

# Precificação de Opções de Taxa de Juros no Brasil: Uma Análise dos Modelos de Black, Vasicek, CIR e Redes Neurais Recorrentes

***Leandro Dos Santos Maciel***

Mestrando em Engenharia Elétrica

Departamento de Controle e Automação (DCA-FEEC-UNICAMP)

Email: maciel@dca.fee.unicamp.br

***Rosangela Ballini***

Professora Doutora do Instituto de Economia da Unicamp (IE-UNICAMP)

***Rodrigo Lanna Franco da Silveira***

Professor Doutor do Instituto de Economia da Unicamp (IE-UNICAMP)

## RESUMO

Nas décadas recentes, a precificação de derivativos de taxa de juros tem chamado grande atenção dos agentes de mercado em geral, assim como de pesquisadores. Gestores de portfólios, *traders*, analistas de risco, *policymakers*, podem extrair informações relevantes de contratos futuros, *swaps* e opções, uma vez que eles desempenham um papel importante em suas estratégias e processos de decisão. Dessa forma, este artigo compara a performance dos modelos de Black (1976), Vasicek (1977) e CIR (1985), tomados como *benchmarks*, com modelos de redes neurais artificiais (ANNs) recorrentes na precificação das opções de compra de IDI no Brasil, com base em dados diários durante o período de Janeiro de 2003 a Julho de 2007. O desempenho dos modelos foi avaliado com base em medidas tradicionais de acurácia de previsão. Além disso, foram realizados testes estatísticos paramétricos e não-paramétricos como os testes AGS, MGN e SIGN para comparação de modelos de predição competitivos. De acordo com as métricas de erros e os testes estatísticos, os modelos de RNAs superaram os modelos de Black, Vasicek e CIR na valoração das opções de compra de IDI, principalmente para as opções *out-of-the-money*.

**Palavras-Chave:** Redes Neurais Recorrentes; Opções de Taxa de Juros; Precificação de Opções; Opções de IDI.

## ABSTRACT

Pricing interest rate derivatives is a challenging task that has attracted the attention of many researchers recently. Portfolio and risk managers, policymakers, traders, find valuable information in forward, swap and option contracts. This information plays an important role in their strategies and decision making process. Thus, this paper compares the performance of Black (1976), Vasicek (1977) and CIR (1985) models, as benchmarks', with recurrent artificial neural networks (ANNs) models in pricing Brazilian IDI calls options using daily data for the period from January 2003 to July 2007. We measure forecast performance for all the estimated models based on summary measures of forecast accuracy. Nevertheless, we performed parametric and nonparametric statistical tests as AGS, MGN and SIGN for competing forecast models. According to the statistical tests and summary forecast measurements, ANNs models are superior to Black, Vasicek and CIR models in IDI calls option pricing, mainly in out-of-the-money options.

**Keywords:** Recurrent Neural Networks; Interest Rate Options; Option Pricing; IDI Options.

**Área ANPEC:** Área 7 - Microeconomia, Métodos Quantitativos e Finanças

## 1 Introdução

Nas décadas recentes, a precificação de derivativos de taxa de juros tem chamado grande atenção dos agentes de mercado em geral, assim como de pesquisadores, devido à sua importância econômica. Gestores de portfólios, *traders*, analistas de risco, *policymakers*, podem extrair informações relevantes de contratos futuros, *swaps* e opções, uma vez que elas desempenham um papel importante em suas estratégias e processos de decisão. Além disso, a curva de rendimentos, ou *yield curve*, é, sem dúvida, uma importante variável econômica, podendo ser modelada por meio das taxas implícitas desses contratos.

De acordo com o Banco de Compensações Internacionais (BIS - *Bank for International Settlements*), no mercado de derivativos mundial, os contratos relacionados à níveis da taxa de juros são os que apresentaram maior volume financeiro. Em 2009, no mercado secundário, estima-se que o valor nocional para contratos de taxa de juros atingiu US\$ 437 trilhões, além de US\$ 342 trilhões em *swaps* de taxa de juros. Ainda, segundo a Associação Internacional de *Swaps* e Derivativos (ISDA - *International Swaps and Derivatives Association*), 80% das 500 maiores companhias do mundo utilizaram derivativos de taxa de juros para controle de seus fluxos de caixa. Dessas mesmas companhias, 75% transacionaram opções de taxas cambiais, 25% opções sobre *commodities* e 10% aplicaram seus recursos em opções sobre índices de ações.

No Brasil, os contratos de derivativos são negociados na Bolsa de Valores, Mercadorias e Futuros (BM&FBOVESPA). Os principais contratos de taxa juros são: Futuro de Taxa Média de Depósitos Interfinanceiros de Um Dia (DI Futuro), Opção de Compra/Venda sobre Índice de Taxa Média de Depósitos Interfinanceiros de Um Dia (Opções de IDI) e Futuro de Cupom Cambial. O DI futuro é o derivativo mais líquido da bolsa, sendo, de acordo com (Burghardt e Acworth, 2010), um dos contratos mais negociados do mundo dentre os derivativos de juros. Segundo a BM&FBOVESPA, na década de 2000, o número médio anual de contratos negociados de DI Futuro foi em torno de 13 milhões, enquanto que as opções de IDI apresentaram o valor médio de 700 mil contratos negociados.

Uma das possíveis explicações para a baixa liquidez dos contratos de opções de IDI é a falta de um modelo de precificação amplamente aceito pelo mercado e sem deficiências teóricas (Barbachan e Ornelas, 2003). Na Bolsa, utiliza-se o modelo de (Black, 1976) para precificar tais opções. Teoricamente, esse modelo é aplicado para opções sobre futuros e não leva em consideração a estrutura a termo da taxa de juros. Dessa forma, os preços teóricos, obtidos pelo modelo, diferem significativamente dos preços observados no mercado, o que gera um receio entre os investidores ao se tomar posições nesses papéis.

As opções de taxa de juros, transacionadas nos mercados dos EUA e Europeu, em geral, têm como ativo objeto um título de renda fixa com vencimento posterior ao vencimento da opção. Assim, ela reflete a expectativa do comportamento futuro da estrutura a termo da taxa de juros entre o período de vencimento da opção e o vencimento do título. Como as opções de IDI refletem o comportamento da taxa de juros de curto prazo na data de avaliação até o vencimento da opção, torna-se necessário adequar os modelos de precificação para a realidade brasileira.

(Vieira Neto e Pereira, 1999) desenvolveram uma fórmula fechada para o cálculo do prêmio das opções de IDI com base no modelo de (Vasicek, 1977). Derivada a fórmula analítica, os autores não verificaram o modelo empiricamente, uma vez que esses papéis ainda não apresentavam liquidez significativa na Bolsa, pois foram introduzidos em 1997. Além disso, o modelo

se baseou num processo de Ornstein-Uhlenbeck, que permite taxas de juros negativas<sup>1</sup>. Posteriormente, (Gluckstern et al., 2002) aplicaram o modelo de (Hull e White, 1990), com base em árvores trinomiais, para precificar as opções de IDI, no período de 1997 a 2000. Seus resultados foram comparados com os obtidos por meio do modelo de (Black, 1976) e indicaram a superioridade do modelo proposto<sup>2</sup>.

(Barbachan e Ornelas, 2003) também derivaram uma fórmula analítica para avaliação das opções de IDI. Com base no modelo de (Cox et al., 1985) (CIR), os autores buscaram superar as limitações do trabalho de (Vieira Neto e Pereira, 1999), por se tratar agora de um modelo que não admite taxas de juro negativas. Os autores avaliaram seu modelo apenas para a série AB17 de abril de 2003 de opções de compra de IDI. Os resultados foram comparados com os obtidos pelo modelo de (Vieira Neto e Pereira, 1999) e mostraram que as fórmulas derivadas com base no modelo CIR resultaram em preços teóricos mais próximos aos observados no mercado em relação às fórmulas analíticas derivadas do modelo de Vasicek.

Por fim, (Almeida e Vicente, 2003) precificaram as opções de IDI com base em um modelo Gaussiano multi-fator dinâmico para a estrutura a termo da taxa de juros brasileira. Os resultados subestimaram os preços das opções. Além disso, conforme os autores, a metodologia envolve uma capacidade computacional muito pesada, o que dificulta sua aplicação prática. Ainda, (Barbedo et al., 2009) implementaram o modelo de (Heath et al., 1992) (HJM) para a valoração das opções de IDI. Os resultados evidenciaram que o modelo HJM, consistentemente, subavalia os papéis. Além disso, verificaram uma relação direta entre a maturidade desses contratos e o erro de precificação, em contraposição a relação negativa do erro com o grau de *moneyness*. Apesar desses modelos citados melhor se adequarem à realidade das opções de IDI, seus resultados ainda mostram que os prêmios teóricos não são capazes de explicar os prêmios do mercado.

Na tentativa de superar as deficiências dos modelos convencionais para precificação de opções, métodos não-paramétricos, baseados em Redes Neurais Artificiais (RNAs) têm se mostrado como uma relevante alternativa, uma vez que possuem elevada capacidade para modelar relações não-lineares, além de não depender de hipóteses restritivas. São inúmeros os trabalhos que utilizam esses modelos para avaliação de opções, dentre eles podemos citar: (Garcia e Gençay, 2000); (Ghaziri et al., 2001); (Gençay e Qi, 2001); (Andreou et al., 2008); (Healy et al., 2002); (Bennel e Sutcliffe, 2004); (Yao et al., 2000); (Liang et al., 2009); (Wang, 2009), (Samur e Temur, 2009). Esses trabalhos, em geral, aplicam modelos de rede neurais progressivas com base nos fatores que afetam diretamente os preços das opções avaliadas. Os principais mercados abordados são os dos Estados Unidos, Inglaterra, Índia, Taiwan, Austrália, Suécia e Japão. Além disso, as opções consideradas nas análises são as referenciadas em índices de ações, futuros, assim como opções sobre taxas cambiais de moedas estrangeiras. Aplicações de RNAs na valoração de opções de taxa de juros podem ser vistas em (Turnbull e Milne, 1991).

Em geral, a superioridade dos modelos de redes neurais é quase um consenso entre os pesquisadores, o que sugere a análise dessa técnica para os diferentes mercados de opções do mundo. Assim, este artigo sugere uma aplicação de modelos de redes neurais recorrentes para a precificação das opções de compra de IDI no período de Janeiro de 2003 a Julho de 2007. Os preços teóricos obtidos serão comparados com os resultantes dos modelos propostos por (Black, 1976), (Vieira Neto e Pereira, 1999) e (Barbachan e Ornelas, 2003), sendo os dois últimos baseados nos modelos de Vasicek e CIR, respectivamente, uma vez que se tratam

---

<sup>1</sup> (Almeida et al., 2003) calibram o modelo proposto por (Vieira Neto e Pereira, 1999), e seus resultados indicaram que a calibração se torna mais robusta em períodos de estabilidade econômico-financeira. Porém, em períodos com grande volatilidade ou pós-crise, o mesmo modelo gera parâmetros instáveis.

<sup>2</sup> (Junior et al., 2003) também testaram um modelo de precificação para as opções de IDI, baseado no modelo de (Hull e White, 1990), mas para o período de 1999 a 2002. Seus resultados indicaram a inconsistência do modelo por meio da avaliação da volatilidade implícito. Porém, mostrou-se aceitável em operações do tipo delta *hedge*.

das metodologias que levam em consideração as características do mercado de opções de IDI, além de apresentarem fórmulas analíticas fechadas para se obter o prêmio desses contratos. As comparações terão como base os preços de fechamento desses papéis, observados no mercado, adotando-se métricas usuais de erro e testes estatísticos.

Esse estudo se destaca ao aplicar tal método de inteligência artificial em um segmento de mercado ainda não avaliado no Brasil, o das opções de taxa de juros. Além disso, nesse artigo avalia-se modelos de redes neurais recorrentes, uma vez que a literatura, em geral, aplica modelos do tipo progressivas ou *feedforward*, que possuem menos potencial em modelar dados temporais ou espaciais. Dessa forma, a seção seguinte apresenta a metodologia, descrevendo os modelos de taxa juros e de redes neurais, assim como as métricas de avaliação de desempenho. Segue-se, então, os resultados e a discussão. Por fim, as conclusões.

## 2 Metodologia

### 2.1 Amostra

Os dados deste trabalho correspondem aos preços de fechamento das opções européias de compra de IDI, transacionadas na BM&FBOVESPA, para diferentes prazos e preços de exercício, que se iniciam em 2 de janeiro de 2003, e com fim em 31 de julho de 2007, totalizando 1.638 dias de negociação e 12.801 observações<sup>3</sup>. Temos associado, para cada contrato de opção o preço de exercício, o preço do contrato DI Futuro de mesmo prazo da opção, o número de contratos negociados, o volume financeiro e prazo para o vencimento em dias úteis<sup>4</sup>. A amostra se compõe dos contratos de IDI mais líquidos negociados na Bolsa, dado que foram considerados somente os papéis com volume de negociação superior a 800 contratos por dia. Ainda, para o mesmo período, utilizou-se os dados da média anualizada diária do CDI (Certificado de Depósitos Interfinanceiros) de um dia, uma vez que essa taxa é a referência dos derivativos de juros no Brasil<sup>5</sup>.

### 2.2 Modelos de Taxa de Juros

Os modelos de estrutura a termo da taxa de juro (ETTJ) podem ser de abordagem direta ou indireta. Modelos de abordagem direta partem da especificação direta do processo estocástico dos preços dos títulos primários, em vez de derivarem este processo de conceitos mais primitivos, como taxa de juro de curto prazo ou a relação de preferência dos agentes econômicos. Já os modelos de abordagem indireta partem da especificação do processo estocástico seguido pela taxa de juro de curto prazo ou pela taxa *forward*. Os modelos mais utilizados na literatura, e em que se baseiam os trabalhos de (Vieira Neto e Pereira, 1999) e (Barbachan e Ornelas, 2003), são os modelos indiretos. Eles dependem da incorporação dos principais fatores que procuram descrever a dinâmica evolutiva da ETTJ. Além disso, se baseiam num processo estocástico para a taxa de juros de curto prazo num mundo neutro ao risco.

A família dos modelos indiretos de um fator pode ser descrita genericamente pela seguinte equação<sup>6</sup>:

$$d(r_t) = [\theta - \alpha r_t] dt + \sigma_t r_t^\beta dW_t \quad (1)$$

<sup>3</sup> A avaliação das opções de venda (*puts*) não foi considerada, devido sua baixa liquidez na Bolsa.

<sup>4</sup> Os dados foram fornecidos pela BM&FBOVESPA.

<sup>5</sup> O CDI é divulgado diariamente pela ANBID (Associação Nacional dos Bancos de Investimento) e os dados foram coletados em: (<http://www.cetip.com.br/> - Acesso em: 13/10/2009).

<sup>6</sup> São ditos modelos de um fator aqueles que consideram apenas a taxa de juros de curto prazo como fator estocástico.

onde  $r_t$  descreve a taxa de juros de curto prazo;  $\theta$  é um parâmetro de ajuste para a curva observada no mercado ou a taxa de juros esperada de longo prazo;  $\alpha$  e  $\sigma_t$  indicam a velocidade de reversão à média e a volatilidade da taxa de juros de curto prazo, respectivamente;  $\beta$  corresponde ao coeficiente de elasticidade sobre  $r_t$ ; e  $W_t$  é um movimento browniano na medida de probabilidade natural.

Com base nesse modelo geral, os demais modelos que avaliam a dinâmica da taxa de juros se mostram como casos particulares. Definida a dinâmica para se modelar a ETTJ, desenvolveram-se fórmulas analíticas fechadas para precificar as opções relacionadas à taxa de juros, de acordo com as características dos contratos de opções padrões transacionados nos mercados norte-americano e europeu, que possuem como ativo objeto um título de renda fixa que vencerá num determinado número de meses depois do vencimento da opção.

As opções de IDI, ao invés de ter como ativo objeto um título de renda fixa com vencimento superior ao seu, negociam a taxa de juros acumulada (Índice DI - IDI) entre a data corrente e a sua própria data de vencimento. Portanto, a forma específica das opções sobre taxa de juros brasileira torna seu preço, bem como os fatores que o afetam, distinto daqueles ligados às opções sobre juros tradicionais<sup>7</sup>. Esses contratos correspondem a uma opção cujo *payoff* é uma função da taxa de juros de curto prazo acumulada entre a data de transação  $t$  e a data de maturação da opção  $T$ . Assim, temos:

$$IDI_t = IDI_t \prod_{i=t}^T (1 + CDI_i) \quad (2)$$

onde  $CDI_i = (1 + CDI_i\%a.a.)^{\frac{1}{252}} - 1$ , sendo que o  $CDI_i$  corresponde ao valor da taxa  $CDI$  no dia  $i$ . Em  $t = 0$ , o valor do  $IDI$  é 100.000, conforme definido pela BM&FBOVESPA em 02 de janeiro de 2003.

Dessa forma, o *payoff* de uma opção de compra de  $IDI$  na data de vencimento  $T$  é dado por:

$$P_t = \max \{0, IDI_T - X\} \quad (3)$$

em que  $X$  representa o preço de exercício da opção.

### 2.2.1 Modelo de Black (1976)

Utilizado pela BM&FBOVESPA, o modelo de (Black, 1976) destina-se a opções sobre futuros, e é bastante utilizado para a avaliação de opções sobre títulos pré-fixados. Trata-se de uma variação do modelo de (Black e Scholes, 1973), com o ativo objeto sendo um contrato futuro ao invés de um ativo à vista.

Assim como a maioria dos modelos, este considera a volatilidade do ativo-objeto como sendo constante ao longo da vida da opção. Entretanto, no caso dos títulos pré-fixados, quanto maior for o tempo até o vencimento, maior será a volatilidade. No caso da vida da opção ser muito curta em relação à vida do título-objeto é que poderíamos considerar uma volatilidade praticamente constante durante a vida da opção. Além disso, vale acrescentar que o modelo de (Black, 1976) assume taxas de juros nominais negativas.

Para as opções de IDI, a Bolsa se baseia no modelo de Black. Neste caso, o ativo-objeto é representado pelo valor do IDI corrigido pela taxa de juros esperada (DI Futuro) até a data do vencimento da opção, ou seja, cria-se um preço a termo do IDI. O problema da volatilidade decrescente é geralmente amenizado por meio da utilização de uma volatilidade “média”. A fórmula analítica para uma opção de compra sobre IDI na data  $t$  é dada por:

<sup>7</sup> Uma opção de IDI envolve as mesmas variáveis financeiras de um contrato de *swap* de taxa de juros, mas possui um padrão de comportamento distinto.

$$c_t = IDI_t N(d_1) - XP(t, T) N(d_2) \quad (4)$$

Em que:

$$d_1 = \frac{\ln\left(\frac{IDI_t}{XP(t, T)}\right) + \frac{\sigma^2(T-t)^3}{6}}{\sigma\sqrt{(T-t)^3/3}} \quad (5)$$

$$d_2 = \frac{\ln\left(\frac{IDI_t}{XP(t, T)}\right) - \frac{\sigma^2(T-t)^3}{6}}{\sigma\sqrt{(T-t)^3/3}} \quad (6)$$

onde  $N(\cdot)$  denota a função de distribuição normal padrão acumulada,  $P(t, T)$  é o preço em  $t$  de um título pré-fixado que paga uma unidade monetária em  $T$ <sup>8</sup>;  $IDI_t$  é o valor do  $IDI$  em  $t$ ; e  $\sigma$  é a volatilidade de um título pré-fixado de curto prazo.

## 2.2.2 Modelo de Vasicek (1977)

(Vasicek, 1977) foi o primeiro a propor o comportamento de reversão à média, fato estilizado das taxas de juros de curto prazo. Isso significa que se a taxa de curto prazo está acima de uma média dita de longo prazo, a tendência dela é de queda. Se estiver abaixo da média esperada de longo prazo, a tendência é de alta. (Vieira Neto e Pereira, 1999) propuseram uma fórmula analítica fechada para precificação das opções de  $IDI$  com base no método de (Vasicek, 1977). Segundo os autores, o prêmio de uma opção de compra de  $IDI$  ( $c_t$ ) é dado por:

$$c_t = IDI_t N(h) - XP(t, T) N(h - k) \quad (7)$$

Sendo:

$$h = \frac{\ln\left(\frac{IDI_t}{P(t, T)X}\right) + \frac{k^2}{2}}{k} \quad (8)$$

$$k^2 = \sigma_t^2 \frac{(4e^{-\alpha\tau} - e^{2\alpha\tau} + 2\alpha\tau - 3)}{2\alpha^3} \quad (9)$$

em que  $IDI_t$  indica o valor do  $IDI$  em  $t$ ;  $X$  é o preço de exercício da opção;  $P(t, T)$  é o preço, em  $t$ , de um título de renda fixa que paga uma unidade monetária em  $T$ ;  $\alpha$  é o parâmetro de reversão à média da taxa de juros de curto prazo;  $\tau = T - t$ , ou seja, o prazo de maturidade opção;  $\sigma_t^2$  indica a volatilidade da taxa de juros no curto prazo; e  $X$  o preço de exercício.

Nesse modelo, temos dois parâmetros não observáveis: a taxa de reversão à média ( $\alpha$ ) e a volatilidade ( $\sigma_t^2$ ) da taxa de juros de curto prazo. Neste trabalho, seguiremos a metodologia de estimação exógena para essas variáveis, por meio da estimação condicional da volatilidade, proposta por (Vieira Neto e Pereira, 1999). De certa forma, os parâmetros  $\alpha$  e  $\sigma$  encontram-se implícitos em  $k$  que, então, passa a ser a única variável não observável da fórmula. Segundo (Vieira Neto e Pereira, 1999), o valor de  $k$  pode ser estimado, uma vez que é a volatilidade de um dado financeiro observável. Essa estimação se baseia na inclusão do efeito duração (*duration*) num modelo GARCH(p,q). Dessa forma, a estimação será dada de acordo com um processo GARCH(p,q)-D.

<sup>8</sup> Naturalmente,  $P(T, T) = 1$ .

Sejam  $R(t, T_n)$  a taxa de juros implícita no preço de ajuste do contrato futuro de DI de  $n$ -ésimo vencimento, isto é,  $R(t, T_n) = \frac{100.000}{PU(t, T_n)} - 1$  e  $r(t)$  o CDI observado na data  $t$ . O GARCH(p,q)-D é composto pelas três relações abaixo:

$$R(t+1, T_n) = R(t, T_n)^* + \sigma(t+1, T_n) \varepsilon_n(t+1) \quad (10)$$

$$\sigma(t+1, T_n)^2 = d(\beta_1, \beta_2, t+1, T_n) + \beta_3 + \sum_{i=1}^p \varphi_i \sigma(t+1-i, T_n)^2 \quad (11)$$

$$+ \sum_{i=1}^q \delta_i [R(t+1-i, T_n) - R(t-i, T_n)]^2$$

$$\varepsilon(t+1) \sim NI(0, 1) \quad (12)$$

onde  $R(t, T_n)^* = \frac{1+R(t, T_n)}{1+r(t)} - 1$  se  $t+1$  não for o primeiro dia útil do mês e  $R(t, T_n)^* = \frac{1+R(t, T_{n+1})}{1+r(t)} - 1$  caso contrário.

A função  $d(\cdot)$  é determinista e tem como objetivo capturar a tendência cíclica de decréscimo da volatilidade conforme diminui o prazo para o vencimento do contrato. Uma possível forma para  $d(\cdot)$  é a própria fórmula de  $k^2$ . Outras formas possíveis incluem:  $d(t, T_n) = \beta_1 (T_n - t)^{\beta_2}$ , onde a curvatura da função é determinada por  $\beta_2$ ;  $d(t, T_n) = \beta_1 (T_n - t)$ , forma linear e mais parcimoniosa em termos de quantidade de parâmetros, que será utilizada neste trabalho.

Dessa forma, para obtenção do prêmio das opções de IDI seguindo o modelo de (Vasicek, 1977), os parâmetros não observáveis serão estimados implicitamente pela estimação da volatilidade com base num processo GARCH(p,q)-D. O processo GARCH(p,q)-D foi parametrizado segundo os critérios de informação de Akaike (AIC) e Schwarz (BIC).

### 2.2.3 Modelo de Cox, Ingersoll e Ross (1985)

O modelo descrito por (Vasicek, 1977) apresenta uma limitação ao permitir que a taxa de juros assumia valores negativos. Entretanto, no modelo CIR, a volatilidade da taxa de juros de curto prazo é  $\sigma_t \sqrt{r(t)}$ , assim, sempre que  $r(t)$  se aproximar de zero, sua raiz quadrada será um número muito pequeno, tornando a volatilidade desprezível, então, a magnitude do termo “ $\sigma_t \sqrt{r(t)}$ ” será dominante, fazendo com que  $r(t)$  se afaste da origem e seja sempre positivo. Com base nesse processo de difusão para a taxa de juros de curto prazo, (Barbachan e Ornelas, 2003) construíram uma forma analítica fechada para o cálculo do prêmio de uma opção de compra de IDI ( $c_t$ ). Nesse caso, temos que:

$$c_t = \psi IDI_t \left[ 1 - \chi^2 \left( \ln \left( \frac{X}{IDI_{t-1}} \right); \eta, \rho \right) \right]$$

$$- X \psi^3 \frac{-\eta}{2} e^{-\mu} \left[ 1 - \chi^2 \left( \ln \left( \frac{X}{IDI_{t-1}} \right) / 3; \eta, \mu \right) \right] / (1 + CDI_t) \quad (13)$$

Em que:

$$\rho = \frac{4CDI_t}{\sigma_t^2} \left( \ln \left( \frac{1 - e^{\theta^* \tau}}{1 - e^{-\theta^*}} \right) \right) \quad (14)$$

$$\eta = \frac{4\theta^* \alpha^* \Delta t}{\sigma_2^t} \quad (15)$$

$$\Psi = \frac{\sigma_2^t}{4\theta^*} + \frac{\sigma_2^t}{4\Delta t e^{\theta^*}} (e^\tau - 1) \quad (16)$$

sendo  $\theta^* = (\theta - \lambda)$ ;  $\alpha^* = \left(\frac{\alpha\theta}{\theta - \lambda}\right)$ ; e  $\mu = \frac{\rho\Psi}{2 + \Psi}$ , em que  $\alpha$  é o parâmetro da velocidade de reversão à média da taxa de juros;  $\theta$  a taxa de juros de longo prazo;  $\sigma^2$  a volatilidade da taxa de juros de curto prazo;  $\tau = T - t$ ;  $\lambda$  o preço de mercado do risco;  $\Delta t = T - t - \xi^9$ ; e  $\chi^2$  denota uma distribuição chi-quadrado não central.

Neste modelo, temos quatro variáveis não observáveis: a velocidade de reversão à média ( $\alpha$ ) e a volatilidade ( $\sigma^2$ ) da taxa de juros de curto prazo, assim como o preço de mercado do risco ( $\lambda$ ) e a taxa de juros de longo prazo ( $\theta$ ). Na estimação desses parâmetros, adotaremos a mesma metodologia de (Barbachan e Ornelas, 2003), exceto para a volatilidade. A velocidade de reversão à média será arbitrada em 0,006. Isso se deve ao fato de que a literatura, em geral, tem estimado, com base em diversos modelos, esse valor para o parâmetro (Barbachan e Ornelas, 2003). Além disso, os autores sugerem o preço de mercado do risco igual a  $-1$ . Ele é negativo, uma vez que o risco é suposto indesejável, e o valor unitário indica que, para cada unidade adicional de risco, o investidor exige uma unidade a mais de retorno.

Os demais parâmetros serão estimados com base em dados históricos. A taxa de juros de longo prazo será estimada por meio da média móvel do CDI over, com uma janela amostral de período de 3 anos. Já o parâmetro da volatilidade será obtido por meio do processo GARCH(p,q)-D, uma vez que permite o cálculo da variância levando em consideração o efeito da *duration*, que melhor se adéqua à *yield curve*.

## 2.3 Redes Neurais Artificiais

Devido sua capacidade em lidar com problemas de reconhecimento de padrões, classificação e predição, tanto para dados transversos quanto para séries temporais ou espaciais, as Redes Neurais Artificiais (RNAs) vêm desempenhando um papel importante nas aplicações em economia e finanças (Shapiro, 2003). Essas técnicas são capazes de modelar tanto estruturas lineares quanto não-lineares, além de serem aproximadores universais de funções e executar tarefas como associação ou memorização, codificação e simulação.

Os modelos de redes neurais são formados por unidades básicas de processamento, denominadas neurônios artificiais. Podemos denotar a saída de um neurônio por:

$$y_i = \Phi \left( \sum_{j=1}^m w_{ij} x_j \right) \quad (17)$$

onde  $x_1, x_2, \dots, x_m$  são os sinais de entrada;  $w_{k1}, w_{k2}, \dots, w_{km}$  são os pesos sinápticos do neurônio  $i$ ;  $u_i$  é a saída do combinador linear devido aos sinais de entrada;  $\Phi(\cdot)$  é a função de ativação; e  $y_i$  é o sinal de saída do neurônio. Em geral, a função de ativação assume a seguinte forma:

$$\Phi(u) = \frac{1}{1 + e^{-au}} \quad (18)$$

onde  $a$  é o parâmetro de inclinação da função. Esta função é de tipo sigmóide, denominada de função logística.

Os neurônios artificiais são dispostos em diversas camadas, o que caracteriza o modelo conhecido como perceptron de múltiplas camadas (*multilayer perceptron* - MLP). Dessa forma,

<sup>9</sup> Segundo (Barbachan e Ornelas, 2003) adota-se  $\xi = \frac{1}{252}$ , que corresponde a um dia útil dividido pelo total de dias úteis durante o período de um ano regular.



uma rede neural é um conjunto de unidades computacionais interconectadas e organizadas em camadas (Haykin, 2009). Os neurônios que recebem os sinais de entrada constituem a camada de entrada. Os neurônios que recebem as saídas dos neurônios da camada de entrada formam a segunda camada, ou camada intermediária/escondida, assim sucessivamente, até a última camada, denominada camada de saída (Haykin, 1998).

Quando as ligações entre os neurônios são unidirecionais (na direção das entradas para as saídas), temos uma rede neural progressiva ou *feedforward*. Uma rede neural recorrente (RNR) tem por base uma rede progressiva com algumas modificações, notadamente, a introdução de realimentação, o que amplia sua potencialidade de modelagem de dados temporais ou espaciais. As realimentações consistem em saídas de neurônios de determinada camada serem reintroduzidas como entradas de neurônios de camadas anteriores ou da própria. Essas possibilidades fazem com que a arquitetura de uma rede recorrente possa tomar diversas formas. Tais realimentações são acompanhadas de defasagens de uma ou mais unidades do tempo, no caso de tempo discreto.

A Figura 1 apresenta a topologia de uma rede neural recorrente do tipo Elman e do tipo Jordan (Mandic e Chambers, 2001), com uma camada intermediária contendo  $Q$  neurônios e uma camada de saída com  $O$  neurônios, sendo  $h^{-1}$  o operador de defasagem. A diferença entre esses dois tipos de redes se dá na realimentação das camadas. Na rede neural recorrente de Elman (RNRE), a recorrência se dá das camadas intermediárias para suas respectivas camadas anteriores, sejam elas de entrada assim como intermediárias, enquanto que na rede de Jordan (RNRJ), a realimentação se dá da camada de saída para a camada de entrada. Neste trabalho, utilizaremos os dois tipos de redes neurais recorrentes da Figura 1 para a precificação das opções de compra de IDI, uma vez que é uma das mais utilizadas para predição de séries temporais (Mandic e Chambers, 2001). Em relação às redes de tipo progressivas, as redes recorrentes permitem que o modelo seja capaz de capturar a dependência temporal entre as variáveis de entrada e saída, o que garante maior aplicabilidade para problemas baseados em predição de séries temporais.

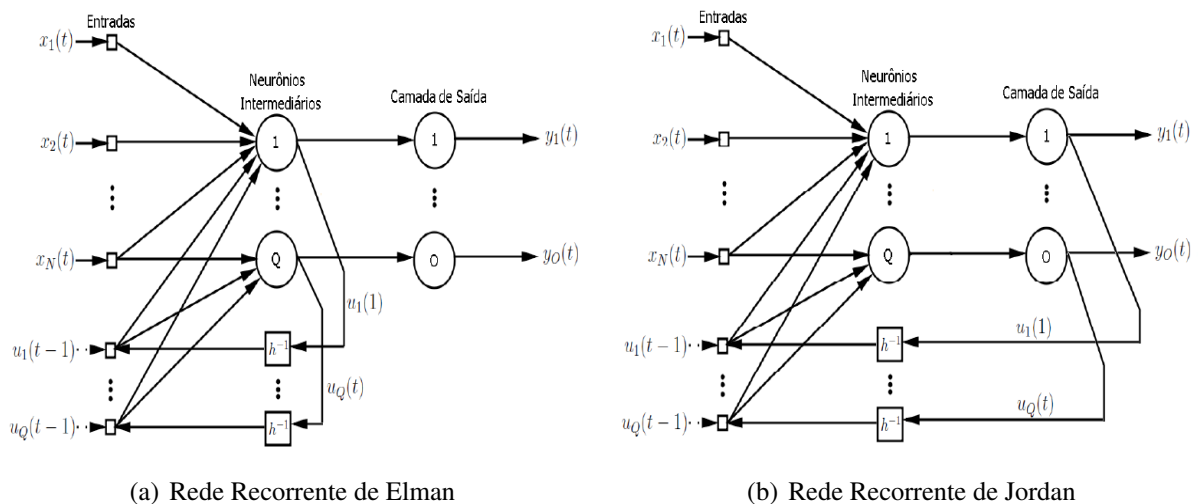


Fig. 1: Modelos de Redes Neurais Recorrentes

Para o treinamento da rede neural, utilizaremos o algoritmo baseado no método de Levenberg-Marquardt. Ele consiste em um aperfeiçoamento do método de Gauss-Newton, que é uma variante do método de Newton. O método de Newton usa a informação da derivada parcial de segunda ordem do índice de performance utilizado para corrigir os pesos sinápticos  $\bar{w}_0$ .

O método de Newton pode ser utilizado para resolver a seguinte equação em uma série de Taylor em torno dos pesos  $\bar{w}_0$ , com base no gradiente  $\nabla$ :

$$\nabla V(\bar{w}) = \nabla V(\bar{w}_0) + (\bar{w} - \bar{w}_0) \nabla^2 V(\bar{w}_0) + \dots \quad (19)$$

em que o índice de performance  $V$  pode ser definido como a soma do quadrado do erro  $e_i$  associado ao  $i$ -ésimo padrão de treinamento da rede, isto é:

$$V(\bar{w}) = \sum_{i=1}^N e_i^2(\bar{w}) \quad (20)$$

Para se obter a regra de atualização de  $\bar{w}$  na equação 19, foram propostos métodos que utilizam aproximações, chamadas Quase-Newton, dentre as quais se situam o método de Gauss-Newton e Levenberg-Marquardt. Dessa forma, a regra de atualização para os pesos sinápticos, segundo o método de Levenberg-Marquardt é dado por:

$$\Delta \bar{w} = - [J^T(\bar{w})J(\bar{w}) + \mu \text{diag}(J^T(\bar{w})J(\bar{w}))]^{-1} J^T(\bar{w}) \bar{e}(\bar{w}) \quad (21)$$

sendo que  $J$  representa a matriz Jacobiana (derivadas parciais) e  $\mu$  um escalar. Dessa forma, o problema de otimização se resume na estimação dos pesos sinápticos com base na função objetivo do erro, considerando a maior aproximação aos prêmios das opções observados no mercado.

### 2.3.1 Redes Recorrentes para o Prêmio das Opções de Compra de IDI

Os modelos de redes neurais recorrente, para obtenção do prêmio das opções de compra de IDI, se caracterizam pelas seguintes variáveis utilizadas como entradas: valor do IDI, atualizado de acordo com o CDI; preço de exercício da opção; maturidade ou tempo de vencimento da opção; e a volatilidade da taxa de juros de curto prazo, obtida por um modelo GARCH(p,q)-D, proposto por (Vieira Neto e Pereira, 1999), uma vez que se mostra como uma alternativa mais realista para a obtenção da variância de taxas de juros ao levar em conta o efeito duração.

A construção da estrutura da rede neural mais adequada aos dados, isto é, escolha do número de camadas e neurônios intermediários se deu com base no critério de informação BIC, em termos do erro quadrático médio (EQM), ou seja:

$$BIC = -2 \left( \frac{\ln(EQM)}{N} \right) + 2 \left( \frac{\ln \Psi}{N} \right) \quad (22)$$

onde  $EQM = \frac{1}{N} \sum_{i=1}^N (c_i - \hat{c}_i)^2$ , sendo  $c_i$  e  $\hat{c}_i$  os preços de mercado e estimado pelo modelo das opções de compra de IDI, respectivamente, e  $\Psi$  indica o parâmetro a ser avaliado, i.e., número de camadas intermediárias e número de neurônios em cada uma delas<sup>10</sup>.

Além disso, para estimação do modelo foi utilizada a técnica estatística de validação cruzada (Stone, 1974). Neste caso, os dados são particionados em conjuntos de treinamento, validação e teste, cujos objetivos são, respectivamente, a estimação dos parâmetros da rede, avaliação da evolução do desempenho da rede durante o treinamento, e, por fim, testar a capacidade de aprendizagem e generalização da rede.

## 2.4 Avaliação dos Modelos de Precificação

Para avaliação dos modelos empregados neste trabalho, comparamos os prêmios das opções de compra de IDI teóricos (i.e., obtidos pelos modelos) em relação aos prêmios observados

<sup>10</sup> (Kaastra e Boyd, 1995) indicam que não existe uma técnica capaz de informar a estrutura ideal para um modelo de rede neural, sendo que, na realidade, envolve escolhas empíricas que variam de acordo com a especificidade dos dados e o objetivo de estimação e previsão.

no mercado. Para tanto, foram utilizadas as métricas tradicionais de erro de predição: Erro Percentual Médio (MEP - *Mean Percentage Error*), Erro Percentual Máximo (MPE - *Maximum Percentage Error*), Raiz do Erro Quadrático Médio (RMSE - *Root Mean Squared Error*), Erro Absoluto Médio (MAE - *Maximum Absolute Error*), Erro Médio Absoluto Percentual (MAPE - *Maximum Absolute Percentage Error*), e o coeficiente de desigualdade de Theil (TIC), definidos como:

$$MEP = \frac{100}{N} \sum_{i=1}^N \frac{|c_i - \hat{c}_i|}{c_i} \quad (23)$$

$$MPE = \max_{i=1, \dots, N} \left\{ 100 \frac{|c_i - \hat{c}_i|}{c_i} \right\} \quad (24)$$

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (c_i - \hat{c}_i)^2} \quad (25)$$

$$MAE = \frac{1}{N} \sum_{i=1}^N |c_i - \hat{c}_i| \quad (26)$$

$$MAPE = \frac{1}{N} \sum_{i=1}^N \left| \frac{c_i - \hat{c}_i}{c_i} \right| \quad (27)$$

$$TIC = \frac{\sqrt{\frac{1}{N} \sum_{t=1}^N (\hat{c}_t - c_t)^2}}{\sqrt{\frac{1}{N} \sum_{t=1}^N \hat{c}_t^2} + \sqrt{\frac{1}{N} \sum_{t=1}^N c_t^2}} \quad (28)$$

Para as métricas usuais de error, seu menor valor indica o melhor modelo, enquanto que o coeficiente de desigualdade de Theil varia entre zero e um, em que zero indica o ajuste perfeito.

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

$$M = \frac{IDI_t}{X \cdot P(t, T)} \quad (29)$$

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* ( $1 - \alpha\% < M < 1 + \alpha\%$ ) 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:

$$c_i = \phi_1 + \phi_2 \cdot \hat{c}_i + \zeta_t \quad (30)$$

no qual  $\phi_1$  e  $\phi_2$  são os parâmetros linear e angular da regressão, respectivamente, e  $\zeta_t$  um ruído branco.

Os resultados empíricos na literatura se baseiam extensivamente em medidas padrões de acurácia, como as métricas de erro apresentadas acima, e não levam em consideração a superioridade de um modelo a outro em termos de significância estatística. Assim, neste trabalho foram aplicados testes paramétricos e não-paramétricos para verificar se existe diferença estatisticamente significativa entre os resultados dos modelos avaliados. Esses testes são apresentados a seguir.

O teste paramétrico aqui descrito se baseia no trabalho de (Ashley et al., 1980), denominado teste AGS. O teste AGS 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 convencionais de precificação (Black, Vasicek ou CIR), tomados como *benchmark*, e de Redes Neurais,  $e_i^B$  e  $e_i^{RNA}$ , respectivamente. Assim, temos que:

$$\Delta e_i^{B,RNA} = e_i^B - e_i^{RNA} = (c_i^B - \hat{c}_i^B) - (c_i^{RNA} - \hat{c}_i^{RNA}) \quad (31)$$

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

$$\Delta e_i^{B,RNA} = \beta_1 + \beta_2 (S_i^{B,RNA} - \mu_S) + v_i \quad (32)$$

Na qual  $v_i$  indica um processo ruído branco.

O teste AGS mostra que  $\beta_1$  representa a diferença entre os erros quadráticos médios de previsão obtidos pelos modelos e  $\beta_2$  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 que a hipótese alternativa consiste em  $H_1 : \beta_1 > 0$  e/ou  $\beta_2 > 0$ . A estatística para este teste, obtida por meio da estimação da equação 32, tem distribuição  $F$  com 2 e  $(N - 2)$  graus de liberdade, assumindo normalidade para os erros. A rejeição da hipótese nula indica que os modelos baseados em Redes Neurais superam os métodos tomados como *benchmark*.

Neste caso, enfatizaremos as comparações entre as metodologias tradicionais com as propostas neste trabalho, baseadas nos modelos de redes neurais. Além disso, iremos comparar a acurácia entre os próprios *benchmarks* e de redes neurais para auferição do melhor método para as fórmulas analíticas e para as metodologias de inteligência artificial.

Outro teste similar, denominado de teste de Morgan-Granger-Newbold (teste MGN), baseado inicialmente no trabalho de (Granger e Newbold, 1977). A utilização desse teste é recomendada para avaliar modelos de predição e é amplamente aplicado quando se relaxa a assunção de não existência de correlação entre os erros dos modelos. O valor da estatística para esse teste pode ser calculado como:

$$MGN = \frac{\hat{\rho}_{s\Delta}}{(1 - \hat{\rho}_{s\Delta})^{1/2}} (N - 1)^{1/2} \quad (33)$$

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 este 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).

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

$$SIGN = \left( \psi(N) - \frac{N}{2} \right) \cdot 2(N)^{1/2} \quad (34)$$

Sendo que o termo  $\psi_t$  denota o número de vezes em que, neste caso, os erros dos modelos tradicionais superaram os erros derivados dos modelos de Redes Neurais. 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 dos modelos do *benchmark* são da mesma proporção que os de inteligência artificial 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 analíticos superam os dos modelos de Redes Neurais numa proporção maior que  $(N/2)$ , o teste indica que os modelos tradicionais são menos acurados que os modelos de Redes Neurais. A estatística para o teste SIGN tem distribuição gaussiana com média zero e variância unitária. Assim, se a estatística SIGN é significativamente elevada, a hipótese nula de equivalência de predição pode ser rejeitada em favor da hipótese alternativa.

### 3 Resultados e Discussão

Os prêmios das opções de compra de IDI foram obtidos de acordo com três modelos tradicionais: modelo de Black, modelo de Vasicek, proposto por (Vieira Neto e Pereira, 1999); e modelo CIR, proposto por (Barbachan e Ornelas, 2003). Além disso, os preços teóricos foram obtidos de acordo com três modelos de redes neurais, sendo um deles uma rede progressiva e as outras duas redes de tipo recorrentes.

De acordo com o critério de informação de Schwarz (BIC), foram selecionadas as topologias de rede que apresentaram os melhores resultados. A Figura 2 apresenta as arquiteturas para os diferentes tipos de rede. Tanto a rede de Elman quanto a rede de Jordan possuem três camadas intermediárias, em que apenas se diferenciam quanto ao número de neurônios intermediários, o que torna a rede de Jordan mais complexa, em termos de parâmetros.

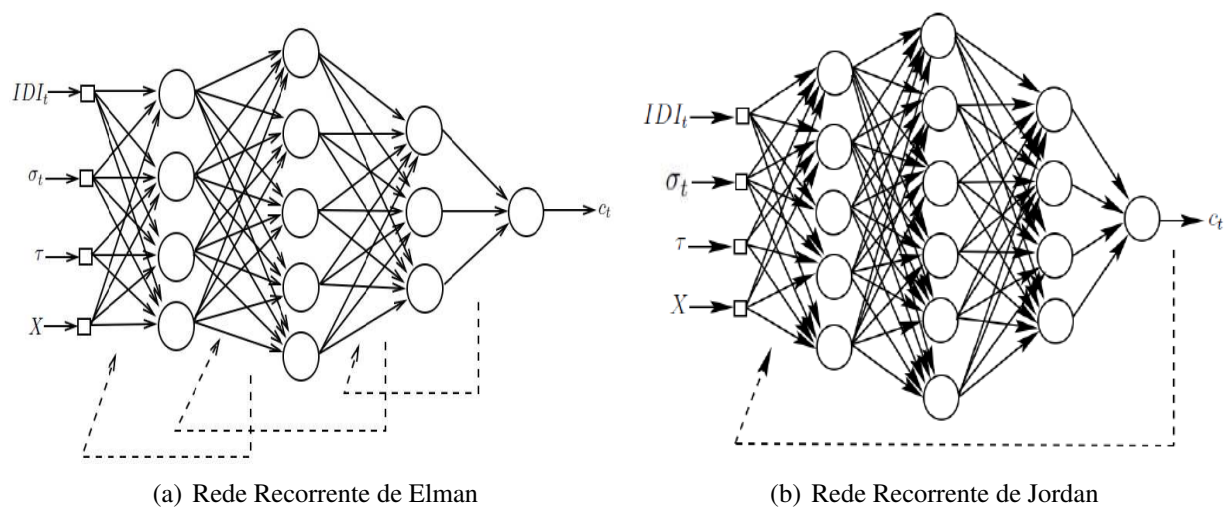


Fig. 2: *Arquiteturas das Redes Recorrentes para Avaliação das Opções de Compra de IDI*

Para se avaliar o desempenho da rede, o método de validação cruzada se caracterizou pelos conjuntos de treinamento, validação e teste que corresponderam, respectivamente, aos seguintes períodos: de 02/01/2003 a 02/02/2006 (70% da amostra), de 03/02/2006 a 08/12/2006 (20% da amostra), e de 11/12/2006 a 31/07/2007 (10% do total da amostra, com 1.270 observações).

Na Tabela 1, apresentamos os resultados, para as métricas de erro, obtidos de acordo com os modelos avaliados. Nesse caso, consideramos apenas o período da amostra teste das redes neurais, uma vez que buscamos avaliar sua capacidade de aprendizado e generalização e, além disso, para os modelos analíticos, essa distinção não altera a essência dos resultados.

Os resultados indicados na Tabela 1 evidenciam a superioridade dos modelos de redes neurais recorrentes propostos para a precificação das opções de compra de IDI, de acordo com os baixos valores obtidos das métricas de erro, assim como para o coeficiente de desigualdade de Theil e de determinação da regressão dos preços teóricos contra os prêmios de mercado. Deve-

Tab. 1: Aderência dos preços calculados pelos modelos tradicionais e de Redes Neurais aos preços de fechamento das *calls* de IDI européias negociadas na BM&FBOVESPA entre 11/12/2006 e 31/07/2007. Amostra com 1.270 observações

Modelos	Métricas						
	MEP	MPE	RMSE	MAE	MAPE	TIC	$R^2$
Black	87,98%	96,15%	44,823	123,434	34,589	0,737	0,553
Vasicek	80,12%	94,54%	33,265	99,872	33,291	0,698	0,622
CIR	74,65%	90,20%	24,983	93,652	29,382	0,507	0,663
RNRE	46,98%	66,82%	8,283	45,214	14,028	0,312	0,910
RNRJ	51,98%	70,99%	7,873	41,862	15,386	0,375	0,892

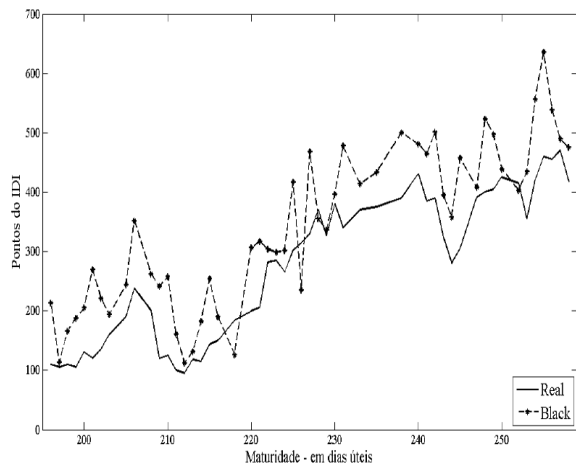
se perceber, sobretudo, que o modelo de precificação baseado em (Black, 1976) apresentou os piores resultados. Os prêmios das opções obtidos pelos modelos convencionais indicam a inconsistência de tais metodologias para o mercado de opções de IDI.

Com isso, as opções de compra de IDI foram separadas de acordo com o grau de *moneyness* e, então, obtivemos os valores das métricas de erro, cujos resultados são indicados na Tabela 2.

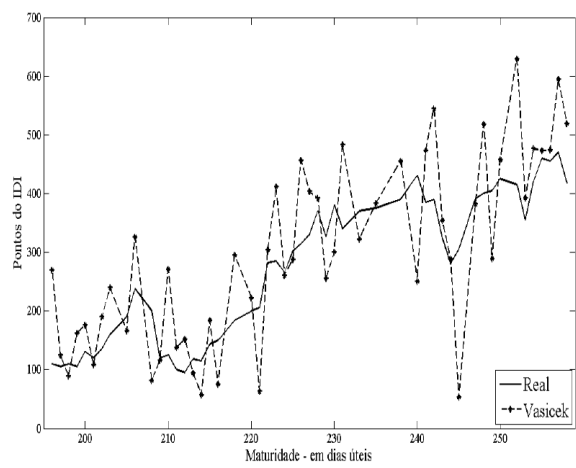
Tab. 2: Aderência dos preços calculados pelos modelos tradicionais e de Redes Neurais aos preços de fechamento das *calls* de IDI européias negociadas na BM&FBOVESPA entre 11/12/2006 e 31/07/2007 de acordo com o grau de *money*. Amostra com 1.270 observações

Modelos	Out-of-the-Money						
	MEP	MPE	RMSE	MAE	MAPE	TIC	$R^2$
Black	92,32%	102,23%	49,287	138,233	41,872	0,797	0,493
Vasicek	86,21%	104,32%	39,422	105,643	39,092	0,788	0,503
CIR	79,83%	94,82%	31,625	105,123	31,726	0,639	0,537
RNRE	48,83%	67,21	9,993	47,321	16,862	0,392	0,876
RNRJ	54,32%	75,42%	8,982	49,421	17,421	0,399	0,802
Modelos	At-the-Money						
	MEP	MPE	RMSE	MAE	MAPE	TIC	$R^2$
Black	86,32%	95,21%	45,421	122,532	33,312	0,772	0,582
Vasicek	81,11%	95,82%	31,312	100,321	34,321	0,701	0,611
CIR	75,23%	91,33%	22,872	94,421	31,312	0,531	0,692
RNRE	48,11%	68,93%	7,983	48,937	12,399	0,374	0,876
RNRJ	54,92%	68,99%	9,32	45,422	13,532	0,366	0,800
Modelos	In-the-Money						
	MEP	MPE	RMSE	MAE	MAPE	TIC	$R^2$
Black	86,99%	95,21%	42,321	121,312	33,873	0,732	0,521
Vasicek	82,32%	96,98%	32,111	102,321	31,321	0,672	0,601
CIR	75,32%	90,89%	26,321	91,321	30,532	0,532	0,688
RNRE	43,83%	64,42%	9,222	43,824	16,384	0,298	0,921
RNRJ	53,32%	68,54%	6,999	42,321	13,313	0,388	0,876

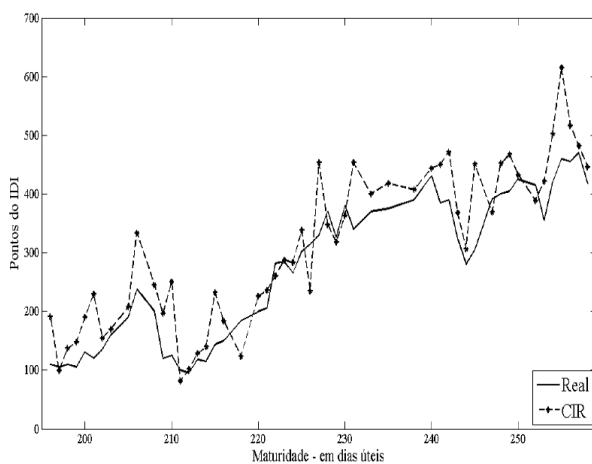
Segundo a Tabela 2, podemos perceber que os modelos teóricos apresentaram baixos erros em opções *at-the-money* e *in-the-money*, se comparados com os obtidos nas opções *out-of-the-money*. Porém, mesmo nesses casos, os modelos de redes recorrentes apresentou os menores valores para as métricas de erro. Na Figura 3 é apresentado a evolução dos prêmios das opções de compra de IDI para um dos strikes mais líquidos do período de dezembro de 2006 a julho de 2007 (amostra teste), isto é, com preços de exercício de 220.000 pontos do IDI.



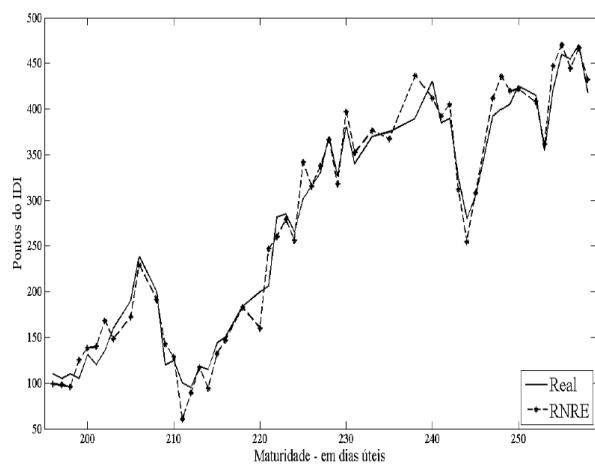
(a) Modelo de Black



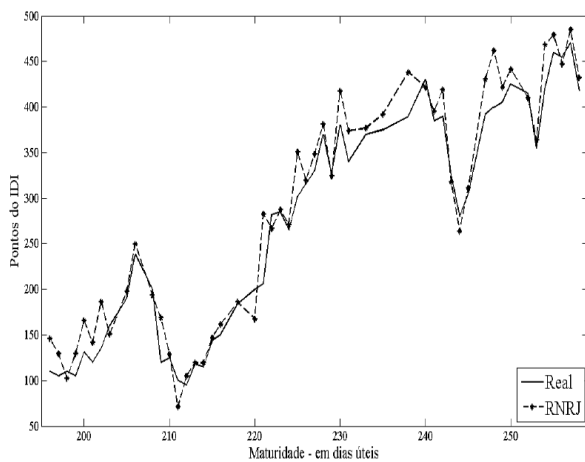
(b) Modelo de Vasicek



(c) Modelo CIR



(d) Rede Recorrente de Elman



(e) Rede Recorrente de Jordan

Fig. 3: Evolução dos prêmios de mercado em relação aos prêmios obtidos pelos modelos teóricos para as opções com strike de 200.000 em diversas maturidades.

É evidente a superioridade dos modelos de redes recorrentes, uma vez que conseguem acompanhar a evolução do prêmio dessas opções de forma mais acurada que os modelos baseados nos processos de Vasicek e CIR, para o comportamento da taxa de juros de curto prazo. Além disso, mesmo com o aumento do prazo de maturidade dessas opções, os resultados dos modelos de redes neurais não se afastam dos prêmios observados no mercado. Pode-se notar, ainda, que

o modelo de Black superestima os prêmios das opções de IDI, confirmando sua inadequação para esse tipo de opção.

Para os testes estatísticos, tomamos como *benchmarks* os modelos de fórmula fechada para então compará-los com os modelos propostos de redes recorrentes. Assim, a hipótese nula indica que os modelos comparados são equivalentes, contra a hipótese alternativa, baseada na assertiva de superioridade dos modelos de redes neurais. A Tabela 3 apresenta os resultados dos testes estatísticos.

Tab. 3: Testes AGS, MGN e SIGN para avaliação dos modelos de precificação para as opções de compra de IDI européias negociadas na BM&FBOVESPA entre 11/12/2006 e 31/07/2007, amostra com 1.270 observações.

Modelos	Teste AGS ( $F_{2,1268}$ )*	Teste MGN ( $t_{1269}$ )*	Teste SIGN $N(0,1)$ *
Black Vs. Vasicek	64,872 (0,0000)	16,324 (0,0000)	-4,623 (0,0001)
Black Vs. CIR	67,724 (0,0000)	17,922 (0,0000)	-4,762 (0,0001)
CIR Vs. Vasicek	14,762 <b>(0,2687)</b>	3,255 <b>(0,4402)</b>	-1,982 <b>(0,2791)</b>
Black Vs. RNRE	107,987 (0,0000)	25,782 (0,0001)	-7,287 (0,0000)
Black Vs. RNRJ	122,238 (0,0000)	31,972 (0,0000)	-8,982 (0,0000)
Vasicek Vs. RNRE	98,872 (0,0000)	34,762 (0,0000)	-8,761 (0,0000)
Vasicek Vs. RNRJ	105,762 (0,0000)	33,999 (0,0000)	-9,872 (0,0000)
CIR Vs. RNRE	102,762 (0,0000)	42,819 (0,0000)	-9,623 (0,0000)
CIR Vs. RNRJ	96,534 (0,0000)	36,829 (0,0000)	-8,777 (0,0000)
RNRE Vs RNRJ	12,653 <b>(0,2318)</b>	4,762 <b>(0,3209)</b>	-1,876 <b>(0,2583)</b>

(\*) Os respectivos p-valores estão representados entre parêntesis.

Pelos resultados da Tabela 3, nota-se que, dentre os modelos de fórmulas analíticas, o modelo de Black apresentou piores resultados, uma vez que os modelos de Vasicek e CIR apresentaram superioridade, em termos estatísticos, em relação à metodologia utilizada pela Bolsa. Entretanto, quando comparados entre si, os testes indicaram equivalência entre os modelos de Vasicek e CIR (valores em negrito). Os modelos de redes recorrentes apresentaram os melhores resultados, em que as estatísticas indicaram que as redes são mais acuradas que os modelos analíticos. Porém, as redes recorrentes se mostraram equivalentes, quando avaliadas conjuntamente.

Os resultados obtidos são animadores, no sentido de que os modelos de redes recorrentes foram capazes de capturar o comportamento gerador do prêmio das opções de compra de IDI de forma mais acurada que os modelos tradicionais. Assim, estes modelos podem ser empregados como instrumentos que garantam confiança aos agentes do mercado, de forma a ampliar a liquidez nesse segmento de taxa de juros no Brasil, sendo de caráter fundamental, uma vez que a maioria das aplicações financeiras se relaciona direta ou indiretamente a variações da taxa de juros.



## 4 Conclusões

A precificação de opções, em geral, é de grande importância para os agentes de mercado, pois esses papéis são capazes de refletir a informação sobre o futuro de seus ativos objeto, diminuindo, assim, o grau de incerteza presente na economia. Entretanto, os preços desses ativos muitas vezes são obtidos por modelos que não incorporam características reais dos ativos subjacentes, fazendo com que seus resultados difiram dos preços de mercado. Dessa forma, alternativas baseadas em inteligência artificial, como as redes neurais, têm se mostrado promissoras na área de precificação de derivativos, sobretudo opções, demonstrada sua capacidade em se modelar padrões não-lineares de tipo entrada-saída.

No mercado de opções brasileiro, não existe um modelo teórico consistente, baseado em fórmulas fechadas, para se precificar as opções sobre o Índice de Depósitos Interfinanceiros de Um Dia (IDI), uma vez que a Bolsa utiliza o modelo de (Black, 1976), que é adequado para opções sobre futuros, não sobre taxa de juros. As opções de IDI são peculiares do nosso mercado e fazem com que seja necessária a adequação dos modelos tradicionais norte-americanos e europeus para o caso brasileiro. Com isso, (Vieira Neto e Pereira, 1999) e (Barbachan e Ornelas, 2003) propuseram modelos analíticos fechados para a precificação das opções de IDI, com base, respectivamente, nos modelos de (Vasicek, 1977) e (Cox et al., 1985). Entretanto, seus resultados ainda evidenciam que os preços teóricos não são capazes de explicar os preços das opções observados no mercado.

Dessa forma, este trabalho avaliou modelos de redes neurais recorrentes de tipo Elman e Jordan para se obter o prêmio das opções de compra de IDI, no período de janeiro de 2003 a julho de 2007. Esses modelos se basearam nas principais variáveis que afetam o preço da opção de IDI: valor do IDI, maturidade do contrato, preço de exercício, e volatilidade da taxa de juros de curto prazo. Neste caso, a volatilidade foi estimada com base em um modelo GARCH(p,q) que leva em consideração o efeito *duration*.

Avaliou-se, então, os modelos de Black, Vasicek e CIR, assim como os métodos de redes neurais para a valoração das opções de IDI. Os preços teóricos dos modelos foram comparados com os preços de fechamento dessas opções, observados no mercado. Para tanto, utilizamos métricas erro convencionais e testes estatísticos paramétricos e não-paramétricos para modelos competitivos de predição.

Os resultados inferiram a superioridade dos modelos de redes neurais em relação aos demais modelos, devido aos baixos erros apresentados. Além disso, para as opções *out-of-the-money*, as redes recorrentes apresentaram os melhores resultados em termos comparativos, uma vez que para as opções *at-the-money* e *in-the-money* os modelos tradicionais não apresentaram erros tão elevados quanto para as opções fora do dinheiro. Para validar os resultados, os testes estatísticos AGS, MGN e SIGN indicaram que a acurácia das redes recorrentes foi superior a dos modelos tradicionais, tomados como *benchmark*. Os mesmos testes indicaram que as redes recorrentes podem ser consideradas equivalentes para a precificação das opções de IDI. Resultado semelhante se verificou com os modelos convencionais, sendo que os modelos de Vasicek e CIR, de acordo com os testes estatísticos, podem também ser considerados equivalentes, mas superiores quando comparados com o modelo utilizado pela Bolsa (modelo de Black), que apresentou os piores resultados.

Com isso, verificou-se a inconsistência do modelo utilizado pela Bolsa, assim como dos modelos de (Vieira Neto e Pereira, 1999) e (Barbachan e Ornelas, 2003), para a precificação das opções de compra de IDI no período considerado. Entretanto, os modelos de redes neurais mostraram um grande potencial para o mesmo fim, resultando em estruturas que permitem captar o processo gerador dos prêmios desses papéis, o que se mostra como uma alternativa para se tentar ampliar a liquidez nesse mercado que, talvez, por deficiências teóricas, apresentam

baixo volume de negociações.

Este trabalho pode ser estendido na avaliação de outros modelos para estimação da volatilidade da taxa de juros de curto prazo, assim como na estimação dos parâmetros não-observáveis do modelo proposto por (Barbachan e Ornelas, 2003). Além disso, novas topologias de rede também podem incrementar os resultados, de forma a se obter um modelo que ainda melhor se ajuste ao comportamento dos prêmios de mercado dessas opções.

## Referências

- Almeida, C. I. R. e Vicente, J. V. (2003). Pricing and hedging brazilian fixed income options, *Anais do VI Encontro da Sociedade Brasileira de Finanças*, Vitória, Brasil.
- Almeida, L. A., Yoshino, J. A. e Schimer, P. P. (2003). Derivativos de renda fixa no brasil: modelo Hull-White, *Política e Planejamento Econômico* **33**(2): 299–333.
- Andreou, P. C., Charalambous, C. e Martzoukos, S. H. (2008). Pricing and trading european options by combining artificial neural networks and parametric models with implied parameters, *European Journal of Operational Research* **185**(3): 1415–1433.
- Ashley, R., Granger, J. W. C. e Schmalense, R. (1980). Advertising and aggregate consumption: an analysis of causality, *Econometrica* **48**(5): 1149–1167.
- Barbachan, J. S. F. e Ornelas, J. R. H. (2003). Apreçamento de opções de IDI usando o modelo CIR, *Estudos Econômicos* **33**(2): 287–323.
- Barbedo, C. H. S., Vicente, J. V. M. e Lion, O. M. B. (2009). Pricing asian interest rate options with a Three-Factor HJM model, *Working paper do Banco Central* .
- Bennel, J. e Sutcliffe, C. (2004). Black-Scholes versus artificial neural networks in pricing FTSE-100 options, *Intelligent Systems in Accounting, Finance and Management* **12**(4): 243–260.
- Black, F. (1976). The pricing of commodity contracts, *Journal of Financial Economics* **3**: 167–179.
- Black, F. e Scholes, M. (1973). The pricing of options and corporate liabilities, *Journal of Political Economy* **81**(3): 637–654.
- Burghardt, G. e Acworth, W. (2010). Volume trends: decline in the West, surge in the East, *The magazine of futures industry* pp. 20–21.
- Cox, J. C., Ingersoll, J. E. e Ross, S. A. (1985). A theory of the term structure of interest rates, *Econometrica* **55**: 385–407.
- Garcia, R. e Gençay, R. (2000). Pricing and hedging derivative securities with neural networks and a homogeneity hint, *Journal of Econometrics* **94**(1–2): 93–115.
- Gençay, R. e Qi, M. (2001). Pricing and hedging derivative securities with neural networks: bayesian regularization, early stopping and bagging, *IEEE Transactions on Neural Networks* **12**(4): 726–734.
- Ghaziri, H., Elfakhani, S. e Assi, J. (2001). Neural networks approach to pricing options, *Neural Network World* **10**(1–2): 271–277.

- Gluckstern, M. C., Francisco, G. e Jr, W. E. (2002). Aplicação do modelo de hull-white à precificação de opções de IDI, *Anais do II Encontro da Sociedade Brasileira de Finanças*, Rio de Janeiro, Brasil.
- Granger, J. W. C. e Newbold, P. (1977). *Forecasting economic time series*, New York: Academic Press.
- Haykin, S. (1998). *Neural networks: a comprehensive foundation*, 2 edn, New Jersey: Prentice Hall.
- Haykin, S. (2009). *Neural networks and learning machines*, 3 edn, New Jersey: Prentice Hall.
- Healy, J., Dixon, M., Read, B. e Cai, F. F. (2002). A data-centric approach to understanding the pricing of financial options, *European Physical Journal B* **27**(2): 219–227.
- Heath, D., Jarrow, R. e Morton, A. (1992). Bond pricing and the term structure of interest rates: a new methodology, *Econometrica* **60**: 77–105.
- Hull, J. e White, A. (1990). Pricing interest rate derivative securities, *Review of Financial Studies* **3**: 573–592.
- Junior, A. F., Greco, F., Lauro, C., Francisco, G., Rosenfeld, R. e Oliveira, R. (2003). Application of hull-white model to brazilian idi options, *Anais do III Encontro da Sociedade Brasileira de Finanças*, Florianópolis, Brasil.
- Kaastra, I. e Boyd, M. (1995). Designing a neural network for forecasting financial and economic time series, *Neurocomputing* **10**(3): 215–236.
- Lehmann, E. L. (1988). *Nonparametric statistical methods based on ranks*, New Jersey: Prentice Hall.
- Liang, X., Zhang, H., Xiao, J. e Chen, Y. (2009). Improving option price forecasts with neural networks and support vector regressions, *Neurocomputing* **72**: 3055–3065.
- Mandic, D. P. e Chambers, J. A. (2001). *Recurrent neural networks for prediction: Architectures, learning algorithms and stability*, Wiley.
- Samur, Z. I. e Temur, G. T. (2009). The use of artificial neural network in option pricing: the case of S&P 100 index options, *Proceedings of the 13th International AFIR Colloquium*, Maastricht - Holanda.
- Shapiro, A. (2003). Capital market applications of neural networks, fuzzy logic and genetic algorithms, *Proceedings of the 13th International AFIR Colloquium*, Maastricht, Holanda.
- Stone, M. (1974). Cross-Validation choice assessment of statistical prediction, *Journal of the Royal Statistical Society* **B36**: 111–133.
- Turnbull, S. M. e Milne, F. (1991). A simple approach to interest rate options, *Financial Studies* **4**(1): 87–120.
- Vasicek, O. A. (1977). An equilibrium characterization of the term structure, *Journal of Financial Economics* **5**: 177–188.
- Vieira Neto, C. A. e Pereira, P. V. (1999). Derivação de uma fórmula para o cálculo do preço livre de arbitragem das opções sobre o índice de depósitos financeiros de um dia - IDI, *Working Paper - IBMEC* **2**.

- Wang, Y. (2009). Using neural network to forecast stock index option price: a new hybrid GARCH approach, *Quality & Quantity* **43**(5): 833–843.
- Yao, J., Li, Y. e Tan, L. (2000). Option price forecasting using neural networks, *Omega* **28**(4): 455–466.