

Apreçamento de opções de dólar no Brasil: uma avaliação dos modelos de redes neurais

Leandro dos Santos Maciel

Graduando no Instituto de Economia da Universidade de Campinas (UNICAMP).

Rosângela Ballini

Professor no Instituto de Economia da Universidade de Campinas (UNICAMP).

Rodrigo Lanna Franco da Silveira

Professor no Instituto de Economia da Universidade de Campinas (UNICAMP).

Departamento de Teoria Econômica

Instituto de Economia, Universidade Estadual de Campinas

Rua Pitágoras, 65 Cidade Universitária Zeferino Vaz

CEP 13083-857 Campinas – São Paulo – Brasil

Resumo

Neste artigo, foi aplicado um modelo de rede neural multicamadas para o apreçamento das opções de taxa de câmbio R\$/US\$, negociadas na Bolsa de Valores, Mercadorias & Futuros, BM&FBOVESPA, no período de janeiro de 2004 a dezembro de 2007. O desempenho do modelo foi comparado com os resultados obtidos pela fórmula analítica de Black (1976), instrumento mais utilizado pelos agentes de mercado na avaliação de opções sobre moeda estrangeira. A volatilidade foi estimada por um processo GARCH (1,1) e a curva de juros foi construída por meio da metodologia de interpolação por *splines* cúbicos. Os preços teóricos obtidos foram comparados com os preços reais observados no mercado. Os resultados revelaram a superioridade do modelo de inteligência artificial, em comparação ao modelo de Black, nos diferentes graus de *moneyness*.

Palavras-chave: Redes Neurais Artificiais, Apreçamento de Opções, Modelo de Black.

1. Introdução

O desenvolvimento do modelo de Black-Scholes – B&S (Black e Scholes, 1973), na década de 1970, representou um grande avanço na teoria de finanças. A partir de um conjunto de hipóteses, Fischer Black e Myron Scholes obtiveram um método de apreçar opções européias sobre ações que não pagam dividendos, sendo o prêmio de uma *call* ou uma *put* uma função do preço do ativo objeto, preço de exercício, tempo de maturidade, taxa de juro livre de risco e volatilidade do preço do ativo-objeto. Tais hipóteses são baseadas nos seguintes pontos: i) o preço da ação segue um movimento geométrico Browniano, em que a distribuição de probabilidade dos preços é lognormal, com média e variância constantes; ii) a taxa de juro livre de risco de curto prazo é conhecida e constante para qualquer vencimento; iii) a volatilidade é constante até o vencimento; iv) há possibilidade de emprestar e tomar emprestado à mesma taxa de juro livre de risco; v) inexistência de pagamento de dividendos durante a vida da opção; vi) o ativo-objeto do contrato é perfeitamente divisível, sendo que sua negociação é contínua e não possui custos de transação ou impostos; vii) não se verificam oportunidades de arbitragem sem risco (Hull, 2006).

Um entrave ao modelo se encontra no fato de que os dados reais com frequência violam suas hipóteses subjacentes. Essas violações podem ser observadas nas seguintes situações: (1) ao invés da descrição de um caminho aleatório, a teoria dos fractais pode ser uma melhor hipótese para o movimento do mercado (Peters, 1994) - ademais, Akgiray (1989) ainda rejeita a hipótese de variância constante para o preço das ações; (2) opções americanas ou bermudas são as mais comercializadas nos mercados - os agentes optam por maiores chances de exercício quando os mercados estão voláteis, assim como quando estão próximos ao pagamento de dividendos; (3) o pagamento de dividendos é uma prática comum, especialmente no mercado de ações.

Na tentativa de superar as violações citadas acima e se adequar aos diversos tipos de opções que surgiram, vários trabalhos foram e vêm sendo desenvolvidos. Neste processo de evolução, modelos baseados em Redes Neurais Artificiais (RNAs) vêm ganhando espaço na literatura de finanças. Trata-se de uma técnica computacional inspirada no funcionamento do sistema nervoso biológico, motivada para a compreensão de relações e padrões de diferentes naturezas, em que permite a modelagem de relações do tipo entrada-saída (*input-output*). Muitos estudos com avaliações de redes neurais aplicadas à modelagem financeira podem ser encontrados na literatura.

Hutchinson et al (1994), por exemplo, avaliaram uma abordagem de redes neurais para o apreçamento e aplicação da operação *delta-hedge* nas opções sobre futuros do índice S&P 500, no período de 1987 a 1991. Seus resultados mostraram que as redes neurais foram capazes de aprender o funcionamento da fórmula de Black-Scholes, com uma amostra para ajuste dos parâmetros composta por preços diários das opções no período de dois anos. Sobretudo, foi possível apreçar e formar operações *delta-hedge* com sucesso para opções não utilizadas na amostra apresentada à rede. Os autores sugeriram que o apreçamento por tal método pode ser mais eficiente quando a dinâmica do ativo objeto da opção é desconhecida.

O estudo de Qi e Maddala (1996), por sua vez, aplicou um modelo de rede neural multicamadas recorrente para o apreçamento de *calls* sobre o índice S&P 500 no período que compreende dezembro de 1994 a janeiro de 1995. Os resultados apresentaram a superioridade do método de inteligência artificial em comparação à fórmula de Black-Scholes. Com base na análise dos pesos da rede neural, confirmaram que quanto menor o preço de exercício ou maior o preço do ativo objeto ou, ainda, quanto mais distante o prazo para o vencimento, maior será o preço da *call*.

Outra pesquisa que se observa o uso de redes neurais para opções é o de Lajbcygier et al (1996). Os autores utilizaram dois modelos para apreçar opções sobre o futuro do índice representativo das ações ordinárias comercializadas no mercado de ações australiano. O primeiro deles teve duas variáveis de entrada (quociente do preço *spot* e o preço de exercício, e o prazo de vencimento do papel). Já o segundo utilizou as mesmas variáveis, adicionando a taxa de juros e a volatilidade (estimada por meio da volatilidade histórica com uma janela de 60 dias), na tentativa de comparar com o modelo utilizado por Hutchinson et al (1994). Os resultados mostraram que o modelo com quatro variáveis de entrada superou o modelo de duas variáveis, assim como a metodologia de Black-Scholes.

Para o mercado brasileiro, Freitas e Souza (2002) avaliaram o modelo de redes neurais para precificar opções de compra européias sobre a ação da Telebrás PN, no período de janeiro de 1995 a dezembro de 1999. A rede neural aplicada pelos autores apresentou resultados superiores em comparação ao modelo de Black-Scholes, especialmente quando as *calls* estavam *in-the-money*.

Outras pesquisas podem ainda ser citadas, como a de Geigle e Aronson (1999) e Malliaris e Salchenberger (1996). Estes primeiros examinaram o desempenho das redes neurais para a avaliação de opções americanas sobre futuro do índice S&P 500 e confirmaram a superioridade do modelo se comparado ao de Black-Scholes. Já Malliaris e Salchenberger (1996) desenvolveram um modelo de rede neural para a previsão da volatilidade futura das opções sobre futuro do índice S&P 100, com base na volatilidade passada e outros fatores que afetavam os preços das opções comercializadas neste mercado. A rede foi estruturada com mais 13 fatores de entrada e a estimativa da volatilidade permitiria aos agentes estabelecer posições estratégicas na antecipação de mudanças nos movimentos do mercado. Os resultados da rede neural foram comparados com a volatilidade implícita das opções, obtidas pelo modelo de Black-Scholes. De acordo com os seus resultados, a estimação da volatilidade pelo modelo de redes neurais foi superior às modelagens de volatilidade histórica e implícita.

Embora a superioridade dos modelos de redes neurais, em comparação aos métodos usuais de apreçamento de opções, seja quase um consenso entre os pesquisadores, Callen et al (1996) afirmam que estas técnicas não apresentam, necessariamente, os melhores resultados aos modelos

lineares de séries temporais, mesmo quando aplicados a séries financeiras, sazonais ou não-lineares. Com isso, é necessária a condução de estudos para diferentes mercados e com a utilização de diferentes modelos de redes neurais na tentativa de inferir sua superioridade ao modelo de Black-Scholes.

Diante deste contexto, este artigo teve o objetivo de aplicar um modelo de rede neural recorrente multicamadas, baseado no algoritmo *backpropagation*, para valoração das opções de compra de taxa de câmbio R\$/US\$, negociadas na Bolsa de Valores, Mercadorias & Futuros – BM&FBOVESPA, no período de 2004 a 2007. A fim de avaliar tal modelo, as opções também foram apreçadas pelo modelo de Black, metodologia utilizada pelos agentes no mercado brasileiro para opções sobre moeda estrangeira, e, em seguida, os resultados foram comparados com os prêmios observados no mercado.

2. Metodologia

2.1 Amostra

A base de dados deste trabalho correspondeu aos preços de fechamento das opções européias de compra de taxa de câmbio R\$/US\$, negociadas na BM&FBOVESPA, para diferentes prazos e preços de exercícios, no período de 02/01/2004 a 31/12/2007 – total de 1.001 dias de negócios e 4.642 observações. Para cada opção, foram obtidos o preço de exercício, o número de contratos negociados, o prazo para o vencimento em dias úteis, o valor de fechamento para cada opção e o preço do contrato futuro de mesmo prazo da opção¹. A amostra teve base nos papéis que tiveram mais de 500 contratos negociados, ou seja, selecionaram-se as *calls* mais líquidas transacionadas no mercado. As subseções seguintes apresentam os métodos de apreçamento utilizados na pesquisa.

2.2. Modelo de Black

Como mencionado anteriormente, a partir de um conjunto de hipóteses, Black e Scholes (1973) propuseram um modelo para cálculo dos prêmios de *calls* (c) e *puts* (p) européias sobre ações que não pagam dividendos – equações (1) e (2).

$$c = S \cdot N(d_1) - X \cdot e^{-rT} \cdot N(d_2) \quad (1)$$

$$p = X \cdot e^{-rT} \cdot N(-d_2) - S \cdot N(-d_1) \quad (2)$$

Em que:

$$d_1 = \frac{\ln(S/X) + (r + \sigma^2/2) \cdot T}{\sigma \cdot \sqrt{T}} \quad (3)$$

$$d_2 = \frac{\ln(S/X) + (r - \sigma^2/2) \cdot T}{\sigma \cdot \sqrt{T}} = d_1 - \sigma \cdot \sqrt{T} \quad (4)$$

Sendo, S o preço à vista (*spot*) do ativo subjacente à opção; X o preço de exercício; T o prazo de vencimento, em anos (base 252 dias úteis); σ a volatilidade do ativo objeto (ao ano); r a taxa de juro de curto prazo livre de risco (ao ano e capitalizada continuamente); $N(\cdot)$ a função de probabilidade Normal cumulativa.

A partir do modelo acima, Merton (1973) estendeu o modelo para inclusão de pagamento de dividendos nas ações e Black (1976) desenvolveu as fórmulas analíticas que permitem a avaliação de opções sobre contratos futuros – equações (5) e (6). Nota-se que o preço *spot*, S , é substituído pelo preço futuro, F .

¹ Dados fornecidos pela BM&FBOVESPA.

$$c = [F \cdot N(d_1) - X \cdot N(d_2)] \cdot e^{-rT} \quad (5)$$

$$p = [X \cdot N(-d_2) - F \cdot N(-d_1)] \cdot e^{-rT} \quad (6)$$

Sendo:

$$d_1 = \frac{\ln(F/X) + \sigma^2 \cdot T/2}{\sigma \cdot T} \quad (7)$$

$$d_2 = d_1 - \sigma \cdot \sqrt{T} \quad (8)$$

Garman e Kohlhagen (1983) fizeram ainda algumas modificações de forma a contemplar o apreçamento de opções sobre taxa de câmbio. No entanto, o modelo de Black, acima descrito, também é amplamente utilizado no mercado financeiro brasileiro para a obtenção dos prêmios das opções de dólar. Assim sendo, o presente trabalho fez uso deste último modelo, sendo, então, as variáveis consideradas: preço de exercício (X); tempo até o vencimento, ao ano, base 252 dias úteis (T); preço do contrato futuro de taxa de câmbio R\$/US\$ de mesmo vencimento da opção (F); a taxa de juros prefixada (ao ano) e a volatilidade.

Cabe salientar como as últimas variáveis (taxa de juros e volatilidade) foram dimensionadas no presente trabalho.

É comum verificar, em diversas séries financeiras, que a hipótese de homocedasticidade (variâncias condicionais do termo errático, u_i , constantes), adotada em vários modelos econométricos, é inapropriada. Consta-se, em geral, que períodos de grandes oscilações de preços são seguidos por intervalos de tempo de relativa estabilidade nas cotações, levando a uma “variação da volatilidade” ao longo do tempo. Neste contexto, Engle (1982) desenvolveu o modelo ARCH (*Autoregressive Conditional Heteroscedasticity*) com o objetivo de estimar a volatilidade de séries com as características citadas acima, expressando a variância condicional em termos do quadrado dos retornos passados. Posteriormente, Bollerslev (1986) estendeu o modelo ARCH, com o desenvolvimento do modelo GARCH (*Generalized ARCH*), em que a variância condicional passa a depender do quadrado dos retornos passados e das variâncias condicionais passadas (Enders, 2004).

Além de apresentar resultados mais robustos, em comparação aos modelos tradicionais, a literatura mostra que os modelos de volatilidade determinística, mencionados acima, apresentam melhores resultados, quando comparados à metodologia da volatilidade histórica (Morais e Portugal, 1999; Jorion, 2003). Dessa forma, no presente trabalho, a estimação da volatilidade do dólar foi calculada com base em um GARCH (1, 1) – equação (9), em que a variância condicional do termo errático depende do termo de erro quadrático e da variância condicional, definidos no instante anterior, $t - 1$.

$$\sigma_t^2 = \alpha_0 + \alpha_1 \mu_{t-1}^2 + \alpha_2 \sigma_{t-1}^2 \quad (9)$$

A razão da adoção do GARCH (1,1) se baseia na sua parcimônia e na sua boa adequação aos dados. Além disso, conforme Silva, Sáfadi e Castro Júnior (2005), poucas restrições nos parâmetros são observadas neste modelo – para a variância ser positiva e fracamente estacionária, é preciso que $\alpha_1 + \alpha_2 < 1$; $\alpha_0 > 0$; $\alpha_1 > 0$; $\alpha_2 > 0$.

Quanto ao cômputo da taxa de juros, adotou-se a metodologia de interpolação por *splines* cúbicos. Monteiro e Salles (2002) compararam diferentes processos de interpolação de taxas de juro no mercado brasileiro e, com base num critério de seleção que combina estratégias de reamostragem do tipo *leave-k-out cross-validation*, mostraram que, a partir dos contratos futuros de juros e *swaps*, o método de interpolação mais adequado à realidade brasileira foi baseado nos *splines* cúbicos. Trata-se de uma metodologia que busca estimar a curva de juro a partir de preços

observados de ativos negociados em função de suas respectivas maturidades e de informações *a priori* sobre a natureza da curva de juros.

O principal objetivo do modelo por *splines* é a construção da curva de juros, via interpolação. Considere os pontos $y_1 = g(x_1), \dots, y_n = g(x_n)$, em que $a = x_1 < \dots < x_n = b$. O problema de interpolação consiste em construir uma função contínua f que permita o cálculo de $g(z)$, sendo z um ponto interior no intervalo definido por a e b ². Neste trabalho, $g(x_j)$ é a taxa de juro pré *versus* DI dos contratos de *swap*, negociados na BM&FBOVESPA, para a maturidade x_j ³. Sendo f um *spline* cúbico, para cada intervalo $\{x_j, x_{j+1}\}$, f deve ser igual a um polinômio cúbico P_j , da forma:

$$P_j(x) = a_{0,j} + a_{1,j} \cdot x + a_{2,j} \cdot x^2 + a_{3,j} \cdot x^3 \quad (10)$$

$$f(x) = P_j(x) \text{ para cada } x \in \{x_j, x_{j+1}\}, j = 1, \dots, n-1.$$

É necessário que f seja contínua e com derivadas, primeira e segunda, contínuas no intervalo $[a, b]$. Além disso, f deve ser igual a g nos pontos x_1, \dots, x_n (Monteiro e Salles, 2002). Para satisfazer essas condições, impõem-se as seguintes condições:

$$P_j(x_j) = g(x_j); \quad j = 1, \dots, n-1;$$

$$P_j(x_{j+1}) = g(x_{j+1}); \quad j = 1, \dots, n-1;$$

$$P_j'(x_j) = s_j; \quad j = 1, \dots, n-1;$$

$$P_j'(x_{j+1}) = s_{j+1}; \quad j = 1, \dots, n-1;$$

$$P_j''(x_j) = P_j''(x_j); \quad j = 2, \dots, n-1. \quad (11)$$

Em que, s_1, \dots, s_n são parâmetros a serem delimitados pelo modelo⁴. Uma vez fixados os valores desses parâmetros, P_j estará determinado e a construção de f será finalizada⁵. Com todas as entradas do modelo de Black obtidas, o próximo passo consiste em determinar o modelo de rede neural aplicado para precificar as opções de compra de dólar consideradas.

2.3. Rede Neural

As Redes Neurais Artificiais (RNA) são modelos computacionais não lineares que, na tentativa de emular o cérebro humano, organizam-se por meio de elementos básicos de processamento denominados neurônios artificiais, que se interligam na formação de redes capazes de processar, armazenar e transferir informações vindas de um meio exterior (Haykin, 2001). A Figura 1 apresenta um modelo de neurônio artificial, conhecido como neurônio de Mcculloch-Pitts (Haykin, 2001) que, de forma similar ao neurônio biológico, recebe informações de outros neurônios ou de um meio externo e as processa de forma a obter uma resposta em sua saída.

² Caso z esteja fora do intervalo definido por a e b , o problema é dito extrapolação.

³ Os dados dos contratos foram fornecidos pela BM&FBOVESPA.

⁴ Essa modelagem foi implementada no *software* Matlab®, assim como todas as aplicações deste trabalho.

⁵ Para mais detalhes do método de *splines* cúbicos ver o trabalho de Monteiro e Salles (2002).

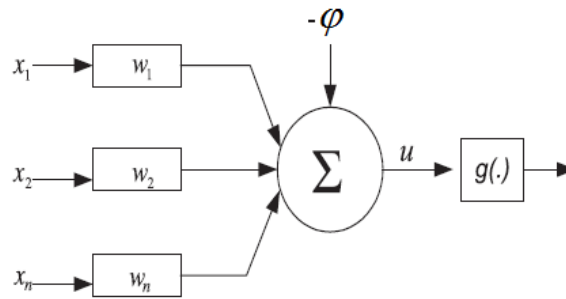


Figura 1: Neurônio Artificial

Fonte: Adaptado de Haykin (2001)

No neurônio artificial da Figura 1, x_1, x_2, \dots, x_n representam os sinais de entrada, w_1, w_2, \dots, w_n são os pesos sinápticos do neurônio, φ é o limiar de ativação do neurônio que define se a saída será disparada, u é a saída do combinador linear em que $u = \sum_{i=1}^N w_i \cdot x_i - \varphi$, e $g(\cdot)$ é uma função de ativação da saída do combinador linear que processa o conjunto de entradas recebidas e as transforma em estado de ativação⁶.

2.3.1. Estrutura das Redes Neurais Artificiais

Os neurônios artificiais podem ser conectados e organizados de várias formas, o que configura a estrutura ou arquitetura da rede neural, com características e aplicações bastante distintas. Em geral, as redes neurais podem ter uma única camada ou múltiplas camadas. Numa rede formada por uma camada única, tem-se a entrada de dados e uma única camada de neurônios em paralelo, que é a própria camada de saída. Esse tipo de rede, entretanto, sofre limitações em representar funções linearmente separáveis. Porém, são superadas pela utilização de redes multicamadas, formadas por uma camada de entrada, uma ou mais camadas intermediárias e uma camada de saída. As arquiteturas das redes neurais também podem ser classificadas como recorrentes e não recorrentes (Russel e Norvig, 1995).

Redes neurais não recorrentes são redes que não possuem realimentação de suas saídas para suas entradas (Haykin, 2001). Podem ser formadas por uma camada única ou por múltiplas camadas. RNAs multicamadas não recorrentes contém um conjunto de entradas, uma ou mais camadas escondidas (ou intermediárias) e uma camada de saída (ver Figura 2).

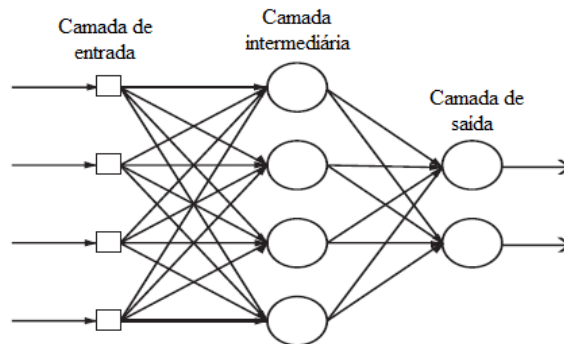


Figura 2: RNA multicamada não recorrente

Fonte: Adaptado de Haykin (2001)

⁶ As principais funções de ativação são: linear, sigmóide ou logística, degrau bipolar e hiperbólica.

No modelo da Figura 2, a camada de saída apresenta a saída da rede. São denominadas camadas intermediárias ou ocultas, as camadas que não possuem ligações diretas com a entrada. Além disso, não existem conexões ligando um neurônio de uma camada a outro de uma camada anterior, nem a um neurônio da mesma camada. As conexões somente são feitas para neurônios da camada seguinte.

Nas redes não recorrentes a saída é apenas uma resposta imediata ao estímulo da entrada, ou seja, as informações prévias não são reutilizadas, produzindo apenas mapeamentos estatísticos. Dessa forma, esse modelo apresenta restrições para representar sistemas variantes no tempo, uma vez que o vetor de saída da rede neural depende apenas do vetor de entrada definido no mesmo instante t (Kolen, 1994).

Já as redes neurais recorrentes se caracterizam pela realimentação das saídas pelas entradas, sendo suas saídas determinadas pelas entradas atuais e pelas saídas anteriores (Haykin, 2001). As interconexões nesse tipo de arquitetura são bem mais complexas que entre redes não recorrentes, pois, quando organizadas em camadas, possuem interligações entre neurônios da mesma camada e entre camadas não consecutivas (Jain, 2001). Neste trabalho, foi aplicado um modelo de rede neural recorrente. Outra questão chave na caracterização de um modelo de rede neural é a utilização do algoritmo de treinamento, descrito a seguir.

2.3.2. Treinamento

O treinamento de uma rede neural consiste no ajuste dos pesos sinápticos w_j e dos limiares φ . O treinamento objetiva fazer com que a rede seja capaz de generalizar um problema, i.e., a rede não precisa conhecer todas as entradas necessárias para obter a saída, de forma que a aplicação de um conjunto de entradas desconhecidas seja capaz de gerar dados de saída desejáveis (Haykin, 2001).

Os algoritmos de treinamento de uma rede neural podem ser supervisionados ou não supervisionados. No treinamento do tipo supervisionado, a rede recebe os vetores de entrada, assim como as saídas desejadas. O treinamento, portanto, ajustará os pesos entre os neurônios de maneira a se obter a saída desejada, ou seja, a rede é treinada para fornecer um conjunto de saída desejada a um estímulo de entrada específico. Na literatura, o algoritmo de treinamento *backpropagation*, ou conhecido como algoritmo de retropropagação do erro, é o mais citado, por isso, foi utilizado neste trabalho.

O mecanismo de funcionamento do algoritmo *backpropagation* é formado por dois passos. No primeiro passo, a entrada é fornecida aos neurônios da camada de entrada da RNA, denominado de passo da propagação. O sinal se propaga para as próximas camadas, ou seja, no sentido das conexões, até a camada de saída, obtendo, assim, o primeiro resultado. Nesse momento, não ocorrem alterações nos pesos da rede. O erro dado pela diferença entre a saída da rede e a resposta desejada é calculado. A soma dos erros quadráticos é obtida para todos os neurônios da camada de saída, isto é:

$$E = \frac{1}{2N} \sum_{l=1}^N \sum_{i=1}^p (y_i^l - \hat{y}_i^l)^2 \quad (12)$$

Sendo, E a função da medida do erro total da rede, n o número de padrões apresentados à rede, p o número de neurônios na camada de saída, y_i^l a saída desejada e \hat{y}_i^l a saída gerada pela rede, para o l -ésimo padrão.

O passo de retropropagação, segundo passo, é caracterizado pelo ajuste dos pesos da rede de forma a minimizar a função E . A variação dos pesos sinápticos é proporcional à direção oposta ao gradiente do erro em relação aos pesos da rede, conforme:

$$\Delta w_{ij} = -\eta \cdot \left(\frac{\partial E}{\partial w_{ij}} \right) \quad (13)$$

Em que, os índices i e j referem-se ao neurônio i da camada posterior e ao neurônio j da camada anterior e η é uma constante de proporcionalidade, chamada taxa de aprendizagem, a qual assume valores no intervalo $(0, 1)$. Esta taxa determina a “velocidade” do passo em direção ao mínimo da função de erro E . Um valor elevado de η produzirá uma rápida aprendizagem, resultando em grandes alterações nos pesos. Entretanto, aumentará o risco da rede tornar-se instável, podendo oscilar ao redor da solução. Por outro lado, se η assumir valores pequenos, as mudanças nos pesos serão menores, requerendo muito mais iterações até a convergência.

Ao calcular a derivada para todas as camadas da rede, as fórmulas para a atualização dos pesos w_{ij} , para cada padrão de entrada individualmente, são:

$$\Delta w_{ij} = \eta \cdot \delta \cdot x_j \quad (14)$$

Onde,

$$\delta = g'_j \left(\sum_{k=1}^{N2} \delta_k w_{kj} \right) \quad (15)$$

Sendo, g'_j a derivada da função de ativação do neurônio j ⁷.

Apesar de ser o algoritmo de treinamento mais utilizado, o *backpropagation* apresenta baixa velocidade de convergência quando próximo de um mínimo da função de erro E , uma vez que a atualização dos pesos é proporcional ao valor do gradiente do erro; além disso, tem-se um problema com os mínimos locais, pois o mínimo global de E não é localizado quando se atinge um mínimo local.

2.3.3. Modelo de Rede Neural

Construiu-se um modelo de apreçamento para as opções de compra de dólar no Brasil pelo método de redes neurais com base nas principais variáveis que influenciam o preço desse ativo: o preço futuro do dólar no mercado (F), o preço de exercício da opção (X), o tempo para o vencimento da opção (T), a volatilidade do preço do dólar (σ) e a taxa de juro livre de risco (r). Os dados foram os mesmos aos aplicados para apreçamento por meio do modelo de Black, sendo a volatilidade estimada por um processo GARCH (1,1) e a taxa de juro livre de risco obtida pelo método de interpolação via *splines* cúbicos⁸.

Na literatura sobre redes neurais, não existe trabalho que especifique os parâmetros ideais para a construção de um modelo de rede que resulte nos melhores resultados. Portanto, o processo envolve escolhas empíricas que variam de acordo com a especificidade dos dados e o objetivo de estimação e previsão (Kaastra e Boyd, 1995). Neste trabalho, foi escolhido o modelo de rede neural que, dentre os vários modelos construídos, apresentou os melhores resultados. Trata-se de um modelo de rede neural multicamadas recorrente, constituído por cinco neurônios compondo a camada de entrada,

⁷ Uma apresentação mais detalhada do algoritmo de *backpropagation* pode ser vista no trabalho de Rumelhart e McClelland (1986).

⁸ Todos os dados inseridos tanto no modelo de Black quanto no modelo de rede neural proposto estão em termos anuais e, além disso, a apuração do tempo até o vencimento da opção foi considerada com base nos dias de negociação, ignorando os dias em que a bolsa esteve fechada.

seis neurônios na camada intermediária e uma camada de saída com um neurônio, representando o preço da opção. A Figura 3 apresenta o modelo de rede neural aplicado para o apreçamento das opções avaliadas.

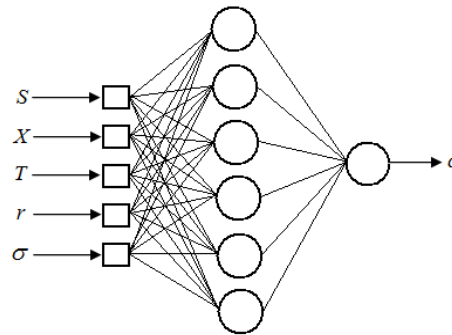


Figura 3: Modelo de Rede Neural para Apreçamento das Opções de Compra de Dólar.

A amostra total foi dividida em três conjuntos: treinamento, validação e teste. Característica básica na construção de um modelo de rede neural, essa divisão permite que os parâmetros da rede sejam ajustados com base no conjunto de treinamento; o erro, avaliado pelo conjunto validação e, finalmente, o conjunto denominado teste apresenta valores nunca vistos pela rede para avaliar seu grau de aprendizagem e generalização. Os conjuntos treinamento, validação e teste possuem um total de 75%, 15% e 10% do total da amostra, respectivamente.

O treinamento da rede se deu com base no algoritmo *backpropagation*. O número de épocas foi fixado em 800⁹, para permitir um maior grau de generalização à rede, e, para que se possa alcançar o menor valor para a função de erro E . O nível de erro foi estipulado em 10^{-4} , uma vez que são níveis de erro aceitáveis para um modelo de estimação e previsão de séries temporais financeiras, e permite que se obtenha um mínimo local da função de erro mais próximo do mínimo global, caso este não seja atingido. Por fim, a taxa de aprendizagem utilizada na estruturação da rede neural foi de 0.06.

2.4. Métricas de Erro

Os preços teóricos das opções, obtidos por meio de cada modelo, foram comparados com os preços realmente observados no mercado. Para avaliação dos modelos, utilizaram-se o Erro Percentual Médio (EPM), o Máximo Erro Percentual (MEP) e a Raiz do Erro Quadrático Relativo Médio (EQRM), representados por:

$$EPM = \frac{100}{N} \sum_{i=1}^N \frac{|c_i - \hat{c}_i|}{c_i} \quad (16)$$

$$MEP = \max \left\{ \frac{100}{N} \sum_{i=1}^N \frac{|c_i - \hat{c}_i|}{c_i} \right\} \quad (17)$$

⁹ Uma época é definida como a realização dos dois passos do algoritmo *backpropagation*: apresentação do dado de entrada, obtenção da primeira saída da rede, cálculo do erro e a etapa de ajuste dos pesos retro-propagando o erro camada por camada.

$$EQRM = \sqrt{\frac{1}{N} \sum_{i=1}^N \left(\frac{c_i - \hat{c}_i}{c_i} \right)^2} \quad (18)$$

Em que, c representa o preço de fechamento, observado no mercado, da opção de compra de dólar, \hat{c} o preço teórico obtido pelos modelos e N o número total de observações.

Os resultados dos modelos ainda foram avaliados de acordo com o grau de *moneyness* (M) das opções, sendo este definido pela relação entre o preço de exercício da opção e o preço a vista do ativo objeto:

$$M = \frac{S}{X \cdot e^{-rT}} \quad (19)$$

Sendo S a cotação da PTAX. Neste caso, as opções da amostra teste foram divididas conforme o grau de *moneyness*: *out-of-the-money* ($M \leq 1 - \alpha\%$), *at-the-money* ($1 - \alpha\% < M \leq 1 + \alpha\%$) e *in-the-money* ($M > 1 + \alpha\%$), considerando $\alpha = 5\%$.

Além disso, calculou-se o coeficiente de determinação, R^2 , da regressão linear dos preços das opções observadas no mercado sobre os preços teóricos:

$$c_i = \phi_1 + \phi_2 \cdot \hat{c}_i \quad (20)$$

Onde, ϕ_1 e ϕ_2 são os parâmetros da regressão.

3. Análise e Discussão dos Resultados

Para confrontar o desempenho do modelo baseado em redes neurais com a fórmula analítica de Black, foram comparados os preços teóricos das opções obtidos por meio de cada modelo com os preços observados no mercado. Nessa comparação, os dados utilizados pertencem ao grupo de teste da estruturação da rede neural, uma vez que o conjunto treinamento e validação foram utilizados para que a rede pudesse determinar seus parâmetros e aprender o processo gerador dos preços de compra das opções de dólar. Dessa forma, o conjunto teste permite avaliar a capacidade de aprendizado e generalização da rede para precificar as opções com dados de entrada nunca antes apresentados à rede. O conjunto teste é composto por negociações que se iniciam em 04/10/2007 e vão até 31/12/2007, compondo uma amostra de 466 observações.

A Tabela 1 apresenta os resultados dos valores das métricas de erro utilizadas neste trabalho para os modelos teóricos avaliados, assim como o coeficiente de determinação da regressão dos preços do mercado sobre os preços teóricos.

Tabela 1. Aderência dos preços calculados pelos modelos Black e de Redes Neurais aos preços de fechamento das *calls* de dólar européias negociadas na BM&FBOVESPA entre 04/10/2007 e 31/12/2007. Amostra com 466 observações.

Modelos	Métricas de erro			
	<i>EPM</i>	<i>MEP</i>	<i>EQRM</i>	R^2
RNA	1,3561	3,6340	0,4563	0,8241
Black	2,0946	4,9129	2,2280	0,6998

Fonte: resultados da pesquisa

Por meio da avaliação dos erros obtidos, ao se observar os resultados da Tabela 1, pode-se verificar que o modelo de redes neurais apresentou preços teóricos mais próximos aos preços de mercado, explicando o baixo EPM, MEP e EQRM, em comparação ao modelo de Black. Além disso, o baixo desempenho do modelo de Black foi observado em toda a amostra dos dados. A Tabela 2 apresenta os valores das métricas de erro para toda a amostra, totalizando 4.642 observações, divididas nos conjuntos de treinamento, validação e teste.

Tabela 2. Aderência dos preços calculados pelos modelos Black e Redes Neurais aos preços de fechamento das opções de compra de dólar européias negociadas na BM&FBOVESPA entre 02/01/2004 e 31/12/2007. Amostra com 4.642 observações.

Modelos	Treinamento			Validação			Teste		
	<i>EPM</i>	<i>MEP</i>	<i>EQRM</i>	<i>EPM</i>	<i>MEP</i>	<i>EQRM</i>	<i>EPM</i>	<i>MEP</i>	<i>EQRM</i>
RNA	0,08	1,23	0,01	0,34	1,76	0,03	1,36	3,63	0,46
Black	2,63	5,04	2,46	1,98	5,21	2,32	2,09	4,91	2,23

Fonte: resultados da pesquisa

Nota-se, pela Tabela 2, que o modelo de Black apresentou erros relativamente iguais nas amostras, uma vez que não faz diferença a distinção entre elas para a aplicação da fórmula analítica. Entretanto, como era de se esperar, nos conjuntos treinamento e validação, para a rede neural, os erros foram muito baixos, se comparados aos próprios resultados do conjunto teste da rede neural e aos resultados do modelo de Black, visto que foram os dados apresentados à rede e que a estruturaram de forma a convergir para o nível de erro estipulado de acordo com o processo de aprendizagem supervisionado.

Os resultados, portanto, foram favoráveis à metodologia das redes neurais. Esse fato se confirma ao se observar o R^2 da regressão dos preços de mercado sobre os preços teóricos (Tabela 1). Entretanto, deve-se notar que, mesmo assim, os preços do mercado não são explicados fundamentalmente pelos preços teóricos, ou seja, outros fatores podem ajudar a determinar os preços dessas opções. Porém, em geral, os valores dos erros para o modelo proposto foram muito pequenos em comparação ao modelo de Black, evidenciando que a rede foi capaz de se aproximar melhor ao processo gerador do preço das opções de compra de dólar no mercado brasileiro.

As médias dos erros para os dois modelos foram elevadas e são significativamente diferentes de zero. Além disso, algumas observações apresentaram erros relativos superiores a 50%, indicando uma inadequação dos modelos para o apreçamento, como também o possível caso de erros grosseiros de informação. Esses dados podem ser vistos como *outliers*, mas, mesmo assim, continuaram na amostra, uma vez que se intencionou uma modelagem de caráter geral. Também, no período avaliado, a dinâmica de comportamento do ativo objeto, preço do dólar, não apresentou um movimento adverso capaz de explicar tais resultados, como uma reversão nas expectativas dos agentes, por exemplo, ou até mesmo a verificação de um fator exógeno que se reverteu numa inflexão da tendência do movimento da taxa cambial entre o dólar e o real.

Foram avaliados, ainda, os resultados dos modelos de acordo com o grau de *moneyness* (M) das opções. Verificou-se que 43% das opções da amostra teste estão dentro do preço, enquanto 31% e 26% estão no dinheiro e fora do dinheiro, respectivamente. A Tabela 3 apresenta os resultados das métricas de erro para o conjunto da amostra teste de acordo com o grau de *moneyness* das opções. Observa-se que tanto o modelo de Black quanto o de rede neural foram mais acurados no apreçamento das opções *in-the-money*. Os maiores níveis de erro foram verificados nas opções *out-of-the-money*, principalmente para o modelo de Black. Na amostra que compõe as opções *at-the-*

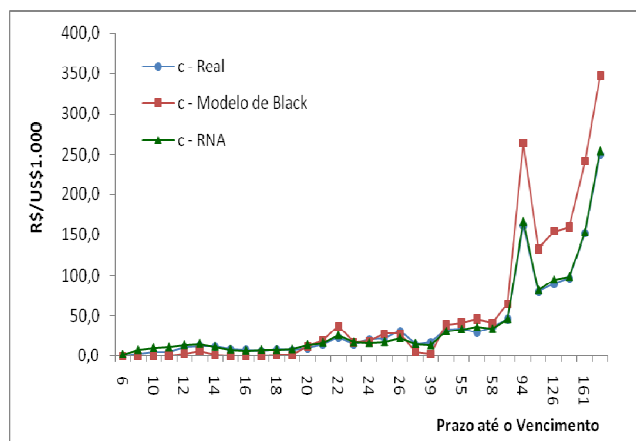
money, os resultados dos modelos foram mais semelhantes, como se pode notar pelas métricas do EPM e do MEP. Entretanto, nos diferentes graus de *moneyness*, os resultados do modelo de RNA foram superiores em comparação ao modelo de Black.

Tabela 3. Aderência dos preços calculados pelos modelos Black e Redes Neurais aos preços de fechamento das opções de compra de dólar européias negociadas na BM&FBOVESPA entre 02/01/2004 e 31/12/2007 de acordo com o grau de *moneyness*. Amostra com 466 observações.

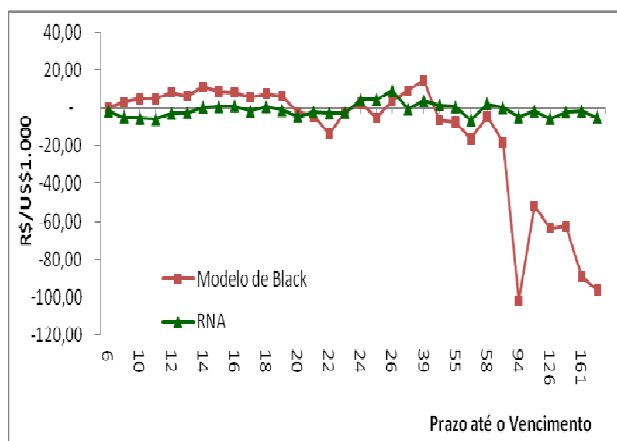
Modelos	<i>Out-of-the-money</i>			<i>At-the-Money</i>			<i>In-the-money</i>		
	<i>EPM</i>	<i>MEP</i>	<i>EQRM</i>	<i>EPM</i>	<i>MEP</i>	<i>EQRM</i>	<i>EPM</i>	<i>MEP</i>	<i>EQRM</i>
RNA	1,51	3,21	0,52	1,24	3,27	0,44	1,07	2,48	0,38
Black	2,33	5,09	2,64	1,89	4,74	2,09	1,66	4,12	1,89

Fonte: resultados da pesquisa.

O Gráfico 1 ilustra o resultado obtido para as opções *in-the-money*. Selecionou-se a *call* mais líquida, que esteve dentro do dinheiro no período de análise – *strike* R\$1.850/US\$1.000 e vencimento janeiro de 2008, observando-se a evolução dos prêmios de mercado e dos resultados dos modelos. Verifica-se que ambos os modelos utilizados apresentaram preços teóricos muito próximos aos preços observados no mercado. Entretanto, o modelo de Black apresentou erros mais significativos de acordo com o aumento do prazo entre a negociação e o vencimento do contrato.



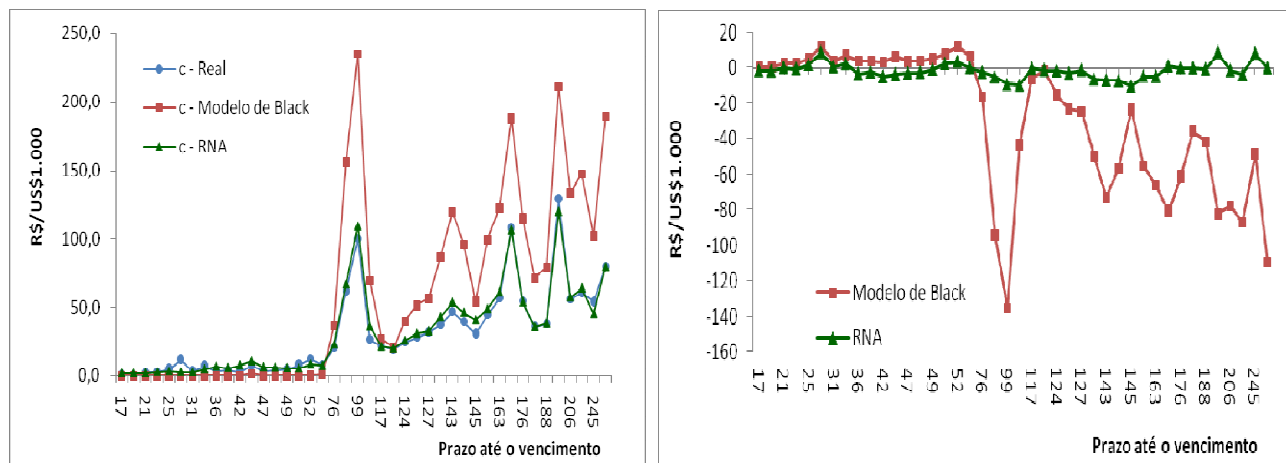
(a) Evolução dos prêmios de mercado, do modelo de Black e da RNA



(b) Diferença entre o prêmio de mercado e os modelos de Black e RNA

Gráfico 1. Comparação entre os modelos de apreçamento da amostra teste para opções *in-the-money*, com *strike* R\$ 1.850/US\$ 1.000 e vencimento em janeiro de 2008.

Para as opções *out-of-the-money*, os resultados do modelo de rede neural se destacam pela sua superioridade em relação ao modelo de Black. No Gráfico 2, como forma de ilustrar o fato acima citado, foi analisada a evolução dos prêmios de mercado e dos prêmios teóricos de uma série fora do dinheiro com alta liquidez – *strike* de R\$ 2.000 e vencimento em janeiro de 2008. É possível observar o bom desempenho obtido pelo modelo de redes neurais para a avaliação das opções estudadas. Consta-se que, para as opções com vencimentos mais curtos, os modelos teóricos apresentaram resultados similares, porém, o modelo de Black resultou em erros de apreçamento maiores de acordo com a elevação da maturidade da opção.



(a) Evolução dos prêmios de mercado, do modelo de Black e da RNA

(b) Diferença entrê o prêmio de mercado e os modelos de Black e RNA

Gráfico 2. Comparação entre os modelos de apreçamento da amostra teste para opções *out-of-the-money*, com *strike* R\$ 2.000/US\$ 1.000 e vencimento em janeiro de 2008.

Os resultados apresentados evidenciam que o modelo de rede neural proposto melhor se adéqua ao mercado de opções de compra de dólar no Brasil, em comparação ao método de Black, modelo bastante utilizado pelos agentes de mercado. O modelo de Black apresentou piores resultados nas opções com baixo grau de *moneyness*, sendo que o modelo de rede neural resultou em preços teóricos muito próximos aos preços reais para os mesmos papéis, evidenciando sua superioridade. Vale ressaltar, ainda, que uma grande vantagem da aplicação de redes neurais para a valoração de opções se encontra no fato do modelo não requerer nenhuma hipótese restritiva com relação às variáveis de entrada e demais pressuposições, uma vez que se trata de um modelo para compreender padrões lineares de alta complexidade.

4. Conclusões

O apreçamento de ativos financeiros apresenta inúmeras dificuldades para os agentes econômicos em geral, uma vez que tais preços são influenciados por choques econômicos, políticos e internacionais, que fazem com que os modelos tradicionais não sejam suficientes para explicar os valores desses ativos transacionados nos mercados, como também os próprios modelos são baseados em hipóteses que não se verificam na realidade. Dessa forma, a chamada engenharia financeira vem apresentando diversas alternativas para avaliação desses instrumentos, como por exemplo, o método de redes neurais. As redes neurais artificiais possuem a habilidade de modelar padrões não lineares e aprender com base em dados históricos. Sua aplicação para apreçamento de opções tem sido apresentada recentemente pela literatura e os resultados encorajam a utilização deste modelo, em comparação aos utilizados pelas instituições financeiras.

Este trabalho avaliou o poder de generalização das redes neurais para a valoração de opções de compra de dólar comercializadas no mercado brasileiro, BM&FBOVESPA. Foram avaliadas as opções para o período de 2004 a 2007, num total de 4.642 observações. Estruturou-se um modelo de rede neural recorrente multicamadas baseado no algoritmo *backpropagation*, método de aprendizagem supervisionado mais utilizado na literatura sobre redes neurais. As mesmas opções foram precificadas pelo modelo de Black, metodologia utilizada pelos agentes no mercado brasileiro para opções sobre moeda estrangeira. Ambos os modelos se basearam nas variáveis fundamentais que afetam o preço de uma opção: preço do ativo objeto, preço de exercício, taxa de juro livre de risco, prazo de vencimento da opção e volatilidade. A taxa de juro livre de risco foi

obtida pelo método de interpolação de *splines* cúbicos e a volatilidade estimada por um processo GARCH (1,1). Os resultados dos modelos, i.e. preços teóricos, foram comparados com os preços reais observados no mercado por meio do Erro Percentual Médio, Máximo Erro Percentual e Erro Quadrático Relativo Médio.

Os resultados mostraram que os preços das opções de compra de dólar comercializadas no mercado brasileiro são obtidos mais acuradamente por meio do modelo proposto de rede neural. Os valores relativamente baixos das métricas de erro permitiram inferir a superioridade do modelo em comparação aos resultados da fórmula analítica de Black. Entretanto a regressão dos preços de mercado sobre os preços teóricos mostrou que os preços de mercado não são fundamentalmente explicados pelos preços teóricos, o que implica erros de avaliação do mercado ou do próprio modelo e, dessa forma, pode permitir arbitragem pelos agentes.

Trabalhos futuros compreendem estudos para a estimação da volatilidade desses contratos, por meio do método de redes neurais. Pode-se estruturar um modelo de rede, baseado na volatilidade implícita dessas opções, para então aprender o processo gerador da volatilidade desses ativos e, então, pela estimação da volatilidade obtida pelo modelo de rede neural, inserir como dado de entrada em uma nova rede essa volatilidade, na tentativa de se aproximar dos preços observados nos mercados. Além disso, novos algoritmos de treinamento podem ser comparados ao *backpropagation*, assim como se testar novas estruturas na intenção de se obter melhores resultados.

5. Referências bibliográficas

- AKGIRAY, V. Conditional heteroskedasticity in time series of stock returns: evidence and forecasts. *Journal of Business*, v.61, p.55-80, 1989.
- BLACK, F; SCHOLES, M. The pricing of options and corporate liabilities. *Journal of Political Economy*, v.81, p.637-659, 1973.
- BLACK, F. The pricing of commodity contracts. *Journal of Financial Economics*, v.3, p.167-179, 1976.
- BOLLERSLEV, T. R. Generalized Autoregressive Conditional Heteroskedasticity. *Journal of Econometrics*, n.51, p.307-327, 1986.
- CALLEN, J. L.; KWAN, C. C. Y.; YIP, P. C. Y.; YUAN, Y. F. Neural network forecasting of quarterly accounting earnings. *Journal of Forecasting*, v.4, n.12, p.475-4482, 1996.
- ENDERS, W. *Applied econometrics time series*. 2. ed. Hoboken: John Wiley & Sons, 2004. 460 p.
- ENGLE, R. F. Autoregressive Conditional Heteroskedasticity with Estimates of the Variance of United Kingston Inflation. *Econometrica*, n.50, p.987-1008, 1982.
- FREITAS, S. O.; SOUZA, A. A. Utilização de redes neurais na precificação de opções. *Resenha BM&F*, n.150, p. 63-73, 2002.
- GARMAN, M. B.; KOHLHAGEN, S. W. Foreign Currency Option Values. *Journal of International Money and Finance*, v.2, p.231-237, 1983.
- GEIGLE, D. S.; ARONSON, J. E. An artificial neural network approach to the valuation of options and forecasting of volatility. *Journal of Computational Intelligence in Finance*, v.7, n.6, p.19-25, 1999.
- HAYKIN, S. *Redes Neurais: Princípios e Práticas*. 2. ed. São Paulo: Bookman, 2001.
- HULL, J. C. *Options, Futures and Other Derivatives*. 6.ed. New York: Prentice Hall, 2006.
- HUTCHINSON, J. M.; LO, A.; POGGIO, T. A nonparametric approach to pricing and hedging derivative securities via learning networks. *Journal of Finance*, v.49, p.851-889, 1994.

- JAIN, I. *Recurrent Neural Networks: Design and Applications*. Departments of Physics and Computer Science and Information Systems American University Washington, D. C.: CRC Press, 2001.
- JORION, P. *Value at Risk: A nova fonte de referência para a gestão do risco financeiro*. São Paulo: Bolsa de Mercadorias e Futuros, 2003.
- KAASTRA, I; BOYD, M. Designing a neural network for forecasting financial and economic time series. *Neurocomputing*, v.10, p. 215-236, 1995.
- KOLEN, J. *Exploring the Computational Capabilities of Recurrent Neural Networks*. Tese (Ph.d.), The Ohio State University, USA, 1994.
- LAJBCYGIER, P.; BOEK, C.; PALANISWAMI, M; FLITMAN, A. Neural networks pricing of all ordinaries SPI options on futures. In: REFENES, A. P.; ABU-MOSTAFA, Y.; MOODY, J.; WEIGEND, A.; editors. Neural networks in financial engineering, Proceedings of 3rd International Conference on Neural Networks in The Capital Markets, Oct 1995, London. World Scientific, p.64-77, 1996.
- MALLIARIS, M.; SALCHENBERGER, L. Using neural networks to forecasting the S&P 100 implied volatility. *Neurocomputing*, v.10, n.2, p.83-95, 1996.
- MERTON, R. C. Theory of rational option pricing. *Bell Journal Economics*, v.4, n.1, p.141-183, 1973.
- MONTEIRO, A.; SALLES, F. M. Interpolação da curva de juros brasileira: métodos e medidas de desempenho. *Resenha BM&F*, n.147, p.13-27, 2002.
- MORAIS, I. A. C.; PORTUGAL, M. S. Modelagem e Previsão de Volatilidade Determinística e Estocástica para a Série do Ibovespa. *Estratégia Econômica*, v.29, n.3, p.303-341, 1999.
- PETERS, E. E. *Fractal market analysis: applying chaos theory to investment and economics*. New York: John Wiley Sons, 1994.
- QI, M.; MADDALA, G. S. Option pricing using artificial neural networks: the case of S&P 500 index call options. In: REFENES, A. P.; ABU-MOSTAFA, Y.; MOODY, J.; WEIGEND, A. Neural networks in Financial engineering, Proceedings of 3rd International Conference On Neural Networks in the Capital Markets, Oct 1995, London. World Scientific, p.78-91, 1996.
- RUMELHART, D. E.; MCCLELLAND, J. L. *Parallel Distributed Processing*. Volume 1. The MIT Press, 1986.
- RUSSELL, S. J.; NORVIG, P. *Artificial Intelligence: a modern approach*. New Jersey: Prentice Hall, 1995.
- SILVA, W. S.; SÁFADI, T.; CASTRO JÚNIOR, L. G. Uma análise empírica da volatilidade do retorno de commodities agrícolas utilizando modelos ARCH: os casos do café e da soja. *Revista de Economia e Sociologia Rural*, v.43, n.1, 2005.