78	0.625	0.712	0.649
257	1.000	1.000	1.000	1.000	1.000	1.000	1.000	**469**	0.538	0.502	0.776	0.846	1.000	0.979	0.997
260	0.906	1.000	1.000	1.000	1.000	1.000	1.000	**473**	0.727	0.577	1.000	0.767	0.953	0.896	0.753
289	0.88	0.923	1.000	0.89	1.000	1.000	0.927	**478**	0.886	0.939	0.939	1.000	1.000	1.000	0.924
290	0.894	0.876	1.000	0.985	1.000	0.997	0.969	**479**	0.827	0.7	1.000	1.000	0.79	1.000	0.715
299	1.000	0.794	0.993	0.962	1.000	1.000	1.000	**484**	0.669	0.567	0.754	0.871	0.825	0.757	0.632
301	1.000	1.000	0.885	1.000	0.938	0.967	1.000	**485**	0.738	0.655	1.000	0.992	0.857	0.977	0.92
304	1.000	1.000	0.882	0.945	1.000	0.846	0.908	**489**	0.79	0.74	0.985	0.843	0.92	1.000	0.773
315	0.936	0.934	1.000	1.000	0.956	0.853	1.000	**490**	0.965	0.83	0.788	0.885	0.862	0.915	0.894
318	1.000	1.000	1.000	1.000	1.000	0.97	1.000	**492**	1.000	1.000	0.992	0.834	1.000	0.716	1.000
346	1.000	0.951	0.97	1.000	1.000	1.000	1.000	**498**	1.000	1.000	1.000	1.000	1.000	1.000	1.000
357	0.716	0.766	0.742	0.753	0.689	0.69	0.706	**507**	1.000	1.000	1.000	1.000	0.959	1.000	1.000
358	0.845	0.92	0.93	1.000	1.000	0.898	1.000	**517**	1.000	1.000	1.000	1.000	1.000	1.000	1.000
406	1.000	0.82	0.794	0.799	1.000	1.000	0.813	**448**	**NA**	1.000	1.000	1.000	1.000	1.000	1.000
416	0.741	0.68	0.717	0.77	0.698	0.797	0.743	**450**	**NA**	0.809	0.853	1.000	1.000	1.000	1.000
419	0.567	1.000	1.000	1.000	1.000	1.000	1.000	**465**	**NA**	**NA**	1.000	1.000	1.000	1.000	1.000
								546	**NA**	1.000	1.000	1.000	0.88	1.000	1.000

Quadro 2: índices de Eficiência de Cooperativas Paranaenses entre Julho de 2009 e Julho de 2015
Fonte: Autor

Observa-se que dos 1.194 índices de eficiência obtidos (27 índices de cada uma das 45 cooperativas, excluindo-se os 21 índices não avaliáveis), 673 foram consideradas 100% eficientes. Mesmo assim se conseguiu 44% de observações que poderiam ajudar os gestores a identificar pontos fracos a serem melhorados.

Os resultados obtidos durante a pesquisa levaram ao questionamento: seriam os índices de eficiência muito diferentes caso calculados de acordo com a versão BCC de DEA voltado para o *output* e não para o *input*? Como o perfil das variáveis poderia indicar que os *inputs* não são de fácil redução, ou mesmo que essa redução não fosse desejável, a versão orientada para o *output* foi investigada. Por isso, os dados selecionados também foram submetidos à análise sob o modelo DEA na versão BCC voltado para o *output* [Banker et al. 1984]. Os resultados obtidos para esta situação não foram relevantemente diferentes dos apresentados neste trabalho. As variações foram pequenas (índice de correlação superior a 99%), indicando que tanto uma diminuição dos níveis de *inputs* ou um aumento dos níveis de *outputs* (ou ainda uma combinação de ambos) seriam eficazes para trazer as cooperativas para um nível desejado de eficiência, conforme .

Observa-se ainda que algumas DMUs foram consideradas 100% eficientes em todos os períodos analisados, a exemplo das cooperativas 453, 498 e 517. Essas poderiam ser consideradas *benchmarks*, ou seja, suas práticas poderiam positivamente as demais. Por outro lado, não existiu nenhuma cooperativa sem classificação máxima (100%) de eficiência em pelo menos uma observação. Isto pode indicar que as diferenças entre as práticas que levam a uma melhora dos níveis de eficiência podem não ser tão evidentes. Ainda assim é possível notar a diferença ao longo dos períodos avaliados entre determinadas cooperativas como, por exemplo, o desempenho da cooperativa 257 (100% eficiente em todas as observações) e a cooperativa 462 (100% eficiente em apenas duas observações): é evidente a diferença entre seus desempenhos. Neste caso, o foco dos esforços destas duas unidades pode estar em direções distintas, ou ainda, pode ser que alguma variável exógena seja responsável pela lacuna de desempenho (peculiaridades da praça de localização, políticas de crédito, entre outras), cabendo investigação pelos gestores.

5. Conclusões

Observou-se durante a presente pesquisa que DEA é uma ferramenta importante para comparar unidades produtivas, com vistas à otimização de seus resultados. Por meio desta técnica, pode-se identificar DMUs com potencial a ser explorado ou para tornarem-se *benchmarks*, evidenciar *inputs* a serem economizados, dentre outras informações relevantes. Entretanto, como analisado no trabalho de [Fernandez et al. 2005], a maneira de se definir eficiência, de se selecionar dados e DMUs a integrarem a análise, assim como a sua quantidade, podem alterar os resultados. Para manter a imparcialidade e reprodutibilidade da seleção de variáveis que irão "alimentar" o modelo, a técnica PCA mostrou-se eficiente e de fácil implementação, com baixo custo computacional.

Assim, observa-se uma característica interessante do modelo: não é de aplicabilidade simples, embora a implementação seja. Como uma das principais restrições da ferramenta é justamente a relação entre variáveis e DMUs a serem incluídas no cálculo, neste caso, para envolver todas as variáveis tabuladas (57 *inputs* e 14 *outputs*), seriam necessárias 1.596 DMUs. Em outras palavras, a aplicação de DEA sem a utilização de PCA seria eficaz em um ambiente com muitas unidades produtivas, podendo ser rotuladas com relativamente poucas variáveis. Um ambiente com estas características não é encontrado com freqüência, já que grande parte da bibliografia revisada possuía, de uma forma ou de outra, seleção de variáveis.

Outra evidência encontrada foi a da necessidade de analistas com conhecimento da ferramenta. A escolha das próprias DMUs e períodos de análise podem alterar os resultados, cabendo a eles (analistas), a sensibilidade inexistente no modelo matemático, para fazer as melhores escolhas. Uma possibilidade para estudo futuro pode ser realizar a integração entre as ferramentas aqui utilizadas com outras que incluam a preferência de decisores na escolha de

variáveis e das DMUs a serem consideradas no modelo, como modelos de maximização de utilidade ou otimização de preferências.

Referências

Anderson, G., Crawford, I. e Leicester, A. (2011). Welfare Rankings from Multivariate Data, a Nonparametric Approach. *Journal of Public Economics*, v. 95, p. 247–252.

Banker, R. D., Charnes, A. e Cooper, W. W. (1984). Some Models for Estimating Technical and Scale Inefficiencies in Data Envelopment Analysis. *Management Science*, v. 30, nº 9.

Boussofiane, A. e Dyson, R. G. (1991). Applied Data Envelopment Analysis. *European Journal of Operational Research*, v.52, p.1-15.

Bowlin, W. F. (1998). Measuring Performance: An Introduction to Data Envelopment Analysis (DEA).*Journal of Cost Analysis*, v.7, 3-27.

Bruce Ho, C. T. e Dash Wu, D. S. (2009). Online banking performance evaluation using data envelopment analysis and principal component analysis. *Computers & Operations Research,* v. 36, p. 1835-1842.

Camacho, J., Pico, J. e Ferrer, A. (2010).Data understanding with PCA: Structural and Variance Information plots. *Chemometrics and Intelligent Laboratory Systems*, v. 100, p. 48–56.

Ceretta, P. S. e Niederauer, C. A. P. (2001). Rentabilidade e Eficiência no Setor Bancário Brasileiro. *Revista Administração Contemporânea*, v. 5, p. 07-26. Set./Dez.

Charnes, A., Cooper, W. W., Lewin, A. Y. e Seiford, L. M. (1997). Data Envelopment Analysis: theory, methodology and application. Massachussets (EUA): Kluwer.

Chen, J., Wan, Z., Zhang, F., Park, N., He, X. e Yin, W. (2015). Operational Efficiency Evaluation of Iron Ore Logistics at the Ports of Bohai Bay in China: Based on the PCA-DEA Model. *Mathematical Problems in Engineering*, v. 16.

Dong, F., Mitchell, P. D. e Colquhoun, J. (2015). Measuring farm sustainability using data envelope analysis with principal components: The case of Wisconsin cranberry. *Journal of Environmental Management*, v. 147, p. 175 -183.

Dong, F., Mitchell, P. D., Knuteson, D., Wyman, J., Bussan, A. J. e Conley, S. (2015). Assessing sustainability and improvements in US Midwestern soybean production systems using a PCA–DEA approach. *Renewable Agriculture and Food Systems*, p. 1-16.

Donthu, A, N., Hershbergerb, E. K. e Osmonbekov, T. (2005). Benchmarking Marketing Productivity Using Data Envelopment Analysis. *Journal of Business Research*, v.58, p. 1474–1482.

Dyson, R. G., Allen, R., Camanho, A. S., Podinovski, V. V., Sarrico, C.S. e Shale, E.A. (2001). Pitfalls and Protocols in DEA. *European Journal of Operational Research*, 132, 245-259, 2001.

Eken, M. H. e Kale, S. (2011). Measuring Bank Branch Performance Using Data Envelopment Analysis (DEA): The case of Turkish bank branches. *African Journal of Business Management*, v. 5, p. 889-901.

Fernandez, C., Koop, B. G. e Steel, M. F. J.(2005). Alternative Efficiency Measures for Multiple-Output Production. *Journal of Econometrics*, v.126, p. 411-444.

Golany, B. e Roll, Y. (1989). An Application Procedure for DEA. *Omega*, v. 17, 237-250.

Gonçalves, A. C., Noronha, C. P., Lins, M. P. E. e Almeida, R. M. V. R. (2007). Análise Envoltória de Dados na Avaliação de Hospitais Públicos nas Capitais Brasileiras. *Revista Saúde Pública*; v. 41, p.427-35.

Haugland, S. A., Myrtveit, I. e Nygaard, A. (2007). Market Orientation and Performance in the Service Industry: A Data Envelopment Analysis. *Journal of Business Research*. v. 60, p. 1191–1197.

Helfand. S. M. e Levine, E. S. (2004). Farm size and the determinants of productive efficiency in the Brazilian Center-West. *Agricultural Economics.* v. 31, p. 241–249.

Holod, D. e Lewis, H. F. (2011). Resolving the deposit dilemma: a new DEA bank efficiency model. *Journal of Banking & Finance*, v.35, p.2801–2810.

Kassai, S. (2002). Utilização da análise por envoltória de dados (DEA) na análise de demonstrações contábeis. Tese de Doutorado - USP.

Kuosmanen, T. e Kortelainen, M. (2007). Valuing Environmental Factors In Cost–Benefit Analysis Using Data Envelopment Analysis. *Ecological Economics*, v.62, p. 56–65.

Liang, L., Yongjun, L. e Shibing, L. (2009). Increasing the discriminatory power of DEA in the presence of the undesirable *outputs* and large dimensionality of data sets with PCA. *Expert Systems with Applications*, v. 36, p. 5895–5899.

Luo, X. (2003). Evaluating the profitability and marketability efficiency of large banks: an application of data envelopment analysis. *Journal of Business Research*, v. 56, p. 627- 635.

Macedo, A. S., Cípola, M. F. e Ferreira, A. F. R. (2010). Desempenho Social no Agronegócio Brasileiro: aplicando DEA no Segmento de Usinas e Processamento de Cana-de-açúcar. *Revista de Economia e Sociologia Rural*, v. 48, n. 1, p. 223-243.

Mahdiloo, M., Noorizadeh, A. e Saen F. R. (2011). Developing a new data envelopment analysis model for customer value analysis. *Journal of industrial and Management Optimization*, v.7, p. 531–558.

Marinho, A. (2003). Avaliação da Eficiência técnica nos Serviços de Saúde nos Municípios do Estado do Rio de Janeiro. *Revista Brasileira de Economia Rio de Janeiro*, v. 57, p. 515-534.

Marinho, E., Soares, F. e Benegas, M. (2004). Desigualdade de Renda e Eficiência Técnica na Geração de Bem-Estar entre os Estados Brasileiros. *Revista Brasileira de Economia do Rio de Janeiro*, v.58, p. 583-608.

Pavanelli, A. M.; Pavanelli, G., Steiner, M. T. A., Costa, D. M. B. e Gusmão, B. G. (2011). Técnicas De Reconhecimento De Padrões Aplicadas À Justiça Do Trabalho. Revista Eletrônica Pesquisa Operacional para o Desenvolvimento, v. 3, n. 2, p. 90 – 106.

Premachandra, I. M. (2001). A note on DEA vs principal component analysis: An improvement to Joe Zhu's approach. *European Journal of Operational Research*, v. 132, p.553-560.

Slottje, D. J. (1991). Measuring the Quality of Life Across Countries. *The Review of Economics and Statistics*, v. 73.

Souza, G. da S. (2006). Significância de Efeitos Técnicos na Eficiência de Produção da Pesquisa Agropecuária. *Revista Brasileira de Economia do Rio de Janeiro*, v. 60, n. 1, p. 69–86.

Souza, U. R. de, Braga, M. J. e Ferreira, M. A. M. (2011). Fatores Associados à Eficiência Técnica e de Escala das Cooperativas Agropecuárias Paranaenses. *Revista de Economia e Sociologia Rural*, Piracicaba, SP, v. 49, n. 3, p. 573-598.

Unsal, G. M. e Orkcu, H. H. (2015). Ranking Decision Making Units with the Integration of the Multi-Dimensional Scaling Algorithm into PCA-DEA. *Hacettepe Journal of Mathematics and Statistics*.

Vaninsky, A. (2006). Efficiency of electric power generation in the United States: analysis and forecast based on data envelopment analysis. *Energy Economics*, v. 28, p. 326–338.

Vilela, D. L., Nagano, M. S. e Merlo, E. M. (2007). Aplicação da Análise Envoltória de Dados em Cooperativas de Crédito Rural. *Revista Administração Contemporânea*, 2ª edição especial.

Weyman-Jones, T. G., Fethi, M. D., Cooper, W. W., Avkiran, N. K., Morita, H., Paradi, J. C.; Meimand, M., Meng, W., Liu, W. e Lovell, C. A. K. (2012). Productivity Analysis in the Service Sector with Data Envelopment Analysis, Amazon Digital Services, Inc.

Zhu, J. (1998). Data envelopment analysis vs. principal component analysis: An illustrative study of economic performance of Chinese cities. *European Journal of Operational Research*, v. 1, p. 50-61.