

Mapeamento de estudos sobre a utilização de Análise por Envoltória de Dados (DEA) na análise de insolvência*

*Silvia Pereira de Castro Casa Nova e Luciana Massaro Onusic***

1 INTRODUÇÃO

As demonstrações contábeis divulgadas pelas empresas têm sido objeto de diversos estudos. Os indicadores contábeis de balanço são utilizados na interpretação dos fenômenos econômicos e financeiros das entidades. Mesmo tendo-se conhecimento das limitações dos números gerados a partir dos relatórios contábeis, os poderes investigativo e preditivo das análises são estudados e, por vezes, comprovados na literatura.

Construíram-se indicadores de previsão de insolvência, com base em instrumentos quantitativos sofisticados. Altman (1968) foi precursor da aplicação de métodos quantitativos em informações contábeis com a finalidade de desenvolver modelos de previsão de falência. Kanitz (1974) foi um dos pioneiros no Brasil a empregar recursos sofisticados no processo de análise de balanços. Posteriormente, outros pesquisadores brasileiros desenvolveram modelos utilizando-se de técnicas estatísticas de regressão logística, análise fatorial, análise discriminante, redes neurais e outras (ELIZABETSKY, 1976 apud SILVA, 2003; MATIAS, 1978; SILVA, 1982 apud SILVA, 2003; ALMEIDA, 2003; MATIAS; SIQUEIRA, 1996). Mais recentemente modelos de análise de insolvência têm sido desenvolvidos com utilização da Análise por Envoltória de Dados (DEA). Esta pesquisa tem como objetivo identificar o estágio atual do estudo da utilização da DEA na análise de falência, no exterior e no Brasil.

* Artigo apresentado originalmente no XXIX EnANPAD. As autoras agradecem o professor Abraham Laredo Sicsú por sua leitura e comentários.

** Silvia Pereira de Castro Casa Nova é pós-doutoranda na EAESP/FGV e professora doutora do Departamento de Contabilidade e Atuária da FEA/USP, São Paulo-SP, e-mail: silvianova@usp.br. Luciana Massaro Onusic é doutoranda em Administração na FEA/USP e professora mestre das Faculdades Sumaré, Cotia-SP, e-mail: lonusic@uol.com.br.

1.1 Histórico da Data Envelopment Analysis

A Análise por Envoltória de Dados originou-se com a tese de doutorado de Edwardo Rhodes, apresentada à *Carnegie Mellon University*, com orientação de W. W. Cooper, em 1978. O objetivo do trabalho era avaliar os resultados de um experimento social de larga escala em educação chamado *Program Follow Through* em escolas públicas americanas. Para tanto o desempenho das escolas participantes do programa (PFT) foi comparado com o de escolas não participantes (NPFT). O desempenho das escolas foi analisado através de uma medida de eficiência que considerava a relação entre resultados desejados (produtos) e recursos consumidos para obtê-los (insumos). No estudo os resultados foram representados por variáveis como, por exemplo, “auto-estima dos alunos” (medida através de testes psicológicos), “notas em leitura” e “notas em matemática”. As variáveis de recursos eram, por exemplo, “tempo gasto pelos pais com as crianças em exercícios relacionados às tarefas escolares” ou “número de professores”. Foi desenvolvido um problema de programação matemática para calcular a eficiência das escolas, considerando múltiplos resultados obtidos e múltiplos recursos utilizados, que redundou na formulação do Modelo CCR (abreviatura de CHARNES, COOPER e RHODES, sobrenomes de seus autores) de Análise Envoltória de Dados e na publicação do primeiro artigo sobre o assunto no *European Journal of Operations Research* em 1978. Desde o trabalho seminal de 1978, muitos artigos foram publicados sobre DEA. A pesquisa na base eletrônica PROQUEST, que reúne mais de 1.000 periódicos de língua inglesa nas áreas de administração, marketing e negócios, o termo Data Envelopment Analysis retorna o total de 787 artigos, desde 1986. Pesquisa semelhante na base EBSCO recuperou 1.811 referências.

1.2 Data Envelopment Analysis – Formulação

Para compreensão da DEA é preciso inicialmente apresentar a nomenclatura usualmente utilizada:

- *Decision Making Units* (DMU's): São as unidades sob avaliação. No estudo original eram escolas, mas há estudos sobre bancos, agências bancárias, programas de pós-graduação,

lojas, universidades, cursos MBA, pesquisadores em Engenharia de Produção no CNPq. É preciso que as unidades sejam homogêneas, ou seja, obtenham os mesmos resultados (produtos) utilizando os mesmo recursos (insumos) com variação apenas de quantidade ou intensidade.

- *Outputs* (Produtos): São os resultados obtidos pelas DMUs. São valores observados. Para universidades podem ser, por exemplo, número de graduados ou número de cursos. Para empresas, o lucro. Para pesquisadores, número de artigos publicados. Devem respeitar ao critério de quanto maior, melhor.
- *Inputs* (Insumos): São os recursos utilizados pelas DMUs para obter os resultados (produtos) desejados. No exemplo das universidades pode ser o número de professores e de funcionários. Para empresas, o Patrimônio Líquido. Para pesquisadores, número de anos de estudo. Deve obedecer ao critério de quanto menor, melhor.

A medida de eficiência calculada pela DEA é então uma generalização da medida de produtividade usual que é a razão (quociente) entre os resultados obtidos e os recursos utilizados, ou seja, para os exemplos citados:

$$\frac{\text{Número de alunos graduados}}{\text{Número de professores}} \quad \text{ou} \quad \frac{\text{Lucro}}{\text{Patrimônio Líquido}}$$

A vantagem da DEA é a possibilidade de relacionar múltiplos produtos e múltiplos insumos em uma medida singular de eficiência, com uma interpretação intuitiva, pois limitada entre 0 e 1 (ou 0 e 100% de eficiência).

A medida de eficiência calculada pela DEA é dada por:

- Seja h_k a eficiência da DMU k , onde K é o número de unidades sendo avaliadas.
- Seja u_j o coeficiente para o resultado j , com $j = 1, 2, \dots, M$, onde M é o número total de tipos de resultados sendo considerados. Também chamado de peso, u_j é uma medida da diminuição relativa da eficiência da unidade k com cada unidade de redução no valor do resultado j .

- Seja v_i o coeficiente para o recurso i , com $i = 1, 2, \dots, N$, onde N é o número total de tipos de recursos sendo considerados. A variável v_i mede o aumento relativo da eficiência com cada redução unitária do valor de recurso.
- Seja Y_{jk} o número observado de unidades do resultado j , gerado pela DMU k , durante um período de tempo.
- Seja X_{ik} o número observado de unidades do recurso i , consumido pela DMU k , durante um período de tempo.
- O conjunto observado de Y_{jk} e X_{ik} para as DMUs em avaliação será denominado de plano de produção.

O objetivo é encontrar o conjunto de pesos u associados com os resultados e de pesos v associados com os recursos que resultarão na máxima eficiência possível para a DMU sendo avaliada. Essa é outra vantagem da DEA: os pesos são atribuídos a cada variável pela análise de forma a colocar a unidade sob avaliação “sob a melhor luz”. Dessa forma é possível incluir um composto (*mix*) de resultados e de recursos mesmo em situações onde não haja pesos ou preços conhecidos.

O indicador de eficiência para uma determinada DMU e é definido por:

$$(1) \quad \text{Maximizar } h_e = \frac{\sum_{j=1}^M u_j y_{je}}{\sum_{i=1}^N v_i x_{ie}},$$

Sujeito à restrição de que, quando o mesmo conjunto de pesos de produtos e insumos (u e v) for aplicado às demais DMUs sob avaliação, nenhuma excederá 100% de eficiência ou uma razão de 1,0.

$$(2) \quad \frac{\sum_{j=1}^M u_j y_{jk}}{\sum_{i=1}^N v_i x_{ik}} \leq 1$$

Onde $k = 1, 2, \dots, K$ e todos os coeficientes u e v devem ser positivos e diferentes de zero.

Para resolver esse programa de programação linear fracionário utilizando um software padrão é necessária uma reformulação. A função objetivo e todas as restrições são razões e não funções lineares. A função objetivo é reformulada como uma função linear acrescentando-se uma restrição de que a soma ponderada dos recursos da unidade sob avaliação seja igual a 1. Assim, a função objetivo fica:

$$(3) \quad \text{Maximizar } h_e = \sum_{j=1}^M u_j y_{je},$$

Sujeito a:

$$(4) \quad \sum_{i=1}^N v_i x_{ie} = 1,$$

Para cada DMU sob avaliação a equação (2) é também reformulada para:

$$(5) \quad \sum_{j=1}^M u_j y_{jk} - \sum_{i=1}^N v_i x_{ik} \leq 0$$

Onde $k = 1, 2, \dots, K$ e todos os coeficientes u e v devem ser positivos e diferentes de zero.

Com base em um conjunto de DMUs e seu plano de produção realizado é possível construir uma curva formada pelas unidades eficientes (fronteira eficiente). Para tanto é necessário resolver o problema de programação linear (PPL) proposto para cada uma das unidades em avaliação, identificando aquelas cujo plano de produção, dados os pesos determinados para seus valores de produtos e insumos, não pode ser superado pelo plano de nenhuma outra unidade do grupo sob avaliação. Essas unidades são ditas eficientes e tornam-se referência para as demais, formando a fronteira

ou curva de eficiência. As unidades que se posicionam abaixo da fronteira são classificadas como ineficientes.

As unidades ineficientes podem ter seu plano de produção (valores de produtos e insumos) projetado na fronteira para determinar quais as alterações que as tornariam eficientes. A projeção pode considerar orientação ao produto ou orientação ao insumo. Na primeira são determinados os aumentos proporcionais nos valores de produtos, mantido o consumo de recursos constante, para que a unidade tomadora de decisão torne-se eficiente. Na orientação ao insumo, a DMU é projetada na fronteira pela redução proporcional no consumo de insumos, mantidos os valores de resultados (produtos) constantes.

Em 1984 BANKER, CHARNES e COOPER apresentaram um outro modelo DEA em artigo publicado na *Management Science*. O modelo foi chamado de BCC (abreviatura do sobrenome dos autores) e permite que as unidades avaliadas apresentem retornos variáveis de escala, ou seja, que o acréscimo em uma unidade de insumo pode gerar um acréscimo não proporcional no volume de produtos.

Segundo Belloni (2000, p. 68) “ao possibilitar que a tecnologia exiba propriedades de retornos à escala diferentes ao longo de sua fronteira, esse modelo admite que a produtividade máxima varie em função da escala de produção”. A formulação matemática do Modelo BCC é:

$$(6) \quad \sum_{j=1}^M u_j y_{jk} - \sum_{i=1}^N v_i x_{ik} \leq 0$$

Sujeito a:

$$(7) \quad \sum_{j=1}^M u_j y_{jk} - \sum_{i=1}^N v_i x_{ik} - u_k \leq 0$$

$$(8) \quad \sum_{i=1}^N v_i x_{ie} = 1$$

$$(9) \quad u_j, v_i \geq 0$$

y_{jk} = valor do produto j para DMU k

x_{ik} = valor do insumo i para a DMU k

u_j = peso calculado para o produto j

v_i = peso calculado para o insumo i

j = número de tipos de produtos = 1, ..., M

i = número de tipos de insumos = 1, ..., N

k = número de DMUs em avaliação = 1, ..., K

Foi introduzida uma variável uk representando os retornos variáveis de escala. Essa variável não precisa atender à restrição de positividade; pode, portanto, assumir valores negativos representando assim os retornos variáveis de escala. Pode identificar retornos crescentes ou decrescentes.

O restante desse trabalho está organizado como se segue: inicialmente são apresentados aspectos conceituais da pesquisa, incluindo a descrição das bases pesquisadas e da organização dos trabalhos recuperados; em seguida é apresentado um resumo dos trabalhos, ressaltando objetivo, metodologia, conclusões e limitações; finalmente as conclusões são apresentadas, com ênfase na discussão de possibilidades de pesquisas futuras.

2 MÉTODOS E TÉCNICAS

2.1 Bases de dados e periódicos incluídos

A definição do escopo da pesquisa considerou o objetivo de mapear os trabalhos que estudaram a utilização de DEA à análise de insolvência. As bases de pesquisa foram os bancos de dados EBSCO, PROQUEST, JStor, Emerald, Periódicos Capes e as revistas RAE, RAE eletrônica, RAUSP, RCFUSP, RAC, Gestão &

Produção e Pesquisa Operacional. Foram incluídos ainda bancos de teses e dissertações e trabalhos de congressos e seminários que tratassem do assunto.

Essa pesquisa secundária foi, no entanto, efetuada de forma não sistemática, havendo a possibilidade de algum artigo pertinente não ter sido incluído.

Por se tratar de um setor com características específicas, os trabalhos aplicados à previsão de insolvência de instituições financeiras não foram incluídos. Existe uma bibliografia relevante sobre o assunto no exterior e no Brasil, que se iniciou na década de 1990 (SIEMS, 1992; BARR; SIEMS, 1994, 1996; BERGER; HUMPHREY, 1997; SIEMS; CLARK, 1997; SIEMS; BARR, 1998; CERETTA; NIEDERAUER, 2001). Foram excluídos ainda os estudos de Troutt, Rai e Zhang (1996) e Pendhakar (2002) por serem propostas de modelos conceituais sem aplicação empírica para teste.

2.2 Definição do escopo e da organização do estudo

Os trabalhos foram recuperados e resumidos. Os focos de atenção foram a metodologia, aplicação, validação e conclusões. As fórmulas serão apresentadas quando os autores propuserem alterações aos modelos originais ou novos algoritmos. Se forem utilizados os modelos originais tal fato será apenas indicado, referenciando-se o modelo conforme abreviatura apresentada na Introdução.

3 ANÁLISE DOS TRABALHOS

3.1 Estudos de análise de insolvência utilizando DEA no exterior

O primeiro estudo utilizando DEA na análise de insolvência foi apresentado por Fernandez-Castro e Smith em 1994, na busca de um modelo não paramétrico para análise do desempenho empresarial. Os quatro problemas-chave relacionados ao uso dos indicadores nas análises de desempenho foram relacionados pelos autores: a escolha das empresas para comparar com as empresas sob exame; a escolha dos indicadores; como os indicadores seriam combinados em uma medida singular de desempenho; e a escolha

de uma meta (*benchmark*) com a qual comparar o desempenho de uma companhia.

Para superar essas dificuldades, os autores examinaram a utilização de um modelo não-paramétrico de análise de indicadores “*que fornece uma medida única de eficiência mesmo quando lida com múltiplos insumos e produtos e torna desnecessária a atribuição de pesos pré-especificados aos mesmos*” (FERNANDEZ-CASTRO; SMITH, 1994, p. 240).

O modelo proposto considerou os indicadores contábeis como produtos, sendo necessário apenas que fossem válidos e respeitassem ao critério quanto maior, melhor, pois se as empresas estiverem operando em condições de mercado similares, então os insumos podem ser considerados imateriais, e admitidos como iguais para todas as empresas. O objetivo da análise foi encontrar as companhias que tivessem, em algum sentido, os melhores indicadores financeiros entre as empresas observadas. O problema linear, considerando N empresas e m indicadores financeiros r_i , foi dado pelas equações 10 a 14:

$$(10) \quad \text{Maximizar } z_0,$$

Sujeito a:

$$(11) \quad \sum_{n=1}^N \lambda_n r_{in} \geq z_0 r_{i0}$$

$$(12) \quad \sum_{n=1}^N \lambda_n = 1$$

$$(13) \quad \lambda_n \geq 0$$

$$(14) \quad z_n \geq 0$$

$r = \text{indicadores financeiros}$

$z =$ incremento proporcional nos indicadores

$\lambda =$ peso da empresa na formação da fronteira para a empresa O

$$i = 1, \dots, m; \quad n = 1, \dots, N$$

As variáveis de decisão do programa linear eram: z_0 – proporção em que os indicadores poderiam ser incrementados, e; λ – peso de cada empresa na formação da fronteira eficiente para a empresa O . O escore de eficiência da empresa O era dado por $1/z_0$.

O modelo foi implementado com a inclusão de seis indicadores financeiros: Caixa/Ativo Total; Ativo Circulante/Passivo Circulante; Capital de Giro/Ativo Total; Passivo Exigível a Longo Prazo/Ativo Total; Lucro Líquido/Ativo Total; Vendas/Ativo Total. Esses indicadores refletiam, respectivamente: Posição de caixa; Liquidez; Posição de Capital de Giro; Alavancagem; Lucratividade; e Giro do Ativo. Uma boa classificação na maioria dos indicadores refletiria uma posição financeira forte.

O modelo foi aplicado aos dados contábeis de 27 setores, com o total de 1.423 empresas, sendo o número mínimo de empresas por setor 7 e o máximo, 136. Para cada um dos setores a amostra incluía uma empresa que faliu no período de janeiro de 1985 a dezembro de 1989 com dados contábeis do ano precedente à falência para essa empresa e para outras empresas do mesmo setor. Os indicadores que assumiram valores negativos foram tratados como produtos fixos, incluindo ao modelo a seguinte restrição:

$$(15) \quad \sum_{n=1}^N \lambda_n r_{in} \geq r_{i0}$$

Os resultados iniciais mostraram que a medida de eficiência das empresas falidas dependia do número de empresas em seu grupo: claramente em setores com números maiores de empresas aumentava a probabilidade de se encontrar empresas eficientes, o que implicava em uma medida de eficiência mais baixa para a empresa falida. Cinco empresas falidas foram classificadas como eficientes.

As razões do erro de classificação foram investigadas. Duas empresas falidas classificadas como eficientes pertenciam a setores com número pequeno de empresas (7 e 15 empresas). Para as outras três empresas foram analisados os pesos atribuídos aos indicadores. Duas empresas atribuíram maior peso ao indicador Alavancagem. A última empresa atribuía maior peso ao indicador Capital de Giro. Para as demais empresas falidas consideradas no estudo, o escore de eficiência variou de 0.14 a 0.96, com mediana de 0.63. A maioria das empresas falidas encontrava-se na metade inferior dos escores de eficiência entre seus pares do mesmo setor. Entretanto, com exceção de apenas dois setores, havia empresas sobreviventes com escores de eficiência inferiores aos das empresas falidas.

Os autores concluíram, examinando os resultados, que a Alavancagem era o indicador que mais freqüentemente impedia que fosse atribuído um escore de eficiência mais baixo para as empresas falidas. Assim, o modelo preliminar sugeria que poderia ser um indicador inadequado, por ter sido determinante na atribuição de um escore de eficiência mais alto em um percentual elevado de empresas falidas.

Os escores de eficiência foram recalculados com sua eliminação. Os escores de eficiência das empresas falidas se reduziram assim como sua posição em relação a seus pares no setor. Outra conclusão interessante foi sobre o indicador do Capital de Giro. O modelo atribuir peso zero em uma variável implicava em folgas diferentes de zero na desigualdade da Formulação 10 e indicava uma *performance* particularmente ruim em um indicador relevante. Os resultados mostraram a preponderância de folgas estritamente positivas (pesos zero) para o indicador Capital de Giro entre as empresas falidas, mostrando que era particularmente útil para prever dificuldades financeiras. Notou-se também que o Giro assumia posição determinante no escore de eficiência, sendo significativo para 20 das 27 empresas.

As conclusões dos autores foram de que o modelo descrito no estudo (FERNANDEZ-CASTRO; SMITH, 1994, p. 247):

é uma solução intermediária entre a simples análise de indicadores e as complexas técnicas de regressão. Oferece a oportunidade de examinar simultaneamente a performance em numerosas dimensões sem exigir especificações estatísticas detalhadas. A falta de uma base estatística

significa que o usuário deva usar julgamento em selecionar os indicadores adequados, uma vez que não há teste de significância disponível. Entretanto, o modelo pode ser prontamente aplicado utilizando um aplicativo padrão DEA e é simples testar alternativas.

Os autores afirmaram que o modelo poderia ser usado para explorar a performance de empresas sob eminência de falir, com sucesso moderado em isolar as empresas falidas das sobreviventes e, ainda assim, se empregado em conjunto com outra ferramenta de análise. O modelo permitiu ainda importantes conclusões sobre o desempenho relativo da empresa e sobre os indicadores determinantes de sua performance. Sobretudo poderia ser utilizado para definir metas e utilizado como uma ferramenta normativa.

As maiores fraquezas do enfoque, segundo os autores, foram sua suscetibilidade aos erros (ou flutuações aleatórias), sua inabilidade para capturar considerações de longo prazo e em fazer julgamentos úteis nas observações não usuais (*outliers*). Foi sugerido que pesquisas futuras fossem aplicadas aos trabalhos intertemporais mais recentes em DEA para verificar se é possível traçar uma posição relativa da empresa no tempo (FERNANDEZ-CASTRO; SMITH, 1994, p. 247). Os pontos fortes do modelo foram: a liberdade de suposições sobre a distribuição; a flexibilidade; e, a independência de qualquer tipo de forma funcional.

Emel *et. al.* (2003) aplicaram uma metodologia baseada em DEA em dados contábeis de 82 indústrias que compunham a carteira de crédito de um dos maiores bancos comerciais da Turquia. Utilizando indicadores financeiros, a análise DEA sintetizou a performance total em único score financeiro de eficiência, batizado pelos autores de score de credibilidade. Entre as vantagens da utilização da DEA foram citadas: (1) conversão de uma multiplicidade de insumos e produtos em um indicador singular de performance adimensional; (2) capacidade de lidar com múltiplos insumos e múltiplos produtos; (3) definição de uma fronteira composta pelas empresas de melhores práticas (fronteira eficiente) para um dado conjunto; (4) mensuração do grau de eficiência relativa para as demais empresas (não eficientes).

A metodologia proposta era composta por sete etapas consecutivas, a saber: Seleção da amostra; Identificação das principais dimensões financeiras; Identificação dos indicadores financeiros; Obtenção das principais componentes financeiras via Análise Fatorial; Seleção final dos indicadores financeiros; Cálculo dos Indicadores de Credibilidade via *Data Envelopment Analysis*; Validação via Análise de Regressão, Análise Discriminante e Análise Julgamental.

Foram obtidas informações financeiras de 100 empresas industriais de setores como produtos florestais, couro, componentes elétricos, alimentos, papel, química, maquinaria, metal, não-metal, plásticos, têxtil e transportes. Os *outliers* (pontos extremos), ou seja, empresas que tinham alguns indicadores financeiros com distância de mais de dois desvios-padrões dos demais foram eliminadas da amostra, resultando em 82 empresas. Os indicadores a serem incluídos no estudo foram selecionados contemplando as dimensões de Liquidez, Atividade, Estrutura Financeira, Lucratividade, Crescimento e Fluxo de Fundos. A experiência dos analistas de crédito do banco foi utilizada na sugestão de indicadores que considerassem a estrutura financeira multidimensional das empresas, resultando em 46 indicadores. Os indicadores foram submetidos a uma Análise Fatorial, com utilização do método de componentes principais e rotação Varimax, sendo mantidos 42 indicadores, agrupados em 11 fatores. Sete fatores representavam características financeiras similares, a saber: Empréstimos bancários; Lucratividade; Alavancagem; Estrutura de Passivos; Liquidez; e, Vendas e Custos.

Com base na experiência dos analistas, análise da literatura e julgamento dos autores as seguintes variáveis foram selecionadas:

Produtos: (*Current assets – inventories*)/*current liabilities* (LR), *Owners' equity/total assets* (OE/TA), *Net profit/total assets* (NP/TA);

Insumos: *Short term bank loans/current liabilities* (STBL/CL), *Current liabilities/net sales* (CL/NS), $1 - (\text{Fixed assets}/\text{owners' equity})$ (ABS).

Foi utilizado o modelo CCR com orientação para insumo. Os indicadores de credibilidade obtidos variaram entre 100 e 2,72. A fronteira eficiente foi formada por 16 empresas. O modelo foi validado considerando:

- **Análise de Regressão:** O indicador de credibilidade foi considerado como variável dependente e os seis indicadores contábeis, como variáveis independentes. As empresas consideradas eficientes (escore de credibilidade = 100) foram excluídas. O modelo de regressão obtido teve estatística F significativa ($4,9 \times 10^{-18}$) e alto poder explanatório ($R^2=78,7\%$). Os coeficientes das variáveis explicativas tinham os sinais esperados e foram também significativos estatisticamente a 5%, com exceção de NP/TA. Portanto, os resultados sugeriam que os escores DEA explicavam com sucesso a variação em cinco dos seis indicadores. Segundo os autores, se a amostra fosse estatisticamente representativa, a equação de regressão poderia ser utilizada para avaliar uma nova solicitação de crédito, sem necessidade de repetir a aplicação DEA.
- **Análise Discriminante:** As empresas foram então classificadas em dois grupos com base nos escores da DEA. O ponto de corte entre empresas “boas” e “ruins” foi definido de forma subjetiva, considerando a distribuição dos escores – 50% das empresas foram consideradas como “boas”. A análise discriminante foi aplicada considerando essas categorias, como variável dependente, e os seis indicadores financeiros, como variáveis independentes. A função discriminante incluiu cinco dos seis indicadores (NP/TA foi excluído), com uma taxa de acerto em relação à classificação DEA de 91,5%. Segundo os autores ficou demonstrado que os resultados da análise DEA podem indubitavelmente ser linearmente aproximados aos resultados obtidos pela Análise Discriminante, uma técnica mais comum.
- O índice de lucratividade NP/TA foi o único – entre os seis considerados no modelo DEA – que não foi estatisticamente significativo na Análise de Regressão e na Análise Discriminante. Assim, de acordo com os autores, esse não seria um critério discriminador entre as empresas “boas” e “ruins”, ou seja, não se poderia afirmar que uma empresa seja “boa” por ser lucrativa ou “ruim” por não ser total ou marginalmente lucrativa. Esse resultado seria contrário à intuição e poderia

ser decorrência de características das empresas que atuavam na Turquia. Em um mercado onde a eficiência fosse maior, a lucratividade teria certamente maior poder explicativo.

- **Análise Julgamental:** A consistência dos resultados da DEA foi confrontada com os pareceres dos analistas de crédito. A taxa de concordância foi de 78%. Assim, em certo grau, os resultados do modelo DEA não estavam distantes das opiniões dos especialistas, apesar de não considerar informações *a priori*, enquanto a análise julgamental é uma mistura de informações *ex post*, incluindo aspectos financeiros e não financeiros.

Para permitir maiores interpretações do significado de escores da DEA para uma empresa determinada, os autores calcularam 46 indicadores financeiros para as 16 empresas que formavam a fronteira eficiente e para as 16 empresas que apresentaram os escores DEA mais baixos (variando de 2,7% a 11%). A estatística t foi então computada com os seguintes resultados: (a) a utilização de empréstimos bancários de curto prazo é inversamente relacionada à eficiência financeira; (b) foram detectados graus de alavancagem menores e maior adequação do capital nas melhores empresas; (c) as melhores empresas apresentavam maiores índices de liquidez; (d) as melhores empresas eram mais lucrativas. Essas conclusões, na opinião dos autores, tinham grande relevância para considerações sobre risco pelos bancos. Assim, os escores de eficiência estavam alinhados com o senso comum. Mais do que isso, os resultados mostraram que os escores DEA maiores eram atribuídos às melhores empresas em termos das dimensões financeiras de Liquidez, Adequação de capital e Lucratividade.

As conclusões foram de que, apesar de os resultados do modelo DEA serem consistentes com as análises tradicionais, ofereceram indicações sobre como as empresas em má situação financeira poderiam melhorar sua eficiência financeira. O estudo apresentado considerou empresas de diversos setores. Na opinião dos autores, para desenvolver um modelo mais consistente, a amostra poderia se limitar a um setor específico. Deveria se ter em mente, no entanto, que seria perdida a possibilidade de comparar os escores de

eficiência entre empresas de diferentes setores. Finalmente, a metodologia apresentada permitiria ao banco monitorar a exposição de sua carteira de crédito. Seria ainda outra ferramenta administrativa para auxiliar em decisões estratégicas e táticas. Em países sujeitos à regulação ou controle governamental, a análise DEA poderia servir como um instrumento de monitoramento das regras de conformidade, inclusive com propósitos consultivos.

Cielen, Peeters e Vanhoof compararam, em estudo de 2004, o desempenho de modelos de programação linear, DEA e *rule induction* (C5.0) em previsão de falência, concluindo que, em termos de acurácia, o modelo DEA superou os demais.

O estudo considerou uma amostra de 367 empresas, para os anos de 1994, 1995 e 1996, sendo 276 empresas em continuidade e 90 empresas falidas. Além das empresas do setor financeiro, a seleção excluiu as empresas que não apresentaram os relatórios anuais completos. Foram também retirados da amostra os *outliers*, isto é, empresas com valores extremos para alguns dos indicadores considerados e que, na opinião dos autores, poderiam influenciar de forma negativa os resultados.

Foram selecionados 11 indicadores financeiros para construção dos modelos, entre os considerados em estudos anteriores de previsão de falência: *Equity/total assets*; *Retained earnings/total assets*; *Tax and social security charges/short term debt*; *Cash/restricted current assets*; *Work in progress, finished goods and contracts in progress/ current assets*; *Amounts payable within one year at the credit institutions/short term debt*; *Operating cash flow before taxes/total assets*; *Cash flow before dividends/total debts*; *Operating profit (loss) after depreciation before financial charges and taxes/total assets*; *Current assets/short term debts*; *Amounts receivable within one year + investments + cash/amounts payable within one year*.

O modelo de programação linear tinha a seguinte formulação: Sejam $n + m$ observações com k variáveis independentes onde n é o número de observações que pertencem ao grupo 1 e m é o número de observações que pertencem ao grupo 2. Seja A uma matriz ($n \times k$) que representa as observações do grupo 1 e B similarmente para o grupo 2.

$$(16) \quad \text{Minimizar } z = C_1 d_1 + C_2 d_2,$$

Sujeito a

$$(17) \quad Ax - d_1 \leq b$$

$$(18) \quad Bx + d_2 \geq b$$

$$(19) \quad d_1, d_2 \geq 0$$

C_1 e C_2 são coeficientes de custo dos erros de classificação
 x e b são livres

Na formulação acima, X era um vetor de variáveis de decisão e b era uma variável real; d_1 e d_2 foram definidos como os vetores de desvios externos que mediam a distância das observações situadas no lado errado, em relação ao hiperplano. O objetivo do modelo era minimizar a sobreposição total entre os grupos, isto é, a soma dos desvios externos das observações classificadas erroneamente.

Os pontos fracos do modelo proposto são: minimizava a soma das distâncias, quando o objetivo é minimizar o custo total das classificações errôneas; não havia consistência entre segmentos diferentes, sendo que, para fatores de custo diferentes, os coeficientes obtidos variavam muito, aconteciam inversões de sinais e o modelo se tornava insolúvel.

Os autores sugeriram que essas falhas poderiam ser resolvidas pelo emprego de um modelo DEA. Os indicadores com correlação positiva seriam definidos como insumos; aqueles com correlação negativa, como produtos. Essas definições seriam as mesmas para todos os segmentos. A formulação do modelo era a seguinte:

$$(20) \quad \text{Minimizar } z = 1C_1 f_1 + 1C_2 f_2,$$

Sujeito a

$$(21) \quad A' y - A'' x - d_1 \leq b$$

$$(22) \quad B' y - B'' x + d_2 \geq b$$

$$(23) \quad d_1 \leq Mf_1$$

$$(24) \quad d_2 \leq Mf_2$$

$$(25) \quad y, x, d_1, d_2 \geq 0$$

f_1 e f_2 são vetores binários

y , x e b são livres

M é um número muito grande

C_1 e C_2 são coeficientes de custo

O terceiro modelo foi desenvolvido a partir de uma técnica não paramétrica baseada em regras ou árvores de decisão, chamada *machine learning* (aprendizado de máquina). As árvores de decisão ou as regras classificavam observações identificadas por variáveis ou atributos; segmentavam sucessivamente regiões do espaço em sub-regiões de acordo com a variável mais informativa. Para especificar o critério para seleção da variável mais explicativa utilizou-se o critério da taxa de ganho informacional, que estabelecia que a informação sobre um evento fosse inversamente proporcional a sua probabilidade e poderia ser mensurado (em *bits*) como menos o logaritmo base 2 daquela probabilidade.

Os resultados de cada modelo foram comparados, em termos de sua acurácia (total, Erro Tipo I e Erro Tipo II) obtida por *cross-validation*, com os resultados do modelo DEA superando os demais. Foram então incorporados aos modelos diferentes custos para os erros Tipo I e II. O modelo DEA mostrou-se mais consistente e menos sensível à mudança dos fatores de custo. Para o modelo de árvore de decisão os diferentes fatores de custo tiveram impacto direto na seleção do atributo mais informativo e diferentes árvores foram construídas, produzindo distribuições

de taxas de acerto mais equânimes. No entanto, o modelo DEA obtido quando considerados os fatores de custo mais baixos foi claramente superior.

Outra característica analisada pelos autores foi a facilidade de compreensão dos modelos. Inicialmente, o modelo de árvore de decisão era mais amigável dado seu resultado ‘sim’ ou ‘não’. Os resultados do modelo DEA foram escores abstratos que eram comparados com um limite que se alterava com as mudanças nos fatores de custo e a proporção entre as populações. Porém, a introdução de fatores de custo para o modelo DEA alterava apenas o valor limite. Para o modelo de árvore de decisão, a cada vez que os coeficientes de custo eram alterados, novas árvores, com outra estrutura e outras variáveis eram obtidas. Portanto, em termos de estabilidade o modelo DEA mostrou-se superior.

Outro fator considerado foi a facilidade de comunicação dos resultados ao usuário final. No modelo DEA havia uma ligação direta entre as variáveis e os coeficientes utilizados, de um lado, e os resultados, de outro. No modelo baseado em árvore de decisão, como uma mesma variável poderia ser utilizada diversas vezes, não se estabeleceria um relacionamento direto entre as variáveis e os resultados. Os testes de acurácia puderam somente ser efetuados por ramo e não por variável. Consequentemente foi muito difícil avaliar o impacto das variáveis.

A conclusão final dos autores foi de que, para uma amostra pequena com informação quantitativa apenas, o modelo DEA superou os demais em termos de acurácia, custo, desenvolvimento e compreensibilidade.

Paradi, Asmild e Simak apresentaram em 2004 o conceito de fronteira DEA de pior prática (*worst practice* DEA), com objetivo de identificar as empresas de pior desempenho e posicioná-las na fronteira. Na opinião dos autores o conceito era particularmente relevante para a aplicação que desenvolveram em avaliação de risco de crédito. O estudo apresentou ainda a técnica de camadas (*layering technique*) como alternativa ao enfoque do ponto de corte.

Segundo os autores, os modelos DEA normais classificam as empresas na fronteira como unidades eficientes e as “envolvidas”

pela fronteira como ineficientes. O modelo DEA de piores práticas utilizaria a mesma formulação mas, ao invés de apontar as empresas de bom desempenho, seu objetivo seria identificar as empresas de desempenho ruim. Isso seria conseguido selecionando variáveis que refletissem falhas de desempenho, ou seja, a estratégia seria selecionar variáveis de produtos que refletissem uma má utilização dos recursos ou resultados indesejáveis como, por exemplo, capital de giro, reclamações em garantia, maus devedores etc.

Os insumos seriam selecionados de forma semelhante podendo ser variáveis como: lucro, vendas, índices de satisfação do cliente etc. As empresas que comporiam a fronteira eficiente seriam, então, aquelas empresas com menores níveis de (bons) insumos e maiores níveis de (maus) produtos. A comparação dos resultados obtidos com a aplicação dos modelos DEA normal e de piores práticas permitiria, ainda, diferenciar as empresas, como apresentado no Quadro 1:

		DEA Piores Práticas	
		Eficiente	Ineficiente
Eficiente	<i>Outliers</i> – não é possível concluir	Empresas de baixo risco	
Ineficiente	Empresas de alto risco	Empresas de risco intermediário – análises adicionais	

Quadro 1 - Cruzamento das análises dos Modelos DEA Normal e de Piores Práticas
 Fonte: Paradi, Asmild e Simak (2004)

Essas informações poderiam ser utilizadas para determinar a que empresas conceder crédito e a que preço. A metodologia foi aplicada a uma amostra de empresas do setor de manufatura, com ativos entre \$10M e \$500M, como forma de aumentar a comparabilidade. Foram selecionadas 17 e 11 empresas que faliram nos anos de 1996 e 1997. O grupo de controle foi composto por empresas selecionadas com bases nos mesmos critérios que não solicitaram falência até 1998: eram 160 e 115 empresas, para 1996 e 1997, respectivamente. Foram consideradas as informações financeiras de um ano antes da falência. O modelo foi construído utilizando os dados de 1995 e a amostra de teste foi composta pelos dados de

1996. Para seleção de variáveis foram identificados os principais estudos sobre os indicadores de saúde corporativa. Foram então comparados os resultados do cálculo da correlação somente entre as variáveis das empresas saudáveis e para a amostra total (falidas e não falidas). A maioria das variáveis que divergiam entre as duas amostras coincidiu com as apontadas na literatura. Foram então selecionadas: *Total assets; Working capital; Earnings before income, tax, depreciation and amortization; Retained earnings; Shareholders equity; Interest expense; Cash flow from operations; Stability earnings; Total liabilities*. As variáveis com valores negativos foram transformadas com base na propriedade de *translation invariance* dos modelos DEA. Diversos modelos foram construídos e testados com combinações dessas variáveis e seus resultados comparados. Os modelos DEA utilizados foram BCC com orientação ao insumo (DEA Normal) e BCC com orientação ao produto (DEA Piores Práticas). Foram comparados: eficiência média das empresas falidas e não-falidas; ponto de corte (considerando-se valores de custo para os erros Tipo I e II de 0,6 e 0,03); taxa de acerto de classificação para empresas falidas e não-falidas.

Constatou-se que os escores de eficiência das empresas falidas eram significativamente mais baixos do que os escores das empresas não-falidas. O ponto de corte, considerando os fatores de custo para cada tipo de erro, privilegiou a classificação correta de empresas falidas, em detrimento das empresas não-falidas. As taxas de acerto variaram entre 94/100 e 81/64, para os modelos normais e 94/100 e 41/21 para os modelos de piores práticas, para os erros Tipo I e Tipo II, respectivamente.

A técnica de camadas foi então empregada para os modelos de piores práticas: a cada etapa, os indicadores de eficiência eram calculados, as empresas eficientes eram identificadas e removidas e o modelo novamente empregado, até todas as empresas falidas terem sido identificadas e removidas. O modelo com melhores resultados teve uma taxa de acerto de 100% das empresas falidas e 61% das empresas não-falidas, na terceira camada. A técnica de camada mostrou ser um enfoque mais flexível para classificação, pois poderia englobar considerações subjetivas, julgamento gerencial, atitude

frente ao risco entre outros aspectos. Quanto mais avessa ao risco a instituição fosse, mais camadas poderia considerar, de forma a eliminar todas as empresas que representassem risco. Além disso, a técnica de camadas poderia ser utilizada para decisões de preço baseadas em risco, uma vez que o risco associado varia de camada para camada, e isto poderia ser refletido no preço.

Por fim, a técnica de camadas foi aplicada cruzando-se os resultados dos modelos normal e de piores práticas. Os melhores resultados foram obtidos combinando as terceiras camadas do modelo normal e de piores práticas, com uma classificação de 100% das empresas falidas e de 78% das empresas não-falidas. Os resultados foram validados utilizando-se os dados de 1996, com taxa de acerto, para a terceira camada, de 100% para empresas falidas e 67% para empresas não-falidas.

As conclusões foram de que a introdução dos modelos de piores práticas em combinação com a técnica de camadas, conduziu a taxas de acertos maiores e a resultados mais influenciados pela amostra, do que o método do ponto de corte. O enfoque de camadas implicou ainda em maior flexibilidade, permitindo incorporar considerações de atitude frente ao risco e estratégias de preço baseados no risco.

3.2 Estudos utilizando DEA em análise de insolvência no Brasil

Almeida e Milioni apresentaram em 2000 um modelo de *credit scoring* baseado em DEA. A proposta foi construir fronteiras de solvência e insolvência a partir de tomadores de crédito que, no passado, se revelaram bons e maus pagadores.

O modelo foi estruturado nas seguintes etapas: (1) Determinação das variáveis de *input* e *output*; (2) Determinação da fronteira de insolvência: foi desenvolvido um modelo DEA de maximização para ser aplicado aos dados das empresas insolventes, previamente selecionadas pelos autores, e definir uma fronteira DEA superior. A eficiência igual a 1 foi o limite superior destas firmas; (3) Determinação da fronteira de solvência: foi utilizado um modelo DEA de minimização para, com base nas informações das empresas solventes, definir uma fronteira DEA inferior para as empresas solventes.

A eficiência igual a 1 foi o limite inferior das empresas solventes; (4) Unificação dos escores: com a utilização simultânea das duas fronteiras envoltórias foram definidas quatro regiões.

A situação das empresas, conforme a região em que se encontrava, era classificada como: Região I: apenas *Empresas solventes*; Região II: apenas *Empresas insolventes*; Região III: continha *Empresas solventes e insolventes*; Região IV: não fornecia informações sobre a situação de solvência das empresas.

As formulações dos modelos DEA de maximização e minimização foram:

Modelo de Maximização (adotado para as empresas insolventes)

$$(26) \text{ Maximizar } \varepsilon_j = \left(\sum A_{js} X_{js} \right) / \left(\sum B_{jk} W_{jk} \right), \text{ para a unidade } j$$

Sujeito a

$$(27) \varepsilon_j = \left(\sum A_{js} X_{js} \right) / \left(\sum B_{jk} W_{jk} \right) \leq 1$$

$$(28) X_{js}, W_{jk} \geq 0$$

Eficiência = ε = Saídas/Entradas

A_{js} = Saída do insumo A_s da unidade j

B_{jk} = Entrada do insumo B_k da unidade j

X_{js} = Peso do insumo A_s

W_{jk} = Peso do insumo B_k

$S=1,2,\dots$ S saídas; $k=1,2,\dots$ K entradas

Modelo de Minimização (adotado para as empresas solventes)

$$(29) \text{ Maximizar } \varepsilon_j = \left(\sum A_{js} X_{js} \right) / \left(\sum B_{jk} W_{jk} \right), \text{ para a unidade } j$$

Sujeito a

$$(30) \varepsilon_j = \left(\sum A_{js} X_{js} \right) / \left(\sum B_{jk} W_{jk} \right) \leq 1$$

$$(31) X_{js}, W_{jk} \geq 0$$

Eficiência = ε = Saídas/Entradas

A_{js} = Saída do insumo A_s da unidade j

B_{jk} = Entrada do insumo B_k da unidade j

X_{js} = Peso do insumo A_s

W_{jk} = Peso do insumo B_k

Os autores ressaltaram “*não ser procedimento usual estabelecer um objetivo de minimizar a relação saídas/entradas tendo como restrição que sejam todas maiores ou iguais a 1 em DEA. No entanto, estamos procurando a envoltória de mínima eficiência na amostra de empresas solventes*”.

A base de dados utilizada era composta por 60 empresas (28 insolventes e 32 solventes) para desenvolver o modelo e 39 empresas para validá-lo. As variáveis consideradas foram: *Produto*: Retorno sobre o Ativo Total; Giro do Ativo Total; *Insumo*: Endividamento total.

O modelo de maximização foi aplicado às empresas insolventes (28 empresas) e a envoltória de insolvência foi determinada. Cinco empresas apresentaram eficiência igual a 1, ou seja, encontravam-se sobre a fronteira. Assim, o modelo para teste exclusivo de insolvência em novas empresas a serem avaliadas (PPL IN) incluía essas cinco empresas na determinação das restrições. Segundo os autores, caso os dados relativos à empresa sob avaliação fossem tais que apresentasse eficiência menor do que 1 seria indicada como candidata a ser classificada como empresa insolvente.

O modelo de minimização foi aplicado às empresas solventes (32 empresas) e a envoltória de solvência foi determinada. Duas empresas apresentaram eficiência igual a 1, ou seja, pertenciam à fronteira. Portanto, o modelo para teste exclusivo de solvência em novas empresas a serem avaliadas (PPL SOL) considerava essas duas empresas na determinação das restrições. Se os dados relativos à empresa sob avaliação fossem tais que apresentasse eficiência maior do que 1 seria indicada como candidata a ser classificada como empresa solvente.

A análise cruzada do escore de empresas sob avaliação em relação às regiões definidas, permitiria sua classificação em situações distintas, indicadas no Quadro 2.

		PPL SOL	
		>1	=1
PLL IN	=1	Região I – Empresas solventes	Região IV – Dados inconclusivos
	>1	Região III – Região de dúvida	Região II – Empresa insolvente

Quadro 2 - Situações em relação ao cruzamento PPL SOL x PPL IN
Fonte: Almeida e Milioni (2000)

A análise proposta aplicada às informações das 60 empresas da amostra de calibração resultou em 82% de taxa de acerto (49 empresas classificadas corretamente). Apenas uma empresa foi classificada na região IV, com dados inconclusivos para decisão. A análise conduziu a 90% de taxa de acerto na amostra de verificação.

Foi proposto ainda um procedimento de aperfeiçoamento dinâmico do modelo. Os autores consideraram as informações de uma empresa solvente da amostra de verificação que foi classificada erroneamente como insolvente pelo modelo. Para corrigir o modelo, propuseram que as informações dessa empresa fossem incluídas nas restrições do PPL SOL, realocando a envoltória. Ressalte-se, no entanto, que antes do novo modelo ser utilizado foi necessário confirmar que as outras empresas originalmente definidoras da fronteira continuavam classificadas como eficientes após a inclusão. Esse procedimento implicou na melhora das

taxas de acerto para 83% e 97%, nas amostras de calibração e verificação, respectivamente.

Em 2004 o estudo de ONUSIC propôs a utilização conjunta das técnicas de Análise por Envoltória de Dados e Regressão Logística como forma de minimizar os erros de classificação das empresas solventes e insolventes em uma determinada base de dados.

Para isto foram aplicados modelos de previsão de insolvência utilizando Regressão Logística (RL) e a Análise por Envoltória de Dados em uma base de dados fornecida pela SERASA relativa aos períodos de 1995 a 2001, composta por 122 empresas (61 solventes e 61 insolventes, pareadas em porte e setor às empresas solventes). As informações da amostra correspondiam a datas de balanço de 3 anos antes da concordata e compreendiam o período entre 1995 e 2001 (p. ex. se a concordata ocorreu em 2001, foram utilizadas informações contábeis de 3 anos antes, ou seja, de 1998, e de outra empresa pareada, da mesma data). Não houve diferenciação de forma jurídica (sociedade por ações, de capital aberto ou fechado, e sociedade limitada), a moeda utilizada foi o Real em valores históricos, da data da publicação do balanço.

As etapas de análise previstas foram: análise exploratória dos dados, com eliminação de outliers; seleção das variáveis pela Análise de Regressão Logística; estudo exploratório da utilização conjunta de Análise de Regressão Logística e DEA; análise de resultados das aplicações, com utilização da correlação de Spearman para análise dos erros de classificação; e, benchmarking, com análise das empresas em que a classificação foi errônea.

Para construção e validação do modelo, a amostra inicial foi subdividida em três aplicações A, B e C, com distribuição do percentual de empresas de 70%/30%, 60%/40% e 50%/50%, respectivamente. As empresas foram selecionadas aleatoriamente para compor cada uma dessas “sub-amostras”.

Na RL as variáveis incluídas foram: Variável dependente – Situação da empresa (solvente ou insolvente); Variáveis independentes – Margem, Endividamento Geral, Endividamento de Curto Prazo. As probabilidades (escores) fornecidas pela RL variaram de 0 a 1. O ponto de corte foi determinado como 0,5. Este ponto de corte foi

escolhido, segundo a autora, de acordo com a proporção da amostra com riscos iguais (50% solvente e 50% insolvente) para a RL (HAIR et al, 1998 apud ONUSIC, 2004).

As mesmas variáveis utilizadas na RL foram consideradas para a DEA, ou seja: Produtos: Margem; Insumo: Endividamento Geral e Endividamento de Curto Prazo. Após o cálculo do escore de eficiência das empresas, foi determinado o ponto de corte que minimizava o erro de classificação das empresas (SIMAK, 1997 apud ONUSIC, 2004). Foi escolhido o modelo DEA com orientação ao insumo e retornos variáveis de escala (BCC) devido à necessidade de efetuar uma transformação de escala na variável Margem, que assumia valores negativos para algumas empresas. Para comparação dos resultados obtidos entre a RL e a DEA foram analisadas apenas as empresas que compunham as amostras de validação, conforme definidas em cada aplicação A, B e C.

O critério proposto para classificação conjunta das empresas foi que, se em qualquer técnica (RL ou DEA) a empresa fosse classificada como insolvente, na análise conjunta a empresa seria considerada insolvente. Em todas as aplicações (A, B e C) o desempenho da DEA foi melhor ao desempenho da RL, com exceção da aplicação C, onde os erros foram iguais (30%). Os resultados em cada aplicação foram resumidos na Tabela 1.

Tabela 1 - Resultados das aplicações

Aplicação	Amostra	Análise	Erro I	Erro II	Acerto
A	30	RL	20%	13%	67%
		DEA	13%	10%	77%
		Conjunta	13%	10%	77%
B	41	RL	22%	12%	66%
		DEA	12%	10%	78%
		Conjunta	10%	12%	78%
C	52	RL	17%	15%	67%
		DEA	10%	15%	75%
		Conjunta	8%	17%	75%

Fonte: Onusic (2004).

O índice de correlação de *Spearman* foi calculado em cada uma das aplicações para verificar se as posições das empresas seguindo

os escores determinados pela DEA estavam correlacionadas com suas posições conforme os escores obtidos pela RL. Para as aplicações A, B e C foram calculados índices de correlação de 0,9249, 0,8341 e 0,7197 respectivamente. A última etapa deste trabalho foi analisar as empresas que foram classificadas erroneamente pelas duas técnicas (RL e DEA), e utilizar o recurso oferecido pela DEA de comparar as empresas ditas não eficientes com as empresas eficientes. Com a comparação de uma empresa classificada erroneamente pelas duas técnicas com uma empresa eficiente em DEA, buscou-se encontrar prováveis explicações para o erro de classificação das empresas.

A tabela a seguir resume as principais características dos trabalhos apresentados:

Tabela 2 - Resumo dos trabalhos relacionando DEA e insolvência

Autor	Indicadores	Amostra	Conclusão
Fernandez-Castro e Smith 1994	Posição de caixa; Liquidez; Posição de Capital de Giro; Alavancagem; Lucratividade; e Giro do Ativo	27 setores 1.423 empresas	Sucesso moderado em isolar as empresas falidas das sobreviventes
Emel et al. 2003	Liquidez, Atividade, Estrutura Financeira, Lucratividade, Crescimento e Fluxo de Fundos	82 indústrias	Consistentes com as análises tradicionais, mas indicações sobre como melhorar sua eficiência financeira
Cielen, Peeters e Vanhoof 2004	11 indicadores financeiros (estudos anteriores)	367 empresas 90 falidas (1994, 1995 e 1996)	Superou os demais em termos de acurácia, custo, desenvolvimento e compreensibilidade
Paradi, Asmild e Simak 2004	<i>Total assets; Working capital; Earnings before income, tax, depreciation and amortization; Retained earnings; Shareholders equity; Interest expense;</i>	17 e 11 falidas 160 e 115 não-falidas (1996 e 1997)	Modelos de piores práticas em combinação com a técnica de camadas, conduziu a taxas de acertos maiores

Tabela 2 - Resumo dos trabalhos relacionando DEA e insolvência (continuação)

Autor	Indicadores	Amostra	Conclusão
	<i>Cash flow from operations;</i> <i>Stability earnings;</i> <i>Total liabilities</i>		
ALMEIDA e MILIONI 2000	Retorno sobre o Ativo Total; Giro do Ativo Total Endividamento total	60 empresas (28 insolventes e 32 solventes) – Desenvolvimento 39 empresas – Validação	Taxas de acerto para 83% e 97%, nas amostra de calibração e verificação, respectivamente
ONUSIC 2004	Margem; Endividamento Geral e Endividamento de	122 empresas (1995 a 2001)	Em todas as aplicações (A, B e C) o desempenho da DEA foi melhor ou igual ao desempenho da RL. Foi proposta a utilização conjunta (RL e DEA) com taxas de acerto de 75 a 78%.

Fonte: elaboração própria.

4 CONCLUSÕES

O presente artigo teve por objetivo mapear trabalhos que relacionaram DEA e análise de insolvência, no Brasil e no exterior. Para tanto, foram pesquisadas as principais bases eletrônicas de publicações de língua inglesa e as principais revistas brasileiras nas áreas de administração, pesquisa operacional e produção. Também foram compulsados os anais de congressos e bancos de teses e dissertações, porém de forma não exaustiva.

Os trabalhos foram resumidos com ênfase na apresentação de sua metodologia, contrapondo pontos controvertidos, possibilidades e limitações. Constatou-se que o tema de pesquisa é bastante recente (data de 1994 o primeiro trabalho publicado) e ainda inexplorado (foram recuperados apenas seis trabalhos, sendo quatro estrangeiros e dois brasileiros). Ressalte-se, no entanto, que a pesquisa teve por limitações não recuperar de forma exaustiva os resultados de pesquisas apresentados em congressos e encontros. Percebe-se ainda um interesse crescente de pesquisadores pelo tema de estudo com um crescimento do número de trabalhos nos últimos anos. É importante ressaltar ainda que foi verificado um relevante número de estudos utilizando DEA para análise de insolvência em insti-

Mapeamento dos estudos sobre a utilização de Análise por Envoltória de Dados (DEA) na previsão de insolvência

tuições financeiras no exterior, com início na década de 1990. No entanto, por se tratar de um setor com características específicas, esses trabalhos não foram incluídos na presente pesquisa.

Assim, conclui-se que a utilização da Análise por Envoltória de Dados na análise de insolvência é um tema promissor para os pesquisadores que queiram investir em sua compreensão.

REFERÊNCIAS

ALMEIDA, H. R.; MILIONI, A. Z. *Análise de Envoltória de Dados na decisão de concessão de crédito*. Anais do XXXII Simpósio Brasileiro de Pesquisa Operacional, Viçosa, 18-20/10/2000.

ALTMAN, E. L. Financial rates, discriminant analysis and the prediction of corporate bankruptcy. *Journal of Finance*, v. 23, n. 4, 1968.

BANKER, R. D.; CHARNES, A.; COOPER, W. W. Some Models for estimating technical and scale inefficiencies in Data Envelopment Analysis. *Management Science*, 30, 9, p. 1078-1092, Sept. 1984

BARR, R. S.; SIEMS, T. F. Bank Failure Prediction Using DEA to Measure *Management Quality*. June, 1996.

BARR, R. S.; SIEMS, T. F. Predicting bank failure using DEA to quantify management quality. *Financial Industry Studies Dallas*, Iss. 1, p. 1 (31 pp.) Jan 1994

BELLONI, J. Â. *Uma metodologia de avaliação da eficiência produtiva de Universidades Federais Brasileiras*. Tese (Doutorado) – Programa de Pós-Graduação em Engenharia de Produção, do Departamento de Engenharia de Produção e Sistemas, da Universidade Federal de Santa Catarina. Florianópolis: UFSC, 2000.

BERGER, A. N.; HUMPHREY, D. B. Efficiency of financial institutions: international survey and directions for future research. *European Journal of Operational Research*, 98, p. 175-212, 1997.

CERETTA, P.S.; NIEDERAUER, C. A. P. Rentabilidade e eficiência do setor bancário brasileiro. *Revista de Administração Contemporânea*, v. 5, n. 3, set./dez. 2001.

CERETTA, P.S. Investigação empírica da eficiência do setor de alimentos. *Gestão e Produção*, v. 6, n. 3, p. 162-169, 1999.

CHARNES, A. *et. al.* *Data Envelopment Analysis: theory, methodology, and application*. Massachusetts (EUA): Kluwer, 1997.

CHARNES, A.; COOPER, W. W.; RHODES, E. *Measuring the efficiency of decision making units*. *European Journal of Operational Research*, 2(6): 429-444, 1978. Apud: CHARNES, A.; COOPER, W. W.; LEWIN, A. Y. & SEIFORD, L. *Data Envelopment Analysis: theory, methodology, and application*. Massachusetts (EUA): Kluwer, 1997.

CIELEN, A.; PEETERS, L.; VANHOOF, K. Bankruptcy prediction using a data envelopment analysis. *European Journal of Operational Research*, 154, p. 526-532, 2004.

ELIZABETSKY, R. *Um Modelo matemático para a decisão no banco comercial*. Trabalho de Formatura – Departamento de Engenharia de Produção, Escola Politécnica da Universidade de São Paulo. São Paulo: POLI-USP, 1976. Apud SILVA, José Pereira da. *Gestão e análise de risco de crédito*. São Paulo: Atlas, 2003.

EMEL, A. B. *et. al.* A Credit score approach for the commercial banking sector. *Socio-Planning Sciences*, 37, p. 103-123, 2003.

FERNANDEZ-CASTRO, A.; SMITH, P. Towards a general non-parametric model of corporate performance. *Omega – International Journal of Management Science*, 22, 3, p. 237-249, 1994.

KANITZ, S.C. Como prever falências. *Exame*, São Paulo: Abril, dez. 1974.

MATIAS, A.B. *Contribuição às técnicas de análise financeira: um modelo de concessão de crédito*. Trabalho de Formatura – Faculdade de Economia, Administração e Contabilidade. São Paulo: FEA-USP, 1978.

MATIAS, A. Borges; SIQUEIRA, J. O. Risco bancário: modelo de previsão de insolvência de bancos no Brasil. *Revista de Administração da USP*, 31(2): 19-28, abr./jun. 1996.

ONUSIC, L. M. *A utilização conjunta das técnicas análise por envoltória de dados e regressão logística na previsão de insolvência de empresas: um estudo exploratório*. Dissertação (Mestrado). São Paulo, 2004

PARADI, J.; ASMILD, M.; SIMAK, P. Using DEA and worst practice DEA in credit risk evaluation. *Journal of Productivity Analysis*, 21, p. 153-165, 2004.

PENDHAKAR, P. C. A Potential use of data envelopment analysis for the inverse classification problem. *Omega*, 30, p. 243-248, 2002.

SEIFORD, L. M.; ZHU, J. Na Acceptance system decision rule with Data Envelopment Analysis. *Computer Ops. Res.* Vol. 25. no. 4. p. 329-332.

SIEMS, T. F. Quantifying management's role in bank survival. Federal Reserve Bank of Dallas – *Economic Review*, p. 29-41, jan. 1992.

SIEMS, T. F.; BARR, R. S. Benchmarking the productive efficiency of U. S. banks. Federal Reserve Bank of Dallas – *Financial Industry Studies*, p. 11-24, dec. 1998.

SIEMS, T. F.; CLARK, J. Rethinking bank efficiency and regulation: how off-balance-sheet activities make a difference. Federal Reserve Bank of Dallas – *Financial Industry Studies*, p. 1-12, dec. 1997.

SILVA, J. P. *Gestão e análise de risco de crédito*. São Paulo: Atlas, 2003.

SILVA, J. P. *Modelos para classificação de empresas com vistas a concessão de crédito*. 1982. Dissertação (mestrado) – Escola de Administração de Empresas de São Paulo, Fundação Getúlio Vargas. São Paulo: EASP-FGV. Apud SILVA, J. P. *Gestão e análise de risco de crédito*. São Paulo: Atlas, 2003.

TROUTT, M. D.; RAÍ, A.; ZHANG, A. The Potential use of DEA for credit applicant acceptance system. Technical Note. *Computers Ops. Res.* 23, 4, p. 405-408, 1996.