



Revista de Administração - RAUSP
ISSN: 0080-2107
rausp@edu.usp.br
Universidade de São Paulo
Brasil

dos Santos Maciel, Leandro; Ballini, Rosangela; Lanna Franco da Silveira, Rodrigo
Apreçamento de opções sobre taxa de câmbio R\$/US\$ negociadas no Brasil: uma comparação entre
os modelos Black e redes neurais artificiais
Revista de Administração - RAUSP, vol. 47, núm. 1, enero-marzo, 2012, pp. 96-111
Universidade de São Paulo
São Paulo, Brasil

Disponível em: <http://www.redalyc.org/articulo.oa?id=223426404007>

- ▶ Como citar este artigo
- ▶ Número completo
- ▶ Mais artigos
- ▶ Home da revista no Redalyc

Apreçamento de opções sobre taxa de câmbio R\$/US\$ negociadas no Brasil: uma comparação entre os modelos Black e redes neurais artificiais

RESUMO

No estudo aqui apresentado, aplicou-se um modelo de rede neural multicamadas para o apreçamento de *calls* sobre taxa de câmbio R\$/US\$, negociadas na Bolsa de Valores, Mercadorias & Futuros de São Paulo (BM&FBovespa), para o período de janeiro de 2004 a dezembro de 2007. A partir dos preços efetivamente praticados no mercado, comparou-se o desempenho entre essa **técnica** e o modelo de Black, utilizando-se **métricas usuais de erro** e **testes estatísticos**. Os resultados obtidos revelaram, em geral, a melhor adequação 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 (BS) (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 para apreçar opções europeias sobre ações que não pagam dividendos, sendo o prêmio de uma *call* e de 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:

- o preço da ação segue um movimento geométrico browniano, em que a distribuição de probabilidade dos preços é lognormal, com média e variância constantes;

*Leandro dos Santos Maciel
Rosangela Ballini
Rodrigo Lanna Franco da Silveira*

Recebido em 16/setembro/2010
Aprovado em 23/mai/2011

Sistema de Avaliação: *Double Blind Review*
Editor Científico: Nicolau Reinhard

DOI: 10.5700/rausp1028

Os autores agradecem ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) e à Fundação de Amparo à Pesquisa do Estado de São Paulo (Fapesp) pelo apoio financeiro.

Leandro dos Santos Maciel, Graduado em Ciências Económicas pelo Instituto de Economia da Universidade Estadual de Campinas, é Mestrando em Engenharia Elétrica no Departamento de Controle e Automação da Faculdade de Engenharia Elétrica e Computação da Universidade Estadual de Campinas (CEP 13083-852 – Campinas/SP, Brasil).
E-mail: maciel@dca.fee.unicamp.br

Endereço:
Universidade Estadual de Campinas
Faculdade de Engenharia Elétrica e Computação
Departamento de Controle e Automação
Avenida Albert Einstein, 400
13083-852 – Campinas – SP

Rosangela Ballini, Graduada em Matemática pela Universidade Federal de São Carlos, Mestre em Ciências da Computação e Matemática Computacional pela Universidade de São Paulo, Doutora em Engenharia Elétrica pela Universidade Estadual de Campinas, é Professora e Pesquisadora no Instituto de Economia da Universidade Estadual de Campinas (CEP 13083-857 – Campinas/SP, Brasil).
E-mail: ballini@eco.unicamp.br

Rodrigo Lanna Franco da Silveira, Graduado em Ciências Económicas pela Faculdade de Economia, Administração e Contabilidade da Universidade de São Paulo, Mestre e Doutor em Economia Aplicada pela Escola Superior de Agronomia “Luiz de Queiroz” da Universidade de São Paulo, é Professor e Pesquisador no Instituto de Economia da Universidade Estadual de Campinas (CEP 13083-857 – Campinas/SP, Brasil).
E-mail: rodrigolanna@eco.unicamp.br

- a taxa de juro livre de risco de curto prazo é conhecida e constante para qualquer vencimento;
- a volatilidade dos retornos do ativo objeto é constante até o vencimento;
- há possibilidade de emprestar e tomar emprestado à mesma taxa de juro livre de risco;
- inexiste o pagamento de dividendos durante a vida da opção;
- o ativo objeto do contrato é perfeitamente divisível, sendo sua negociação contínua e não possui custos de transação ou impostos;
- não se verificam oportunidades de arbitragem sem risco.

Vale observar que extensões do modelo BS foram desenvolvidas. Exemplos disso são Merton (1973), que ajustou tal modelo para a possibilidade de pagamento de dividendos nas ações, e Black (1976), que estendeu o método de forma a avaliar opções sobre contratos futuros. Por apresentar hipóteses bastante restritivas, o modelo de BS e suas extensões apresentam alguns vieses sistemáticos, bastante documentados na literatura, que levam a diferenças entre o prêmio calculado pelo modelo e aquele constatado no mercado (BENNEL e SUTCLIFFE, 2004; ANDREOU, CHARALAMBOUS e MARTZOUKOS, 2008).

Com o intuito de superar as limitações dos modelos acima citados, assim como adequar-se aos diversos tipos de contratos que surgiram, a literatura vem desenvolvendo metodologias alternativas para o apreçamento de opções. Nesse processo de evolução, modelos não paramétricos baseados em redes neurais artificiais (RNA) vêm ganhando espaço na literatura de finanças. Trata-se de uma técnica computacional inspirada no funcionamento do sistema nervoso biológico humano, motivada para a compreensão de padrões de diferentes naturezas, adequada para a modelagem de relações do tipo entrada-saída (*input-output*). As vantagens dos modelos de RNA resumem-se no fato de não dependerem das hipóteses restritivas a que os modelos paramétricos estão submetidos, além de terem a capacidade de modelar relações não lineares e de serem aplicáveis a qualquer tipo de opção, conferindo-lhes alta flexibilidade. Por outro lado, uma restrição do modelo de RNA baseia-se na necessidade de uma quantidade de dados históricos elevada, a fim de que o processo de aprendizado e generalização das relações do passado para o presente sejam eficazes (HAYKIN, 2001).

Malliaris e Salchenberger (1993) e Hutchinson, Lo e Poggio (1994) foram os pioneiros na aplicação de RNA no apreçamento de opções. Essa primeira pesquisa, que utilizou modelos de RNA e BS, obteve o prêmio das opções sobre índice de ações S&P 100 negociadas durante o primeiro semestre de 1991. Os resultados indicaram um melhor desempenho do método de redes neurais nas opções fora do dinheiro. Já o segundo estudo utilizou RNA e BS para apreçamento e aplicação da operação *delta-hedge* nas opções sobre futuros do índice S&P 500, no período de 1987 a 1991, indicando, em grande parte das análises, superioridade das redes neurais.

A partir dos dois trabalhos acima descritos, várias pesquisas foram conduzidas, sendo essas, em grande parte, direcionadas para as opções sobre ações ou índices de ações. Como exemplo, no mercado norte-americano, podem ser citados Qi e Maddala (1996), Geigle e Aronson (1999), Garcia e Gençay (2000), Ghaziri, Elfakhani e Assi (2000), Gençay e Qi (2001) e Andreou, Charalambous e Martzoukos (2008). Em geral, os autores aplicaram modelos de redes neurais para apreçamento de opções sobre o índice S&P 500, tendo resultados superiores em comparação com aqueles obtidos pela fórmula analítica de Black-Scholes.

Análises similares foram realizadas nos mercados do Reino Unido por De Freitas *et al.* (2000), Healy *et al.* (2002) e Bennel e Sutcliffe (2004), os quais apontaram melhor desempenho do método de inteligência artificial para apreçar as opções sobre o índice FTSE 100, em contraposição aos modelos tradicionais de avaliação. Vale ainda citar estudos feitos com opções sobre índices de ações negociadas nos mercados: da Alemanha – Anders, Korn e Schmitt (1998), Hanke (1999) e Ormoneit (1999); da Suécia – Amilon (2003); de Taiwan – Lin (2005); do Japão – Yao, Li e Tan (2000); da Austrália – Lajbcygier *et al.* (1996); do Brasil – Freitas e Souza (2002). A aplicação de modelos de inteligência artificial também foi realizada, em menor escala, em opções sobre taxa de juros (WHITE, 1998) e sobre taxa de câmbio (CARELLI, SILANI e STELLA, 2000).

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 essas técnicas não apresentam, necessariamente, os melhores resultados em relação aos modelos lineares de séries temporais. Com isso, é necessária a condução de estudos para diferentes mercados e com a utilização de diferentes modelos de redes neurais na tentativa de analisar se existe superioridade em relação ao modelo de Black-Scholes.

Diante desse contexto, neste artigo teve-se o objetivo de aplicar um modelo de rede neural multicamadas, com base no algoritmo de retropropagação do erro para atualização dos parâmetros, para valoração das opções de compra sobre 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 mais utilizada pelos agentes no mercado brasileiro para opções sobre moeda estrangeira. Em seguida, os resultados foram comparados com os prêmios observados no mercado com base em métricas usuais de erro e testes estatísticos. Neste trabalho, insere-se uma contribuição importante na literatura, uma vez que as abordagens usuais não costumam avaliar os resultados dos modelos com testes estatísticos, mas apenas com as métricas tradicionais de erro de previsão. Além disso, o mercado de opções brasileiro ainda apresenta poucos estudos nessa área, assim como não se verificam aplicações para o tipo de opção avaliado no presente artigo, o que amplia a relevância do es-

tudo que visa aprimorar o apreçamento de tais contratos no mercado em consideração.

2. METODOLOGIA

2.1. Amostra

A base de dados deste trabalho correspondeu aos preços de fechamento das opções europeias de compra sobre taxa de câmbio R\$/US\$, negociadas na BM&FBovespa, para diferentes prazos e preços de exercícios, no período de 02 de janeiro de 2004 a 31 de dezembro de 2007 – total de 1.001 dias de negócios e 4.642 observações⁽¹⁾. Para cada opção, obteve-se 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 opção⁽²⁾. A amostra utilizada foi composta pelos papéis que tiveram mais de 500 contratos negociados, ou seja, selecionaram-se as *calls* mais líquidas transacionadas no mercado. Nas subseções seguintes, apresentam-se 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) europeias sobre ações que não pagam dividendos, dadas pelas equações [1] e [2], respectivamente.

$$c = S \cdot N(d_1) - X \cdot e^{-r \cdot T} \cdot N(d_2) \quad [1]$$

$$p = X \cdot e^{-r \cdot T} \cdot N(-d_2) - S \cdot N(-d_1) \quad [2]$$

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 da descrição 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], em que o preço *spot*, S , é substituído pelo preço futuro, F .

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

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

em que:

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

$$d_2 = d_1 - \sigma \cdot \sqrt{T} \quad [8]$$

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, assim como modelos próprios baseados nas fórmulas de Black-Scholes ajustados às condições de mercado e inclusive modelos de RNA, são utilizados no mercado financeiro brasileiro para obtenção dos prêmios das opções de dólar. Assim sendo, no presente estudo fez-se uso do modelo Black como *benchmark*, 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); taxa de juros prefixada (ao ano) e volatilidade da opção (σ).

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

É comum verificar que, em diversas séries financeiras, a hipótese de homocedasticidade (variâncias condicionais do termo errático, u_t , constantes), adotada em vários modelos econométricos, é inapropriada. Constatou-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. Nesse 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 esse modelo apresentar resultados mais robustos, em comparação aos 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 estimativa da volatilidade do dólar foi calculada com base em um processo 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) baseia-se em sua parsimônia e em sua boa adequação aos dados. Além disso, conforme Silva, Sáfadi e Castro Júnior (2005), poucas restrições nos parâmetros são observadas nesse modelo – para a variância ser positiva e fracamente estacionária, é preciso que $\alpha_1 + \alpha_2 < 1$; $\alpha_0 > 0$; $\alpha_1 > 0$; $\alpha_2 > 0$ ⁽³⁾.

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. Considerem-se 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 ⁽⁴⁾. Neste trabalho, $g(x_j)$ é a taxa de juro pré *versus* DI dos contratos de *swap*, registrados na BM&FBovespa, para a maturidade $x_j^{(5)}$. 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:

$$\begin{aligned} P_j(x) &= a_{0,j} + a_{1,j} \cdot x + a_{2,j} \cdot x^2 + a_{3,j} \cdot x^3 \\ f(x) &= P_j(x) \text{ para cada } \{x_j, x_{j+1}\}, j = 1, \dots, n - 1. \end{aligned} \quad [10]$$

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:

$$\begin{aligned} P_j(x_j) &= g(x_j); j = 1, \dots, n - 1; \\ P_j(x_{j+1}) &= g(x_{j+1}); j = 1, \dots, n - 1; \\ P'_j(x_j) &= s_j; j = 1, \dots, n - 1; \\ P'_j(x_{j+1}) &= s_{j+1}; j = 1, \dots, n - 1; \\ P_{j-1}(x_j) &= P''_j(x_j); j = 2, \dots, n - 1. \end{aligned} \quad [11]$$

Em que s_1, \dots, s_n são parâmetros a serem delimitados pelo modelo⁽⁶⁾. Uma vez fixados os valores desses parâmetros, P_j estará determinado e a construção de f será finalizada⁽⁷⁾. Com todas as entradas do modelo de Black obtidas, o próximo passo consiste em determinar o modelo de rede neural aplicado para especificar as opções de compra de dólar consideradas.

2.3. Redes neurais artificiais

As redes neurais artificiais (RNA) são modelos computacionais não lineares que constituem sistemas de processamento distribuído paralelo que, por serem baseados no funcionamento do sistema nervoso biológico, possuem propensão natural de armazenar conhecimento experimental (HAYKIN, 2001). Assim como o cérebro humano é capaz de transmitir e/ou armazenar informações por meio de impulsos nervosos entre os neurônios, as RNAs possuem conexões interneurais, representadas por pesos sinápticos, com função de armazenar e processar o conhecimento.

Além de considerarem as não linearidades presentes entre as entradas e as saídas, as RNAs possuem características bastante atraentes. A primeira delas baseia-se no fato de não requererem *a priori* informações sobre o modelo, pois aprendem a partir de exemplos e, assim, capturam relações funcionais entre os dados, mesmo se tais relações sejam dificilmente descritíveis. Para que uma RNA possa aprender e garantir um bom desempenho no resultado, o conjunto de dados disponíveis deve ser suficientemente grande. Nesse sentido, esses modelos podem ser tratados como um método estatístico multivariado não paramétrico e não linear. Já a segunda característica tem relação com sua capacidade de generalização. Após a fase de aprendizagem, quando já possuem algum conhecimento sobre as características do conjunto de dados, as RNAs conseguem inferir sobre o comportamento da população mesmo se os dados amostrais contêm ruído nas informações (ZHANG, PATUWO e HU, 1998).

A unidade básica de processamento de uma RNA é denominada neurônio artificial. Matematicamente, a saída do neurônio é dada por:

$$y = g(\sum_{i=1}^N x_i w_i) = g(z) \quad [12]$$

sendo x_1, x_2, \dots, x_n os sinais de entrada; w_1, w_2, \dots, w_n os pesos sinápticos associados a esse neurônio; $g(\cdot)$ a função de ativação⁽⁸⁾. Neste artigo, a função de ativação adotada é a logística dada por:

$$g(z) = \frac{1}{1 + \exp^{-yz}} \quad [13]$$

na qual y representa o ponto de inflexão da função e $g(z) \in (0,1)$.

A organização dos neurônios de uma RNA é feita em camadas, sendo a primeira denominada camada de entrada e a última, camada de saída. O número de camadas intermediárias e o número de neurônios em cada uma delas são escolhidos de acordo com a maior ou menor necessidade de processar as entradas da rede. Assim sendo, os números de camadas e de neurônios aumentam, caso o conjunto de dados exija um processamento mais refinado para capturar características não lineares.

Do ponto de vista estrutural, as RNAs podem ser classificadas como redes neurais estáticas (não recorrentes) ou dinâmicas (recorrentes) (HAYKIN, 2001). A principal diferença entre as duas estruturas é a presença ou não de conexões que realimentem os neurônios presentes na estrutura da rede. Entre essas estruturas, destacam-se as redes neurais multicamadas (*Multi Layer Perceptron – MLP*), amplamente utilizadas na literatura e empregadas neste trabalho.

A estrutura (ou arquitetura) da rede MLP consiste de uma camada de entrada, uma ou mais intermediárias e outra de saída. A primeira camada, cujos neurônios são denominados unidades de entrada, difunde as entradas para as camadas seguintes sem modificação alguma. As camadas intermediárias transmitem informações por meio das conexões entre a entrada e a saída da rede. Os neurônios a ela pertencentes são denominados unidades intermediárias. Finalmente, a camada de saída transmite a resposta da rede neural à entrada mediante os neurônios denominados de unidades de saída. A figura 1 ilustra a arquitetura de uma rede neural MLP com uma camada intermediária com k neurônios e m neurônios na camada de saída (BALLINI, 2000).

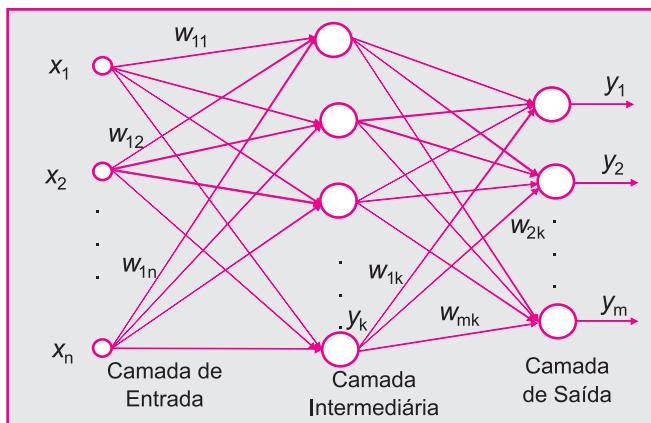


Figura 1: Estrutura da Rede Neural MLP

Para o ajuste dos parâmetros (pesos)⁽⁹⁾ da rede neural MLP, o método mais difundido e utilizado na literatura é o de retropropagação do erro (*backpropagation*) (RUMELHART e MCCLELLAND, 1986). Esse método consiste em duas fases. Na primeira, conhecida como fase de propagação direta (*forward*), as entradas são apresentadas e propagadas pela rede, camada a camada, calculando a saída de cada neurônio. Durante essa fase, os pesos são fixos e a saída calculada é comparada com a saída desejada, resultando em um erro para cada unidade de saída. Na segunda fase, o erro calculado é propagado da camada de saída para a camada de entrada, fase de propagação reversa (*backward*), e os pesos são ajustados de acordo com a regra de correção do erro, originando o termo retropropagação do

erro. A segunda fase pode ser formulada como um problema de otimização não linear irrestrita, sendo os pesos sinápticos as variáveis de decisão. O objetivo é minimizar o somatório dos erros quadráticos de todos os neurônios de saída. No algoritmo de retropropagação tradicional, para resolver o problema de otimização, adota-se o método do gradiente descendente, que considera apenas condições de primeira ordem a partir da função objetivo. Além disso, a inicialização dos pesos foi obtida de acordo com o algoritmo de geração de números aleatórios do software Matlab®. Apesar de existirem outras técnicas de inicialização dos pesos, conforme Chan, Lam e Wong (2000), os pesos aleatórios gerados inicialmente por essa técnica não impedem a convergência do algoritmo de treinamento.

Dessa forma, o objetivo do processo de treinamento é ajustar os parâmetros livres (pesos) da rede, minimizando uma função objetivo representada pelo erro quadrático médio, ou seja:

$$\min E = \min \left(\frac{1}{2N} \sum_{l=1}^N \sum_{i=1}^m (d_i^l - y_i^l)^2 \right) \quad [14]$$

sendo E a função da medida do erro total da rede, N o número de padrões⁽¹⁰⁾ apresentados à rede, m o número de neurônios na camada de saída, d_i^l a saída desejada e y_i^l a saída gerada pela rede, para o l -ésimo padrão.

O método de retropropagação aplica uma correção Δw_{ij} aos pesos sinápticos w_{ij} proporcional à direção oposta ao gradiente $\partial E / \partial \Delta w_{ij}$, definida pela regra delta como segue:

$$\Delta w_{ij} = -\eta \cdot \left(\frac{\partial E}{\partial w_{ij}} \right) \quad [15]$$

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 $(0, 1)$. Essa taxa determina a **velocidade** do passo em direção ao mínimo da função de erro E . Um valor elevado de η produzirá uma rápida convergência, resultando em grandes alterações nos pesos. Entretanto, aumentará o risco de a 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 mais iterações até a convergência⁽¹¹⁾.

2.3.1. Rede neural MLP para apreçamento de opções

Construiu-se um modelo de apreçamento para as opções de compra de dólar no Brasil usando uma rede neural MLP com base nas principais variáveis que influenciam o preço desse ativo: preço futuro do dólar no mercado (F), preço de exercício da opção (X), tempo para o vencimento da opção (T), volatilidade do preço futuro do dólar (s) e taxa de juro livre de risco (r)⁽¹²⁾. Escolheram-se essas variáveis por serem as mesmas utilizadas pelo modelo de Black para o apreçamento de opções, 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 cúbicos⁽¹³⁾.

Na literatura sobre redes neurais, não existe uma técnica que especifique qual é a estrutura ideal, ou seja, qual o número de camadas intermediárias e o número de neurônios nessa camada 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, 1996).

Neste trabalho, para a seleção da estrutura de rede neural mais adequada e para estimar os pesos sinápticos da rede neural MLP, o conjunto total de dados foi particionado em três subconjuntos denominados de: treinamento, validação e teste. O subconjunto de treinamento é usado durante a fase de estimação dos parâmetros da rede. O subconjunto de validação é empregado para avaliar a evolução do desempenho da rede durante a fase de ajuste dos parâmetros. Já o conjunto de teste é utilizado após a determinação da estrutura mais adequada, sendo esse o período propriamente dito de previsão (BALLINI, 2000). Os conjuntos treinamento, validação e teste possuem um total de 75%, 15% e 10% do total da amostra, respectivamente. O conjunto teste tem, essencialmente, a função de diagnosticar o modelo de rede estimado, uma vez que o nível de erro nessa amostra implica se o modelo está bem especificado ou não.

Assim, a partir dos conjuntos de treinamento e de validação, o número de camadas intermediárias foi determinado pelo menor erro quadrático médio. Para a determinação do número de neurônios intermediários, foram utilizados os critérios de informação e qualidade de ajuste propostos por Akaike (1974), Shibata (1976), Schwarz (1978) e Hannan e Quinn (1979), baseados no erro quadrático médio, definidos como (LAURENT e PETERS, 2001):

$$Akaike = -2 \left(\frac{\ln(EQM)}{N} \right) + 2 \left(\frac{K}{N} \right) \quad [16]$$

$$Shibata = -2 \left(\frac{\ln(EQM)}{N} \right) + \ln \left(\frac{N+2K}{N} \right) \quad [17]$$

$$Schwarz = -2 \left(\frac{\ln(EQM)}{N} \right) + 2 \left(\frac{\ln(K)}{N} \right) \quad [18]$$

$$Hannan-Quinn = -2 \left(\frac{\ln(EQM)}{N} \right) + 2 \left(\frac{K \ln(\ln(N))}{N} \right) \quad [19]$$

em que EQM representa o erro quadrático médio, $EQM = \frac{1}{N} \sum_{i=1}^N (c_i - \hat{c}_i)^2$, sendo c o preço de fechamento, observado no mercado, da opção de compra de dólar, \hat{c} o preço teórico obtido pelos modelos, N o número de observações e K o número de

neurônios na camada intermediária. Esses critérios apresentam um compromisso entre a minimização do erro quadrático médio e a complexidade da rede (número de neurônios da camada intermediária)⁽¹⁴⁾, por isso foram selecionados neste trabalho, além de serem amplamente difundidos na literatura, porém, outras técnicas podem ser verificadas em Medeiros, Teräsvirta e Rech (2006).

2.4. Métricas de erro

Os preços teóricos das opções, obtidos por meio de cada modelo, foram comparados com os preços observados no mercado. Para avaliação dos modelos, utilizaram-se o erro percentual médio (EPM), o erro percentual máximo (EP_{\max}), a raiz do erro quadrático médio (REQM) e o coeficiente de desigualdade de Theil (TIC)⁽¹⁵⁾:

$$EPM (\%) = \frac{100}{N} \sum_{i=1}^N \frac{|c_i - \hat{c}_i|}{c_i} \quad [20]$$

$$EP_{\max} (\%) = \max_{i=1, \dots, N} \left\{ 100 \frac{|c_i - \hat{c}_i|}{c_i} \right\} \quad [21]$$

$$REQM (R$/US\$) = \sqrt{\frac{1}{N} \sum_{i=1}^N (c_i - \hat{c}_i)^2} \quad [22]$$

$$TIC = \frac{\sqrt{\sum_{i=1}^N (\hat{c}_i - c_i)^2}}{\sqrt{\sum_{i=1}^N \frac{\hat{c}_i^2}{N}} + \sqrt{\sum_{i=1}^N \frac{c_i^2}{N}}} \quad [23]$$

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. Cabe observar que EPM e EP_{\max} são descritos em termos percentuais e REQM em termos absolutos.

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

$$M = \frac{S}{X \cdot e^{-r \cdot T}} \quad [24]$$

sendo S a cotação da PTAX. Nesse caso, as opções da amostra teste foram divididas conforme o grau de *moneyness*: *out-of-the-money* ($M \leq 1 - \alpha\%$), *at-the-money* ($1 - \alpha\% < M < 1 + \alpha\%$) 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:

$$c_i = \phi_1 + \phi_2 \cdot \hat{c}_i + \zeta \quad [25]$$

na qual ϕ_1 e ϕ_2 são os parâmetros linear e angular da regressão, respectivamente, e ζ um ruído branco.

Os resultados empíricos na literatura baseiam-se extensivamente em medidas padrões de acurácia, como as métricas de erro apresentadas (EPM, EP_{max}, REQM, TIC), e não levam em conta a superioridade de um modelo em relação a outro em termos de significância estatística. Assim, foram aplicados testes paramétricos (AGS e MGN) e não paramétricos (SIGN) para verificar se existe diferença estatisticamente significativa entre os resultados dos modelos avaliados. Esses testes são apresentados a seguir.

2.4.1. Teste AGS

O teste paramétrico aqui descrito baseia-se no trabalho de Ashley, Granger e Schmalensee (1980), denominado teste AGS. Esse teste permite a avaliação da existência de significância estatística entre a diferença do erro quadrático médio (EQM) entre modelos de previsão. Seja $\Delta e_i^{B,RNA}$ a diferença entre os erros de previsão obtidos pelos modelos de Black e de redes neurais, e_i^B e e_i^{RNA} , respectivamente. Assim:

$$\Delta e_i^{B,RNA} = e_i^B - e_i^{RNA} = (c_i^B - \hat{c}_i^B) - (c_i^{RNA} - \hat{c}_i^{RNA}) \quad [26]$$

sendo $i = 1, \dots, N$. Define-se, ainda, $S_i^{B,RNA}$ a soma dos erros de previsão e μ_S a média amostral de $S_i^{B,RNA}$. O teste pode ser descrito segundo a equação:

$$\Delta e_i^{B,RNA} = \beta_1 + \beta_2 (S_i^{B,RNA} - \mu_S) + v_i \quad [27]$$

na qual v_i é um processo de ruído branco, que se resume em uma sequência de erros aleatórios com média e variância constantes.

O teste AGS mostra que β_1 representa a diferença entre os erros quadráticos médios de previsão obtidos pelos modelos e β_2 é proporcional à diferença entre as variâncias dos erros de previsão dos modelos. O teste de significância tem como hipótese nula $H_0: \beta_1 = \beta_2 = 0$, enquanto a hipótese alternativa consiste em $H_1: \beta_1 > 0$ e/ou $\beta_2 > 0$. A estatística para este teste, obtida por meio da estimativa da equação [27], tem distribuição F com 2 e $(N - 2)$ graus de liberdade, assumindo normalidade para os erros⁽¹⁶⁾. A rejeição da hipótese nula indica que o modelo rede neural supera o método de Black.

2.4.2. Teste MGN

Nesta subseção, é descrito outro teste paramétrico para igualdade entre acurácia de previsão. Trata-se do teste de Morgan-Granger-Newbold (teste MGN), baseado inicialmente no trabalho de Granger e Newbold (1977). A utilização desses testes é recomendada para avaliar modelos de previsão e quando se relaxa a assunção de não existência de correlação

entre os erros dos modelos. O teste estatístico para esse teste pode ser calculado como:

$$MGN = \frac{\hat{\rho}_{s\Delta}}{(1 - \hat{\rho}_{s\Delta}^2)^{1/2}} (N - 1)^{1/2} \quad [28]$$

em que $\hat{\rho}_{s\Delta}$ é o coeficiente de correlação estimado entre $S_i^{B,RNA}$ e $\Delta e_i^{B,RNA}$, e N representa o número de observações. A estatística para o teste MGN tem distribuição t com $(N - 1)$ graus de liberdade. Para esse teste, se as previsões são igualmente acuradas, então a correlação entre $S_i^{B,RNA}$ e $\Delta e_i^{B,RNA}$ será zero (hipótese nula).

2.4.3. Teste SIGN

O *Significance Test* (teste SIGN), derivado do trabalho de Lehmann (1988), é uma avaliação não paramétrica para modelos de previsão que não requer suposições como normalidade e não correlação serial dos erros. A estatística para esse teste pode ser calculada como:

$$SIGN = \left(\psi(N) - \frac{N}{2} \right) / 2(N)^{1/2} \quad [29]$$

o termo ψ_t denotando o número de vezes em que, neste caso, os erros do modelo de Black superaram aqueles derivados do modelo de rede neural. A hipótese nula é $H_0: \psi(N) = (N/2)$, contra a hipótese alternativa de que $H_1: \psi(N) > (N/2)$. Assume-se que N é o número de observações previstas por ambos os modelos e, se os erros de previsão do modelo de Black são da mesma proporção que os do modelo de rede neural em exatamente $(N/2)$ observações, de acordo com o teste SIGN, a acurácia dos modelos pode ser considerada equivalente, sem diferença estatisticamente significativa. Entretanto, se a proporção de erros do modelo de Black supera os do modelo de rede neural numa proporção maior que $(N/2)$, o teste indica que o modelo de Black é menos acurado que o modelo de rede neural. A estatística para o teste SIGN tem distribuição Normal com média zero e variância unitária. Assim, se a estatística SIGN é significativamente elevada, a hipótese nula de equivalência de previsão pode ser rejeitada em favor da hipótese alternativa.

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 pertenceram ao grupo de teste da estruturação da rede neural, uma vez que o conjunto treinamento e validação foram utilizados para determinar a estrutura de rede mais adequada, estimar os pesos sinápticos e aprender o processo gerador dos preços de compra das opções de dólar. Dessa forma,

o conjunto teste, ou conjunto de previsão, permitiu avaliar a capacidade de aprendizado e generalização da rede para prever as opções com dados de entrada nunca antes apresentados à rede. O conjunto teste foi composto por negociações que se iniciam em 04 de outubro de 2007 e vão até 31 de dezembro de 2007, totalizando 466 observações.

Na tabela 1, apresentam-se os resultados de escolha do número de neurônios intermediários. De acordo com essa tabela, nota-se que os critérios de informação mostraram que a rede composta por seis neurônios na camada intermediária apresentou os melhores resultados, exceto para o critério de Shibata, que indicou cinco neurônios. Verificou-se que, com o aumento do número de neurônios, a complexidade da rede aumentava, mas seu desempenho declinava. Portanto, construiu-se um modelo de rede neural multicamadas, constituído por cinco neurônios compondo a camada de entrada, seis neurônios na camada intermediária e uma camada de saída com um neurônio, representando o preço da opção. Na figura 2, apresenta-se o modelo de rede neural aplicado para o apreçamento das opções avaliadas.

O número de épocas (iterações) foi 300⁽¹⁷⁾, com base na aprendizagem padrão a padrão, com um nível de erro de 10^{-4} , uma vez que são níveis de erro aceitáveis para um modelo de estimativa e previsão de séries temporais financeiras. Por fim, a taxa de aprendizagem foi fixada em 0,6.

Na tabela 2 são apresentados os resultados dos valores das métricas de erro, definidas na seção 2.4, para os modelos teóricos avaliados, assim como o coeficiente de determinação da regressão dos preços do mercado sobre os preços teóricos.

Por meio da avaliação dos erros (tabela 2), é possível verificar que o modelo de redes neurais apresentou preços teóricos mais próximos aos preços de mercado, explicando o baixo EPM, EP_{max} e REQM, em comparação ao modelo de Black. O coeficiente de desigualdade de Theil (TIC) mostrou um ajuste superior do modelo de rede neural aos dados das opções, pois apresentou um valor bem mais próximo de zero em comparação ao modelo de Black. Adicionalmente, o coeficiente de determinação da equação [25], regressão dos preços de mercado sobre os preços teóricos, confirmou as análises expostas acima, dado o maior R^2 quando utilizados os prêmios gerados pelo modelo de RNA. Nessa última análise, cabe notar que os preços do mercado não foram explicados fundamentalmente pelos preços teóricos, o que leva a concluir que outros fatores estariam determinando os prêmios dessas opções.

De forma complementar, na tabela 3 apresentam-se os valores das métricas de erro para toda a amostra (4.642 observações), divididas nos conjuntos de treinamento, validação e teste⁽¹⁸⁾. Os resultados mostram que o baixo desempenho do modelo de Black foi observado em toda a amostra dos dados.

Tabela 1

Critérios de Informação para Determinação da Estrutura da Rede Neural para Apreçamento das Opções de Compra de Dólar

K	Akaike	Shibata	Schwarz	Hannan-Quinn
1	4,19987	4,19237	4,12367	4,21745
2	4,28612	4,01838	4,28613	4,19971
3	4,30183	4,11831	4,16523	4,28415
4	4,27616	4,09171	4,09522	4,17673
5	4,23677	4,00183*	4,07612	4,22217
6	4,16775*	4,00813	4,06263*	4,14316*
7	4,36712	4,07213	4,19028	4,15713
8	4,74322	4,10089	4,09235	4,14755
9	4,28126	4,18327	4,20013	4,35619
10	4,50132	4,46721	4,12832	4,37563
11	4,63125	4,36215	4,28361	4,40001
12	4,55153	4,28631	4,19716	4,38662
13	4,67524	4,42877	4,21326	4,38372
14	4,46532	4,37626	4,17632	4,45165
15	4,66823	4,64126	4,57923	4,67163

Nota: * Denota o melhor modelo.

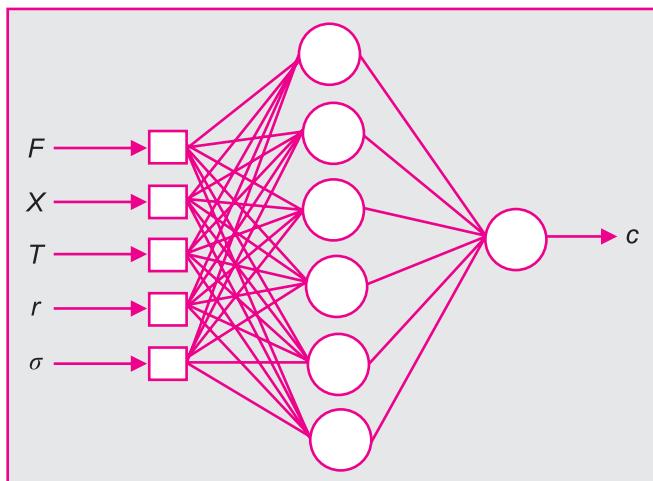


Figura 2: Rede Neural MLP para Apreçamento das Opções de Compra de Dólar

Os erros desse modelo apresentaram-se relativamente iguais nos três conjuntos de dados, uma vez que não faz diferença a distinção entre eles para a aplicação da fórmula analítica. Entretanto, como era de esperar, nos conjuntos treinamento e validação, para a rede neural, os erros foram menores, 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.

Vale observar que as médias dos erros para os dois modelos foram elevadas e significativamente diferentes de zero para o período teste (04 de outubro de 2007 a 31 de dezembro de 2007). Além disso, algumas previsões apresentaram erros relativos superiores a 100%, de acordo com o EP_{max} , 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, 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 estavam dentro do dinheiro, enquanto 31% e 26% estão no dinheiro e fora do dinheiro, respectivamente. Na tabela 4, apresentam-se os resultados das métricas de erro para o conjunto da amostra teste de acordo

Tabela 2

Aderência dos Preços Calculados pelos Modelos Black e Redes Neurais aos Preços de Fechamento das Calls de Dólar Europeias Negociadas na BM&FBovespa entre 04 de outubro de 2007 e 31 de dezembro de 2007 – Amostra com 466 Observações

Modelos	Métricas de Erro				
	EPM	EP_{max}	REQM	TIC	R^2
RNA	1,21	3,13	0,46	0,16	0,79
Black	4,32	7,86	2,23	0,51	0,66

Tabela 3

Aderência dos Preços Calculados pelos Modelos Black e Redes Neurais aos Preços de Fechamento das Opções de Compra de Dólar Europeias Negociadas na BM&FBovespa entre 02 de janeiro de 2004 e 31 de dezembro de 2007 – Amostra com 4.642 Observações

Modelos	Treinamento			Validação			Teste		
	EPM	EP_{max}	REQM	EPM	EP_{max}	REQM	EPM	EP_{max}	REQM
RNA	0,15	1,08	0,01	0,11	0,89	0,03	1,09	2,03	0,46
Black	3,42	6,23	2,46	2,82	4,44	2,32	3,66	5,21	2,23

com o grau de *moneyness* das opções. Ambos os modelos, de Black e de rede neural, foram mais acurados no apreçamento das opções *in-the-money*. Os maiores níveis de erros 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 perceber pelas métricas do EPM e do EP_{max}. Nota-se, de forma geral, nos diferentes graus de *moneyness*, que os resultados do modelo de RNA foram superiores em comparação ao modelo de Black.

Para as opções *out-of-the-money*, os resultados do modelo de rede neural destacaram-se por sua superioridade em relação ao modelo de Black. No gráfico 1 (partes a e b), página 106, 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. Constata-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.

Apesar dos resultados favoráveis ao modelo de inteligência artificial, é necessário verificar se a diferença entre os erros obtidos pelos modelos é estatisticamente significativa. A tabela 5 elucida essa análise por meio dos resultados dos testes AGS, MGN e SIGN.

O resultado do teste AGS, que examina os erros de previsão em termos do erro quadrático médio dos modelos avaliados (sendo o modelo de Black tomado como *benchmark*), indicou que o modelo de rede neural apresentou superioridade significativa na precificação das opções de compra de dólar no Brasil. O nível de significância apresentado (0,0001) indicou a rejeição da hipótese nula, ou seja, os erros do modelo de rede neural foram inferiores em comparação aos do modelo de Black, em termos estatísticos. Contudo, o resultado do teste MGN não apresentou evidência significativa da melhor adequação do modelo de rede neural. Neste caso, a significância do teste MGN permitiu aceitar a hipótese nula e inferir que os modelos foram igualmente acurados. Por fim, o teste não paramétrico (SIGN) mostrou a superioridade do modelo de rede neural para a precificação das opções, rejeitando a equivalência de predição dos modelos.

Esses testes foram também conduzidos nos diferentes graus de *moneyness* das opções, conforme consta na tabela 6.

Tabela 4

Aderência dos Preços Calculados pelos Modelos Black e Redes Neurais aos Preços de Fechamento das Opções de Compra de Dólar Europeias Negociadas na BM&FBovespa entre 04 de outubro de 2007 e 31 de dezembro de 2007 de acordo com o Grau de Moneyness – Amostra com 466 Observações

Modelos	Out-of-the-Money			At-the-Money			In-the-Money		
	EPM	EP _{max}	REQM	EPM	EP _{max}	REQM	EPM	EP _{max}	REQM
RNA	1,34	2,76	0,52	1,02	2,11	0,44	1,21	2,51	0,38
Black	3,63	6,02	2,64	1,97	3,25	2,09	3,07	5,89	1,89

Tabela 5

Testes AGS, MGN e SIGN para Avaliação do Modelo de Rede Neural na Precificação das Opções de Compra de Dólar Europeias Negociadas na BM&FBovespa entre 04 de outubro de 2007 e 31 de dezembro de 2007, Tendo como Benchmark o Modelo de Black – Amostra com 466 Observações

Modelo	Teste AGS (F _{2,464}) [*]	Teste MGN (t ₄₆₅) [*]	Teste SIGN N(0,1) [*]
RNA	104,0987 (0,0000)	0,36249 (0,7170)	-2,74357 (0,0030)

Nota: *O nível de significância para os testes foi de 5%. Suas estatísticas estão representadas entre parênteses.

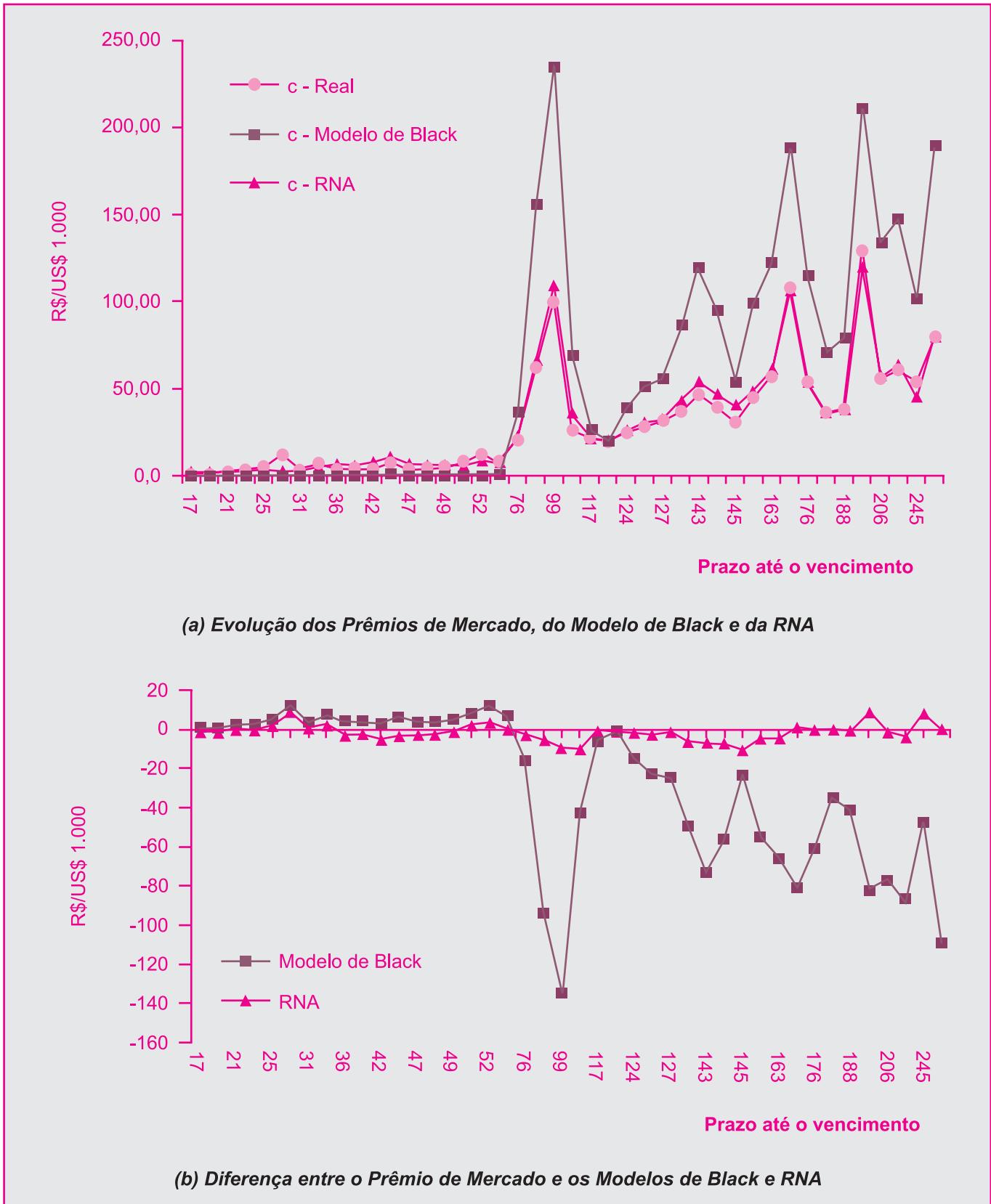


Gráfico 1: Comparação entre os Modelos de Apreciação da Amostra Teste para Opções Out-of-the-Money, com Strike R\$ 2.000/US\$ 1.000 e Vencimento em Janeiro de 2008

Tabela 6

Testes AGS, MGN e SIGN para Avaliação do Modelo de Rede Neural na Precificação das Opções de Compra de Dólar Europeias Negociadas na BM&FBOVESPA entre 02 de janeiro de 2004 e 31 de dezembro de 2007 de Acordo com o Grau de Moneyness – Amostra com 466 Observações

Moneyness	Teste AGS ($F_{(2,264);(2,112);(2,84)}^*$)	Teste MGN ($t_{(265);(113);(85)}^*$)	Teste SIGN $N(0,1)$
Out-of-the-money	234,7329 (0,0000)	235,98216 (0,0000)	-1,27707 (0,1007)
At-the-money	156,8730 (0,0001)	1,28364 (0,1999)	-3,22615 (0,0006)
In-the-money	123,8622 (0,0001)	0,55438 (0,5795)	-3,26983 (0,0005)

Notas: *Os graus de liberdade entre parênteses são correspondentes às amostras das opções *out-of-the-money*, *at-the-money* e *in-the-money*, respectivamente.

O nível de significância para os testes foi de 5%, e suas estatísticas estão representadas entre parênteses.

Os resultados obtidos confirmaram as observações anteriores, exceto para as opções fora do dinheiro. O teste AGS indicou a melhor adequação do modelo de redes neurais para a precificação das opções em questão para todos os graus de *moneyness*, apresentando níveis de significância baixos, que indicam a rejeição da hipótese nula. O mesmo resultado verificou-se com o teste SIGN. Já o teste MGN indicou que os resultados do método de inteligência artificial foram estatisticamente mais acurados que os do modelo de Black apenas para as opções *out-of-the-money*.

Os resultados apresentados evidenciam que o modelo de rede neural proposto ajusta-se melhor 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*, e o modelo de rede neural resultou em preços teóricos muito próximos aos preços reais para os mesmos papéis, evidenciando sua superioridade.

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 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 pela literatura e os resultados encorajam a utilização desse modelo, em comparação aos modelos mais tradicionais como o de Black-Scholes e seus derivados.

Nesse sentido, o objetivo neste trabalho foi analisar o apreçamento de *calls* sobre taxa de câmbio R\$/US\$, negociadas na BM&FBovespa, mediante o uso de técnica de redes neurais, comparando os resultados com os do modelo de Black. Foram avaliadas as opções para o período de 2004 a 2007, num total de 4.642 observações. Estruturou-se um modelo de rede neural multicamadas com algoritmo *backpropagation* para atualização dos pesos sinápticos. As mesmas opções foram precificadas pelo modelo de Black, considerado como *benckmark*. Os modelos consideraram as 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, isto é, preços teóricos, foram comparados com os preços reais observados no mercado por meio do erro percentual médio, erro percentual máximo e raiz do erro quadrático médio e do coeficiente de desigualdade de Theil, assim como com testes estatísticos paramétricos (AGS e MGN) e não paramétricos (SIGN).

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. Os testes estatísticos AGS e SIGN indicaram a superioridade do modelo de rede neural em termos estatísticos, com nível de significância de 5%, em relação à fórmula analítica de Black. O teste MGN, por sua vez, inferiu igualdade para o resultado dos métodos, exceção feita às opções *out-of-the-money*.

Deve-se observar que os resultados se limitam a uma amostra de um período específico, o que resulta em algumas limitações e, além disso, não foram considerados no modelo Black ajustes de assimetria e curtose por meio de um *smile* de volatilidade implícita nos preços negociados no mercado. Trabalhos futuros podem avançar na estimativa da volatilidade desses contratos, por meio de abordagens de modelos de inteligência artificial, tais como redes neurais. Assim, esses resultados poderiam ser comparados com estimativas da volatilidade obtidas por meio de modelos de tipo EGARCH ou de volatilidade estocástica, por exemplo, e também com modelos derivados das fórmulas de BS ajustadas às condições de mercado.♦

NOTAS

- (1) A escolha desse período justifica-se pelo fato de a taxa de câmbio R\$/US\$ apresentar relativa estabilidade, de forma a não influenciar os resultados dos modelos de precificação. Deve-se acrescentar ainda que, apesar de tratar-se de uma amostra considerável, uma limitação dos modelos de RNAs é que eles podem ajustar-se bem à amostra e não à população.
- (2) Os dados foram fornecidos pela BM&FBovespa.
- (3) Outros modelos para estimar a volatilidade poderiam ser usados, tais como modelos EGARCH e de volatilidade estocástica, porém, devido à grande popularidade na literatura de modelos GARCH, optou-se por tal metodologia. Além disso, não é o objetivo neste artigo comparar diferentes metodologias para estimativa da volatilidade para essas opções.
- (4) Caso z esteja fora do intervalo definido por a e b , o problema é dito extrapolação.
- (5) Os dados dos contratos foram fornecidos pela BM&FBovespa.
- (6) Essa modelagem foi feita no *software* Matlab®, assim como todas as aplicações deste trabalho.
- (7) Para mais detalhes do método de *splines* cúbicos, ver o trabalho de Monteiro e Salles (2002).
- (8) As principais funções de ativação são: linear, sigmoide ou logística, degrau bipolar e hiperbólica. Destaque-se que as funções de ativação devem ser contínuas.
- (9) A etapa de ajuste dos pesos da rede neural é conhecida como processo de aprendizagem ou período de treinamento da rede neural.
- (10) Número de padrões ou número de pares entrada-saída desejada.
- (11) Uma apresentação mais detalhada do algoritmo de *backpropagation* pode ser vista em Haykin (2001).
- (12) Apesar dos benefícios apontados por Azoff (1994) em normalizar as variáveis de entrada, neste trabalho não se realizou tal procedimento, uma vez que ele não afetou os resultados obtidos, assim, para poupar cálculos desnecessários, as entradas foram mantidas em nível.
- (13) Todos os dados inseridos, tanto no modelo de Black quanto no modelo de rede neural proposto, estão em termos anuais. 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.
- (14) A complexidade da rede poderia também ser representada pelo número de variáveis de entrada. Entretanto, para análise de comparação, definiu-se *a priori* que as entradas seriam as mesmas do modelo de Black, ou seja, preço futuro, preço de exercício, maturidade do contrato, taxa de juros livre de risco e volatilidade do preço futuro.
- (15) O valor do TIC varia entre zero e um, em que zero indica o ajuste perfeito.
- (16) Os procedimentos do teste AGS também são descritos em Bessler e Brandt (1992), Kastens e Brester (1996) e Kiani, Bidarkota e Kastens (2005).
- (17) Denomina-se de época a apresentação de todos os dados do conjunto de treinamento durante o

NOTAS

processo de aprendizagem. Neste trabalho, para a atualização dos pesos foi empregada a chamada aprendizagem padrão a padrão, ou seja, a cada par de entrada-saída apresentado à rede, os pesos são atualizados. Nesse caso, para um conjunto de treinamento com N pares entrada-saída, em uma época, é realizada a atualização de pesos N vezes. Ao final de uma época, o erro quadrático médio é calculado e, se o valor for menor que o nível de

erro estipulado, o processo de aprendizagem é encerrado. Caso contrário, inicia-se uma nova época.

- (18) O conjunto de treinamento foi composto pelos dados no período de 02 de janeiro de 2004 a 11 de maio de 2007, o conjunto de validação de 14 de maio de 2007 a 03 de outubro de 2007 e, por fim, a amostra teste de 04 de outubro de 2007 a 31 de dezembro de 2007.

REFERÊNCIAS

- AKAIKE, H. A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, New York, v.19, n.6, p.716-723, Dec. 1974.
[DOI: <http://dx.doi.org/10.1109/TAC.1974.1100705>]
- AMILON, H. A neural network versus Black-Scholes: a comparison of pricing and hedging performances. *Journal of Forecasting*, Hoboken, New Jersey, v.22, n.4, p.317-335, July 2003.
- ANDERS, U.; KORN, O.; SCHMITT, C. Improving the pricing of options: a neural network approach. *Journal of Forecasting*, Hoboken, New Jersey, v.17, n.5-6, p.369-388, Sept./Nov. 1998.
- ANDREOU, P.C.; CHARALAMBOUS, C.; MARTZOUKOS, S.H. Pricing and trading European options by combining artificial neural networks and parametric models with implied parameters. *European Journal of Operational Research*, v.185, n.3, p.1415-1433, Mar. 2008.
[DOI: <http://dx.doi.org/10.1016/j.ejor.2005.03.081>]
- ASHLEY, R.; GRANGER, C.W.J.; SCHMALENSEE, R. Advertising and aggregate consumption: an analysis of causality. *Econometrica*, Chicago, v.48, n.5, p.1149-1167, July 1980.
[DOI: <http://dx.doi.org/10.2307/1912176>]
- AZOFF, E.M. *Neural Network time series forecasting of financial markets*. Chichester: John Wiley and Sons, 1994. 196p.
- BALLINI, R. *Análise e previsão de vazões utilizando modelos de séries temporais, redes neurais e redes neurais nebulosas*. 2000. Tese (Doutorado) – Universidade Estadual de Campinas, Campinas, São Paulo, Brasil, 2000.
- BENNEL, J.; SUTCLIFFE, C. Black-Scholes versus artificial neural networks in pricing FTSE 100 options. *Intelligent Systems in Accounting, Finance and Management*, Hoboken, New Jersey, v.12, n.4, p.243-260, Mar. 2004.
[DOI: <http://dx.doi.org/10.1002/isaf.254>]
- BESSLER, D.A.; BRANDT, J.A. An analysis of forecasts of livestock prices. *Journal of Economic Behavior & Organization*, v.18, n.2, p.249-263, July 1992.
[DOI: [http://dx.doi.org/10.1016/0167-2681\(92\)90030-F](http://dx.doi.org/10.1016/0167-2681(92)90030-F)]
- BLACK, F. The pricing of commodity contracts. *Journal of Financial Economics*, v.3, n.1-2, p.167-179, Jan./Mar. 1976.
[DOI: [http://dx.doi.org/10.1016/0304-405X\(76\)90024-6](http://dx.doi.org/10.1016/0304-405X(76)90024-6)]
- BLACK, F.; SCHOLES, M. The pricing of options and corporate liabilities. *Journal of Political Economy*, Chicago, v.81, n.3, p.637-659, May 1973.
- BOLLERSLEV, T.R. Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, v.31, n.51, p.307-327, June 1986.
[DOI: [http://dx.doi.org/10.1016/0304-4076\(86\)90063-1](http://dx.doi.org/10.1016/0304-4076(86)90063-1)]
- CALLEN, J.L.; KWAN, C.C.Y.; YIP, P.C.Y.; YUAN, Y.F. Neural network forecasting of quarterly accounting earnings. *Journal of Forecasting*, Hoboken, New Jersey, v.12, n.4, p.475-482, Dec. 1996.
- CARELLI, A.; SILANI, S.; STELLA, F. Profiling neural networks for option pricing. *International Journal of Theoretical and Applied Finance*, v.3, n.2, p.183-204, Apr. 2000.
[DOI: <http://dx.doi.org/10.1142/S0219024900000097>]
- CHAN, M.C.; LAM, C.C.; WONG, C.C. Financial time series forecasting by neural network using conjugate gradient learning algorithm and multiple linear regression weight initialization. *Computing in Economics and Finance*, Hong Kong, n.61, 2000.
- DE FREITAS, J.F.G.; NIRANJAN, M.A.; GEE, A.H.; DOUCET, A. Sequential Monte Carlo methods on train neural network models. *Neural Computation*, Cambridge, v.12, n.4, p.955-993, Apr. 2000.
[DOI: <http://dx.doi.org/10.1162/089976600300015664>]
- ENDERS, W. *Applied econometrics time series*. 2nd ed. Hoboken: John Wiley & Sons, 2004. 460p.
- ENGLE, R.F. Autoregressive conditional heteroskedasticity with estimates of the variance of United Kingdom inflation. *Econometrica*, Chicago, v.50, n.4, p.987-1008, July 1982.
[DOI: <http://dx.doi.org/10.2307/1912773>]

REFERÊNCIAS

- FREITAS, S.O.; SOUZA, A.A. Utilização de redes neurais na especificação de opções. *Resenha BM&F*, São Paulo, n.150, p.63-73, 2002.
- GARCIA, R.; GENÇAY, R. Pricing and hedging derivative securities with neural networks and a homogeneity hint. *Journal of Econometrics*, v.94, n.1-2, p.93-115, Jan. 2000. [DOI: [http://dx.doi.org/10.1016/S0304-4076\(99\)00018-4](http://dx.doi.org/10.1016/S0304-4076(99)00018-4)]
- GARMAN, M.B.; KOHLHAGEN, S.W. Foreign currency option values. *Journal of International Money and Finance*, v.2, n.3, p.231-237, Dec. 1983. [DOI: [http://dx.doi.org/10.1016/S0261-5606\(83\)80001-1](http://dx.doi.org/10.1016/S0261-5606(83)80001-1)]
- 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, Nov./Dec. 1999.
- GENÇAY, R.; QI, M. Pricing and hedging derivative securities with neural networks: bayesian regularization, early stopping and bagging. *IEEE Transactions on Neural Networks*, v.12, n.4, p.726-734, July 2001. [DOI: <http://dx.doi.org/10.1109/72.935086>]
- GHAZIRI, H.; ELFAKHANI, S.; ASSI, J. Neural networks approach to pricing options. *Neural Network World*, Prague, v.10, n.1-2, p.271-277, 2000.
- GRANGER, C.W.J.; NEWBOLD, P. *Forecasting economic time series*. New York: Academic Press, 1977. 333p.
- HANKE, M. Neural networks versus Black-Scholes: an empirical comparison of the pricing accuracy of two fundamentally different option pricing methods. *Journal of Computational Intelligence in Finance*, v.7, n.1, p.26-34, Jan./Feb. 1999.
- HANNAN, E.J.; QUINN, B.G. The determination of the order of an autoregression. *Journal of the Royal Statistical Society, Hoboken*, New Jersey, v.41, n.2, p.190-195, 1979.
- HAYKIN, S. *Redes neurais: princípios e práticas*. 2.ed. São Paulo: Bookman, 2001. 900p.
- HEALY, J.; DIXON, M.; READ, B.; CAI, F.F. A data-centric approach to understanding the pricing of financial options. *The European Physical Journal B – Condensed Matter and Complex Systems*, Berlin, v.27, n.2, p.219-227, May 2002.
- HUTCHINSON, J.M.; LO, A.; POGGIO, T. A nonparametric approach to pricing and hedging derivative securities via learning networks. *Journal of Finance*, Berkeley, California, v.49, n.3, p.851-889, July 1994.
- JORION, P. *Value at risk: a nova fonte de referência para a gestão do risco financeiro*. São Paulo: Bolsa de Mercadorias & Futuros, 2003. 487p.
- KAASTRA, I.; BOYD, M. Designing a neural network for forecasting financial and economic time series. *Neurocomputing*, v.10, n.3, p.215-236, Apr. 1996. [DOI: [http://dx.doi.org/10.1016/0925-2312\(95\)00039-9](http://dx.doi.org/10.1016/0925-2312(95)00039-9)]
- CASTENS, T.L.; BRESTER, G.W. Model selection and forecasting ability of theory-constrained food demand systems. *American Journal of Agricultural Economics*, Oxford, v.78, n.2 p.301-312, May 1996. [DOI: <http://dx.doi.org/10.2307/1243704>]
- KIANI, K.; BIDARKOTA, P.; KASTENS, T. Forecast performance of neural networks and business cycle asymmetries. *Applied Financial Economics Letters*, London, v.1, n.4, p.205-210, July 2005. [DOI: <http://dx.doi.org/10.1080/17446540500143848>]
- LAJBCYGIER, P.; BOEK, C.; PALANISWAMI, M.; FLITMAN, A. Comparing conventional and artificial neural network models for the pricing of options on futures. *Neurovest Journal*, v.4, n.5, p.16-24, Sept./Oct. 1996.
- LAURENT, S.; PETERS, J. GARCH 2.0: an ox package for estimating and forecasting various ARCH models. In: FORECASTING FINANCIAL MARKETS, 8., 2001, London. Proceedings... United Kingdom, 2001.
- LEHMANN, E.L. *Nonparametric statistical methods based on ranks*. Upper Saddle River, New Jersey: Prentice Hall, 1988. 463p.
- LIN, C. The valuation of Taiwan stock index option price – comparison of performances between Black-Scholes and neural network model. *Journal of Statistics & Management Systems*, Amsterdam, v.8, n.2, p.355-367, July 2005.
- MALLIARIS, M.; SALCHENBERGER, L. A neural network model for estimating option prices. *Journal of Applied Intelligence*, Berlin, v.3, n.3, p.193-206, Sept. 1993. [DOI: <http://dx.doi.org/10.1007/BF00871937>]
- MEDEIROS, M.C.; TERÄSVIRTA, T.; RECH, G. Building neural network models for time series: a statistical approach. *Journal of Forecasting*, Hoboken, New Jersey, v.25, n.1, p.49-75, Jan. 2006.
- MERTON, R.C. Theory of rational option pricing. *Bell Journal Economics*, Santa Monica, California, v.4, n.1, p.141-183, Spring 1973. [DOI: <http://dx.doi.org/10.2307/3003143>]
- MONTEIRO, A.; SALLES, F.M. Interpolação da curva de juros brasileira: métodos e medidas de desempenho. *Resenha BM&F*, São Paulo, 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. *Estudos Econômicos*, São Paulo, v.29, n.3, p.303-341, jul./set. 1999.

REFERÊNCIAS

- ORMONEIT, D. A regularization approach to continuous learning with an application to financial derivatives pricing. *Neural Networks*, v.12, n.10, p.1405-1412, Dec. 1999.
[DOI: [http://dx.doi.org/10.1016/S0893-6080\(99\)00078-7](http://dx.doi.org/10.1016/S0893-6080(99)00078-7)]
- 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. (Ed.). *Neural networks in financial engineering: proceedings of the Third International Conference on Neural Networks in the Capital Markets*. New York: World Scientific, 1996. p.78-91.
- RUMELHART, D.E.; MCCLELLAND, J.L. *Parallel distributed processing*. Massachusetts: The MIT Press, 1986. v.1, 611p.
- SCHWARZ, G.E. Estimating the dimension of a model. *The Annals of Statistics*, v.6, n.2, p.461-464, Mar. 1978.
[DOI: <http://dx.doi.org/10.1214/aos/1176344136>]
- SHIBATA, R. Selection of the order of an autoregressive model by Akaike's information criterion. *Biometrika*,
- Oxford, v.63, n.1, p.117-126, Apr. 1976.
[DOI: <http://dx.doi.org/10.1093/biomet/63.1.117>]
- 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*, Brasília, v.43, n.1, p.119-134, jan./mar. 2005.
- WHITE, A.J. A genetic adaptive neural network approach to pricing options: a simulation analysis. *Journal of Computational Intelligence in Finance*, v.6, n.2, p.13-23, Mar./Apr. 1998.
- YAO, J.; LI, Y.; TAN, L. Option price forecasting using neural networks. *Omega*, v.28, n.4, p.455-466, Aug. 2000.
[DOI: [http://dx.doi.org/10.1016/S0305-0483\(99\)00066-3](http://dx.doi.org/10.1016/S0305-0483(99)00066-3)]
- ZHANG, Q.; PATUWO, B.; HU, M. Forecasting with artificial neural networks: the state of the art. *International Journal of Forecasting*, v.14, n.1, p.35-62, Mar. 1998.
[DOI: [http://dx.doi.org/10.1016/S0169-2070\(97\)00044-7](http://dx.doi.org/10.1016/S0169-2070(97)00044-7)]

ABSTRACT

Pricing R\$/USD exchange rate options: a comparison between the Black model and the artificial neural networks model

In this study, a multilayer neural network model was applied to the pricing of R\$/USD exchange rate call options traded on the São Paulo Securities, Commodities and Futures Exchange (BM&FBovespa) from January 2004 to December 2007. Based on the actual market prices, the performances of a neural network model and the Black model were compared, using the usual error metrics and statistical tests. Overall, the results showed that the artificial intelligence model outperformed the Black model for the different degrees of moneyness.

Keywords: artificial neural networks, options pricing, Black model.

RESUMEN

Valoración de opciones sobre tasa de cambio R\$/US\$ cotizadas en Brasil: una comparación entre los modelos Black y redes neuronales artificiales

En este estudio se aplicó un modelo de red neuronal multicapa para la valoración de opciones de compra sobre el tipo de cambio R\$/US\$, cotizadas en la *Bolsa de Valores, Mercadorias e Futuros* (BM&FBovespa), para el período comprendido entre enero de 2004 y diciembre de 2007. A partir de los precios efectivamente practicados en el mercado, se comparó el desempeño entre las redes neuronales y el modelo de Black, con el uso de métricas habituales de error y pruebas estadísticas. Los resultados mostraron, en general, la mejor adecuación del modelo de inteligencia artificial, en comparación con el modelo de Black, en diferentes grados de monetización.

Palabras clave: redes neuronales artificiales, valoración de opciones, modelo de Black.