

Podemos prever a taxa de cambio brasileira? Evidência empírica utilizando inteligência computacional e modelos econométricos



Leandro dos Santos Coelho

André Alves Portela Santos

Newton Carneiro Affonso da Costa Jr.

Resumo

As abordagens de inteligência computacional, tais como sistemas nebulosos e redes neurais artificiais, têm-se gradualmente estabelecido como ferramentas robustas para a tarefa de aproximação de sistemas não-lineares complexos e previsão de séries temporais. Em aplicações envolvendo a área de Finanças, evidências empíricas anteriores indicam que modelos de inteligência computacional são mais precisos, dada sua maior capacidade em capturar não-linearidades e outros fatos estilizados presentes em séries financeiras. Nesse sentido, este artigo investiga a hipótese de que os modelos matemáticos de redes neurais perceptron multicamadas, redes neurais função de base radial e o sistema nebuloso Takagi-Sugeno (TAKAGI; SUGENO, 1985) são capazes de fornecer uma previsão fora-da-amostra mais acurada que os modelos auto-regressivos de médias móveis (ARMA) e auto-regressivo de médias móveis supondo heterocedasticidade condicional auto-regressiva (ARMA-GARCH). O desempenho de previsão um-passo-à-frente dos modelos foi comparado utilizando-se séries de retorno da taxa de câmbio real/dólar (R\$/US\$) com frequências de 15 minutos, 60 minutos, 120 minutos, diária e semanal. Resultados indicam que o desempenho dos modelos está diretamente relacionado à frequência observada das séries. Além disso, os modelos de redes neurais obtiveram um desempenho superior em relação aos demais modelos considerados. A avaliação da estratégia de negociação estabelecida com base nas previsões geradas pelos modelos indicou que estratégias baseadas em modelos de redes neurais forneceram retornos superiores em relação àquelas baseadas em modelos ARMA e ARMA-GARCH e também em relação à estratégia buy-and-hold.

Palavras-chave: Previsão. Inteligência computacional. Modelos não-lineares. Modelos lineares. Séries temporais. Redes neurais. Sistemas nebulosos.

1 Introdução

A literatura relacionada à previsão de séries temporais financeiras tem registrado, desde a década de 90, importantes avanços relacionados à incorporação de novas metodologias que tentam determinar padrões de relacionamentos presentes nos dados do mercado financeiro. Essas metodologias, em sua maioria, computacionalmente intensivas, caracterizam-se pela capacidade de identificação e previsão de sistemas dinâmicos não-lineares, ou seja, sistemas em que as variáveis do ambiente possuem um complexo padrão de inter-relacionamento que se altera ao longo do tempo.

A crescente constatação, reportada em uma série de trabalhos científicos, acerca da presença de um comportamento não-linear em séries financeiras tem gerado críticas quanto ao uso de modelos lineares na explicação

(e previsão) da dinâmica dos preços (PÉREZ-RODRÍGUEZ et al., 2005; ABHYANKAR et al., 1995). Pode-se argumentar que a metodologia econométrica tradicional, baseada em modelos com especificações simples e parâmetros constantes, como, por exemplo, os modelos ARMA, é incapaz de replicar a dinâmica inerente às séries econômicas e financeiras. Esse aspecto tem estimulado pesquisadores de diversas formações acadêmicas a aplicarem técnicas de inteligência computacional em uma série de problemas em Finanças, como, por exemplo, apreçamento de ativos (MORELLI et al., 2004), seleção de investimentos (EAKINS; STANSELL, 2003), teoria dos jogos (BARR; SARACENO, 2005) e previsão de séries temporais financeiras (MALHOTRA; MALHOTRA, 2002; TSAIH et al., 1998; KAASHOEK; VAN DJIK,

2002; OCAL, 2000; PÉREZ-RODRIGUEZ et al., 2005). Essas abordagens são adequadas para lidar com características intrínsecas às séries temporais financeiras, como, por exemplo, presença de múltiplos regimes no processo gerador de dados e não-linearidades. Entretanto, deve-se destacar que o debate acerca dos ganhos advindos do uso de modelos não-lineares ainda não atingiu um consenso (CLEMENTS et al., 2004), estimulando a continuidade da pesquisa nas áreas de seleção, estimação e avaliação de modelos não-lineares.

Dois abordagens relevantes, entre tantas outras, utilizadas para previsão de séries temporais são os sistemas nebulosos Takagi-Sugeno (TS) e as redes neurais *perceptron* multicamadas (MLP) e redes neurais função de base radial (RBF). O sistema nebuloso TS possui simultaneamente um alto grau de não-linearidade e estrutura simples (TAKAGI; SUGENO, 1985). O problema da identificação na modelagem envolvendo o sistema nebuloso TS consiste, basicamente, em duas partes: identificação da estrutura e identificação dos parâmetros. A identificação da estrutura está relacionada à determinação das partes antecedente e conseqüente das regras de ação, ou seja, consiste na partição do espaço de dados, extração de um número de regras e determinação da estrutura dos elementos de saída (equações), respectivamente. Finalmente, a tarefa de ajuste (ou aprendizado) dos parâmetros do sistema é realizada de modo a minimizar um determinado critério de desempenho baseado nos erros de saída do sistema. As redes neurais artificiais, originalmente inspiradas na funcionalidade das redes de neurônios biológicos, são, por sua vez, capazes de aprender relações funcionais complexas a partir de um conjunto limitado de dados de treinamento. Elas podem atuar como modelos “caixa-preta” para identificação de sistemas não-lineares multivariados e univariados e são passíveis de treinamento, utilizando dados de entrada e saída obtidos no sistema. A rede neural usual consiste em múltiplos elementos simples de processamento, chamados neurônios, interconexões entre eles e pesos atribuídos a essas interconexões. A informação relevante permanece armazenada nos pesos.

Nos últimos anos, pesquisadores (TSENG; TZENG, 2002; SHAPIRO, 2002; LEÓN et al., 2002) têm proposto um variado espectro de metodologias de identificação e previsão não-linear baseadas em sistemas nebulosos para lidarem com sistemas não-lineares que apresentam incertezas. As redes neurais, por sua vez, têm recebido considerável atenção nos últimos anos em função de suas habilidades em realizar aprendizagem, sendo aplicadas em um grande número de problemas na área de Finanças.

Este artigo investiga a hipótese de que os modelos de redes neurais MLP, redes neurais RBF e os sistemas nebulosos TS são capazes de fornecer uma previsão

fora-da-amostra mais acurada que os modelos auto-regressivos de médias móveis (ARMA) e os modelos ARMA supondo heterocedasticidade condicional auto-regressiva (ARMA-GARCH). O artigo está organizado da seguinte forma: a seção 2 descreve os modelos de previsão de séries temporais utilizados nesse estudo; nas seções 3 e 4 são apresentados a metodologia do estudo e os resultados, respectivamente; e, finalmente, na seção 5, são discutidos os resultados e apresentadas sugestões para trabalhos futuros.

2 Modelos de previsão de séries temporais

2.1 Modelos ARMA e ARMA-GARCH

Dentro da literatura de econometria financeira, duas classes de modelos amplamente estabelecidos na área de previsão de séries econômicas e financeiras são os modelos ARMA e ARMA-GARCH. De acordo com a metodologia proposta por Box e Jenkins (1976), um processo auto-regressivo de médias móveis ARMA (p, q) pode ser representado pela Equação 1.

$$y_t = \phi_0 + \phi_1 y_{t-1} + \dots + \phi_p y_{t-p} + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \dots + \theta_q \varepsilon_{t-q} \quad (1)$$

Sendo $\{\varepsilon_t\}$ um ruído branco com média zero e variância σ^2 . Os termos $\phi_0, \phi_1, \dots, \phi_p$ e $\theta_1, \dots, \theta_q$ são parâmetros que devem ser estimados. Este modelo é classificado linear na média.

Considerando o fato estilizado de ocorrência variância residual não-constante, foi também utilizado o modelo ARMA com um processo de heterocedasticidade condicional auto-regressiva – modelo ARMA (p, q) – GARCH (p, q) (MEADE, 2002). Este modelo é linear na média e não-linear na variância. A partir da Equação 1, é considerado o processo gerador de $\{\varepsilon_t\}$ como dado pelas Equações 2 e 3.

$$\varepsilon_t = v_t \sqrt{h_t}, v_t \sim \text{IID}(0,1) \quad (2)$$

$$h_t^2 = \alpha_0 + \sum_{i=1}^q \alpha_i \varepsilon_{t-i}^2 + \sum_{i=1}^p \beta_i h_{t-i}^2 \quad (3)$$

Em que h_t^2 é a variância condicional no período t . A condição para a existência do momento de quarta ordem e, portanto, de estacionaridade no modelo é $\sum_{i=1}^q \alpha_i + \sum_{i=1}^p \beta_i < 1$.

O modelo para a variância condicional descrito na Equação 3 é simétrico no que se refere aos impactos dos choques de volatilidade. Entretanto, séries temporais financeiras geralmente possuem respostas assimétricas aos choques de volatilidade. Nesse sentido, Zakoian (2002) propôs um modelo ARCH assimétrico (TARCH – Threshold ARCH) dado pela Equação 4.

$$h_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 h_{t-1}^2 + \lambda \varepsilon_{t-1}^2 I_{t-1} \quad (4)$$

Em que $I_{t-1} = 1$ se $\varepsilon_{t-1} < 0$. Desta forma, “boas” e “más” notícias têm impactos diferentes na variância condicional h_t^2 . De modo similar, Nelson (1991) propôs o modelo assimétrico EGARCH (Exponential GARCH) dado pela Equação 5.

$$\log(h_t^2) = \omega + \beta \log(h_{t-1}^2) + \alpha \left(\left| \frac{\varepsilon_{t-1}}{h_{t-1}} \right| - \sqrt{\frac{2}{\pi}} \right) + \gamma \frac{\varepsilon_{t-1}}{h_{t-1}} \quad (5)$$

A estimação dos parâmetros dos modelos ARMA pode ser obtida por meio de várias maneiras, como, por exemplo, mínimos quadrados ordinários ou pelas equações de Yule-Walker. No caso dos modelos ARMA-GARCH simétricos e assimétricos, a estimação dos parâmetros é usualmente obtida por meio da maximização da função de verossimilhança associada a cada modelo por meio de algoritmos de otimização numérica. Neste artigo, utilizou-se o *software* E-VIEWS® e MATLAB® para a estimação de parâmetros e previsão por meio de modelos ARMA e ARMA-GARCH. Um detalhamento das metodologias de estimação de parâmetros para essas classes de modelos é apresentado em Tsay (2002) e Hamilton (1994).

2.2 Modelos de redes neurais

2.2.1 Redes neurais *perceptron* multicamadas

As redes neurais fornecem, usualmente, conhecimento quantitativo não-paramétrico. São adequadas para identificação e aprendizado de sistemas que apresentam um comportamento complexo.

O *perceptron* multicamadas pode ser visto como a generalização do *perceptron* de uma camada. A rede neural *perceptron* multicamadas (MLP) contém três tipos de camadas: a camada de entrada, a(s) camada(s) escondida(s) ou intermediária(s) e a camada de saída. Qualquer neurônio de uma camada é capaz de estabelecer uma conexão com um neurônio da camada seguinte. Os neurônios da camada de entrada recebem sinais do ambiente externo e os transmitem para os neurônios da próxima camada. Os neurônios da camada de saída, por sua vez, enviam a informação recebida dos neurônios da(s) camada(s) intermediária(s) para o ambiente externo, conforme apresentado na Figura 1.

A saída de uma rede neural MLP com uma camada intermediária e n_1 neurônios intermediários (e um componente linear) pode ser descrita pela Equação 6.

$$y_{t+h} = w_0' Z_t + \sum_{i=1}^{n_1} w_{i1}' g(Z_t) + \varepsilon_{t+h} \quad (6)$$

Em que $g(x) = 1/(1 + e^x)$ e $Z_t = (1, y_t, y_{t-1}, \dots, y_{t-n+1})$.

O processo de treinamento de uma rede neural MLP é supervisionado e usualmente realizado por meio do algoritmo de retropropagação do erro ou *backpropagation* (BP). O algoritmo BP está baseado em uma regra de

correção do erro, e pode ser visto como uma generalização do algoritmo de filtro adaptativo ou um caso especial do algoritmo de mínimos quadrados.

O algoritmo BP é um método iterativo baseado no gradiente que busca minimizar o erro quadrático entre a saída observada e a saída desejada. Para minimizar a função-objetivo, o algoritmo BP utiliza uma técnica de busca baseada no gradiente, a regra delta generalizada. Entretanto, o algoritmo BP apresenta algumas limitações, como, por exemplo, baixa velocidade de aprendizado. Alternativas viáveis para o aumento da velocidade de treinamento são, por exemplo, o uso de procedimento heurístico para a modificação adaptativa dos parâmetros, prescrição de uma seleção específica de valores iniciais para os pesos e a utilização da segunda derivada no processo de treinamento. O algoritmo BP utiliza apenas a primeira derivada (curvatura) da superfície do erro, e seu desempenho é sensível à configuração do coeficiente de aprendizado. Com a utilização da segunda derivada (taxa de mudança da curvatura), o tempo de treinamento pode ser significativamente reduzido.

Neste trabalho, utilizou-se o algoritmo de treinamento Levenberg-Marquardt (LM) (LEVENBERG, 1944; MARQUARDT, 1963) o qual consiste em uma modificação do método de Newton e representa, freqüentemente, uma alternativa eficiente ao algoritmo de retropropagação do erro tradicional. A estratégia do algoritmo consiste na expansão da função $f(x, w) = t$ em uma série de Taylor. O objetivo do algoritmo é minimizar a função-custo, $J(w)$, com relação ao peso w . O algoritmo LM apresenta alto custo computacional e requer um grande número de operações.

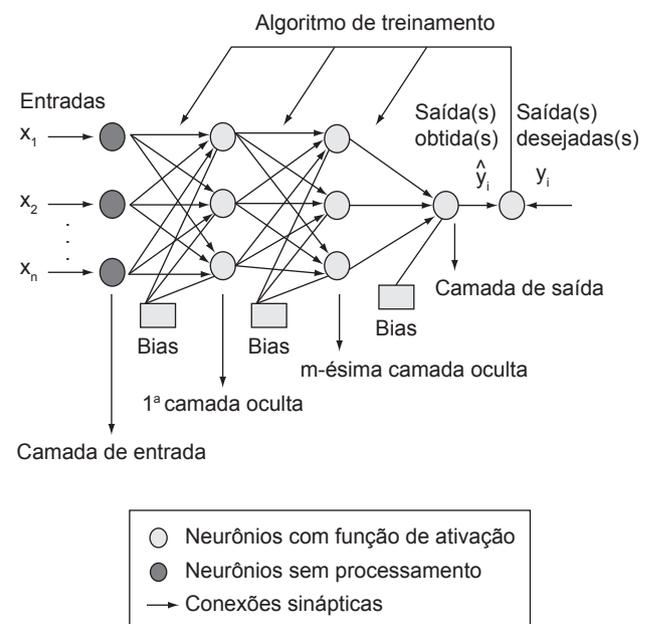


Figura 1. Rede neural *perceptron* multicamadas.

2.2.2 Redes neurais função de base radial

A aproximação não-linear realizada por uma rede neural função de base radial (RBF) com m neurônios com uma camada intermediária pode ser descrita na Equação 7.

$$y_{t+h} = \sum_{i=0}^m w_i \varphi(\|Z_t - c_i\|) \quad (7)$$

Sendo $\varphi(\cdot)$ a função de base radial e c_i o centro da i -ésima função de base radial. $\varphi(\cdot)$ é usualmente definida como uma função Gaussiana do tipo $\varphi(x) = \exp(-x^2 / 2\sigma^2)$ em que σ define a largura de cada função. A solução de mínimos quadrados para os pesos w_i satisfaz a relação $(A^T A)w = A^T b$ em que A é a matriz com elementos A_{ij} representando o elemento de saída do j -ésimo neurônio para a i -ésima entrada. O método de clusterização utilizado no experimento envolvendo as redes neurais RBF foi o método clássico k -médias. A estrutura geral de uma rede neural RBF pode ser vista na Figura 2.

2.3 Sistema nebuloso Takagi-Sugeno

A identificação de um modelo nebuloso Takagi-Sugeno (TS) envolve duas tarefas principais: ajuste dos parâmetros e otimização da estrutura. O procedimento de ajuste dos parâmetros busca a estimação de um conjunto factível de parâmetros para uma dada estrutura. O procedimento de otimização da estrutura busca: a) identificar uma estrutura ótima para modelos de ajuste local; b) identificar quais são as variáveis relevantes da parte antecedente das regras de ação; e c) determinar uma partição adequada do espaço de dados (MARINKE et al., 2005). O modelo TS consiste em um conjunto de regras lingüísticas SE-ENTÃO que podem ser representadas pela Equação 8:

$$R^{(j)} : \text{SE } z_1 \text{ E } A_1^j \text{ E } \dots \text{ E } z_m \text{ E } A_m^j \text{ ENTÃO,} \\ g_j = w_0^j + w_1^j u_1^j + \dots + w_{q_j}^j u_{q_j}^j \quad (8)$$

Em que as sentenças SE definem a parte antecedente (premissa) enquanto as funções da regra ENTÃO cons-

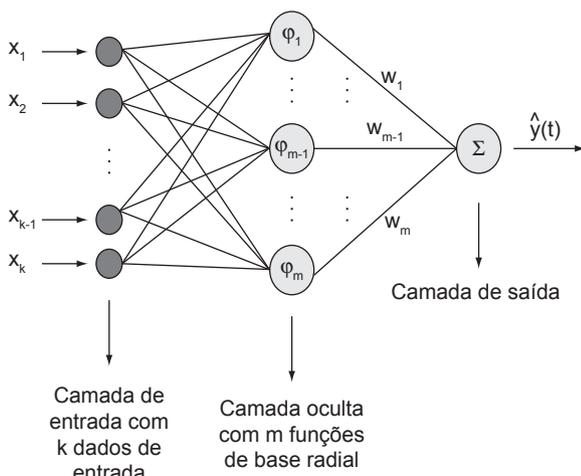


Figura 2. Estrutura geral de uma rede neural RBF.

tituem-se na parte conseqüente do sistema nebuloso; $\underline{z} = [z_1, \dots, z_m]^T$ é o vetor de entradas da parte premissa; e A_i^j são *labels* dos conjuntos nebulosos; $\underline{u} = [u_1^j, \dots, u_{q_j}^j]^T$ representa o vetor de entradas para a parte conseqüente de $R^{(j)}$ que é constituída de q_j termos; $g_j = g_j(\underline{u})$ denota a j -ésima regra de saída que é um polinômio linear dos termos de entrada do conseqüente u_i^j ; e $\underline{w}^j = [w_0^j, w_1^j, \dots, w_{q_j}^j]^T$ são os coeficientes polinomiais que formam o conjunto de parâmetros dos conseqüentes. A cada *label* lingüístico A_i^j é associada com a função de pertinência; enquanto $\mu_{A_i^j}(z_i)$ é descrita pela Equação 9:

$$\mu_{A_i^j}(z_i) = \exp\left[-\frac{1}{2} \frac{(z_i - m_{ij})^2}{\sigma_{ij}^2}\right] \quad (9)$$

Em que m_{ij} e σ_{ij} são o valor médio e o desvio-padrão da função de pertinência do tipo Gaussiana, respectivamente.

A união de todos estes parâmetros define o conjunto de parâmetros da premissa. O poder da regra $R^{(j)}$ representa seu grau de excitação e é regido pela Equação 10.

$$\mu^j(\underline{z}) = \mu_{A_1^j}(z_1) \cdot \mu_{A_2^j}(z_2) \cdot \dots \cdot \mu_{A_m^j}(z_m) \quad (10)$$

Os conjuntos nebulosos que dizem respeito a uma regra formam uma região nebulosa (*cluster*) dentro do espaço da premissa, $A_1^j \times A_2^j \times \dots \times A_m^j$, com uma distribuição de pertinência descrita pela Equação 10. Dado os vetores de entrada \underline{z} e \underline{u}^j , $j = 1, \dots, M$, a saída final do sistema nebuloso é inferida pela média ponderada das saídas locais $g_j(\underline{u}^j)$ (Equação 11).

$$y = \sum_{j=1}^M v_j(\underline{z}) \cdot g_j(\underline{u}^j) \quad (11)$$

Em que M denota o número de regras e $v_j(\underline{z})$ é a intensidade do poder normalizado de $R^{(j)}$ que é definida pela Equação 12:

$$v_j(\underline{z}) = \frac{\mu_j(\underline{z})}{\sum_{j=1}^M \mu_j(\underline{z})} \quad (12)$$

Neste artigo, a identificação da estrutura do modelo nebuloso TS foi baseada no algoritmo de agrupamento (*clustering*) nebuloso Gustafson-Kessel (GK) (GUSTAFSON; KESSEL, 1979) para a parte antecedente das regras e no algoritmo de mínimos quadrados para a parte conseqüente das regras. Um diagrama ilustrando a estrutura de um sistema nebuloso TS é mostrado na Figura 3.

2.4 Principais diferenças entre os modelos utilizados

Os modelos de redes neurais e sistemas nebulosos utilizam funções não-lineares para descrever e prever o comportamento do sistema, como pode ser visto nas Equações 6, 7 e 9. Dessa forma, esses modelos são consi-

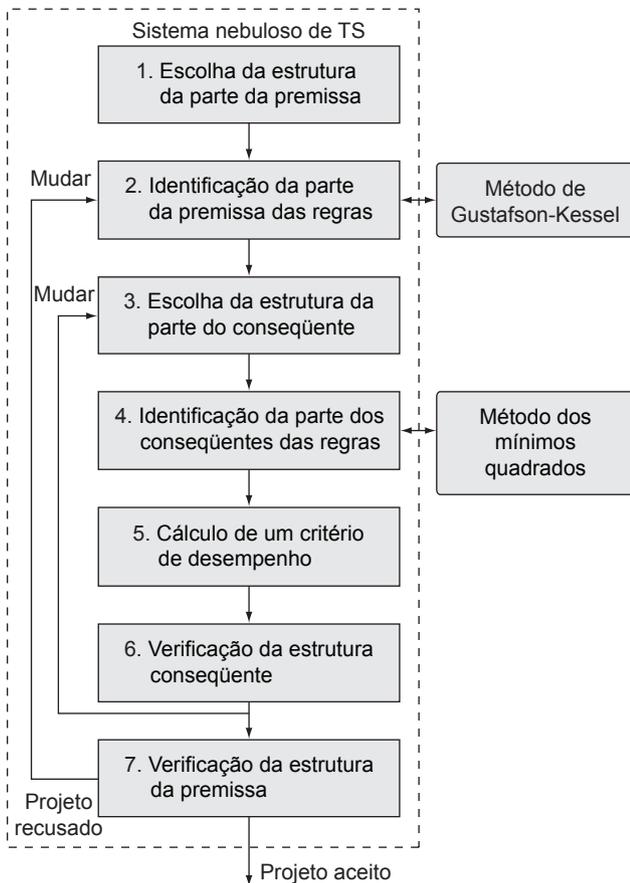


Figura 3. Estrutura de um sistema nebuloso TS.

derados não-lineares na média. Os modelos ARMA, por sua vez, descrevem o comportamento do sistema utilizando uma função linear. Portanto, essa classe de modelos é considerada linear na média. Finalmente, os modelos ARMA-GARCH utilizam uma função linear para descrever o comportamento do primeiro momento (média) e uma função não-linear para descrever o comportamento do segundo momento (variância). Dessa forma, esses modelos são classificados como não-linear na variância (CAMPBELL et al., 1997). Do ponto de vista da estimação, os modelos de redes neurais e sistemas nebulosos utilizados classificam-se como não-paramétricos, enquanto que os modelos ARMA e ARMA-GARCH classificam-se como paramétricos.

3 Metodologia e estatística descritiva das séries temporais

Neste artigo, foram utilizadas séries da primeira diferença dos logaritmos (log-retornos) da taxa de câmbio real/dólar (R\$/US\$) com frequências de 15 minutos, 60 minutos e 120 minutos de 01/01/2002 a 01/01/2003 e séries com frequência diária e semanal de 01/01/2000 a 01/01/2004. As Figuras 4 a 8 mostram o comportamento

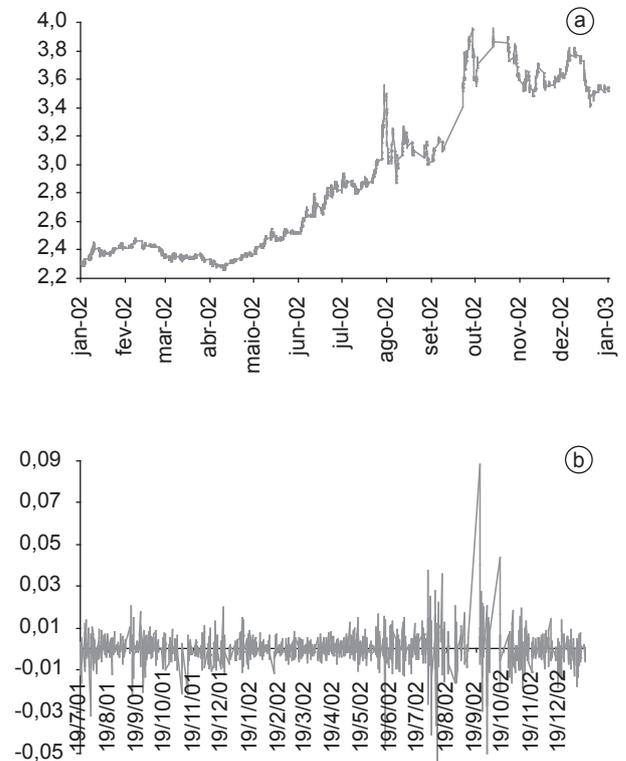


Figura 4. a) Taxa de câmbio brasileira com frequência de 15 minutos; e b) a primeira diferença dos logaritmos.

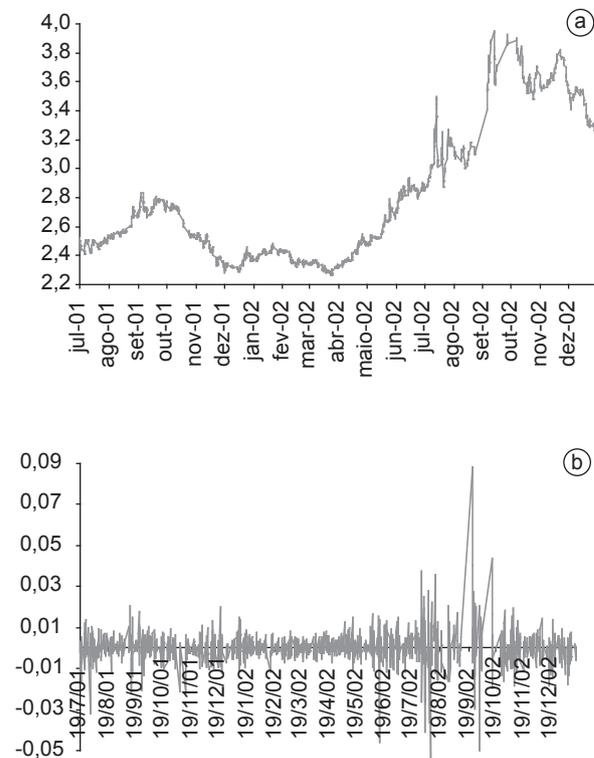


Figura 5. a) Taxa de câmbio brasileira com frequência de 60 minutos; e b) a primeira diferença dos logaritmos.

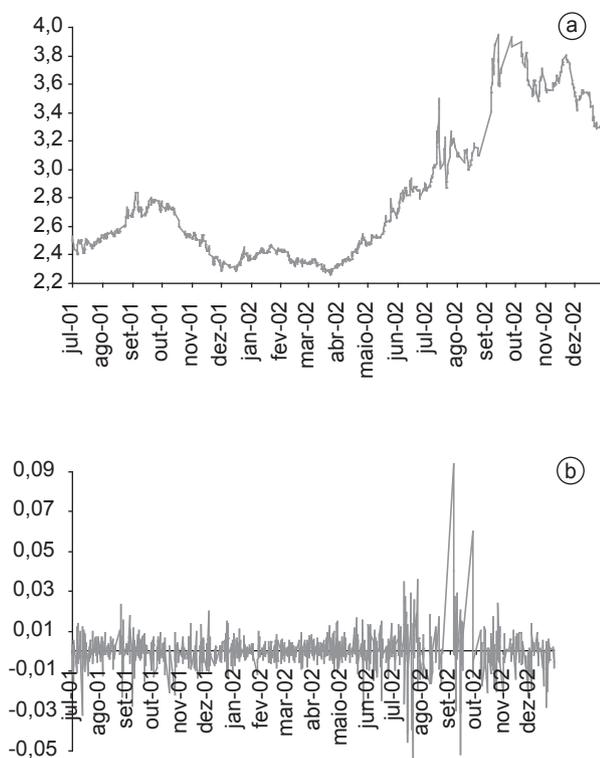


Figura 6. a) Taxa de câmbio brasileira com frequência de 120 minutos; e b) a primeira diferença dos logaritmos.

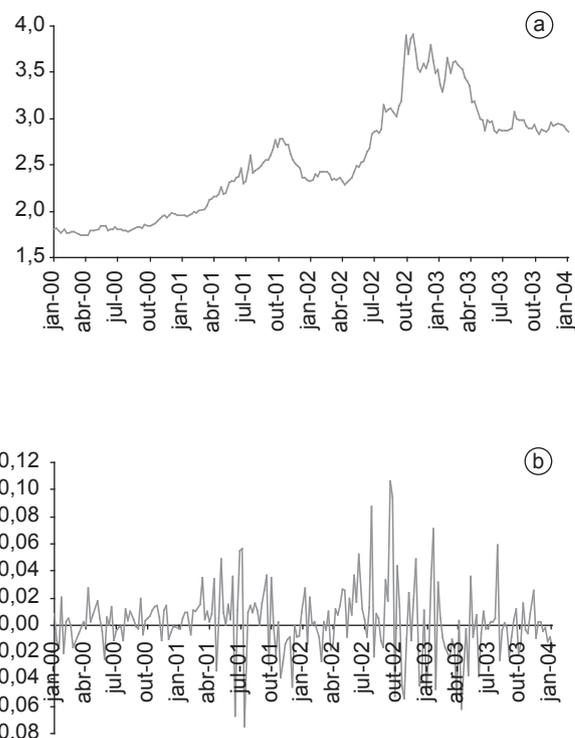


Figura 8. a) Taxa de câmbio brasileira com frequência de semanal; e b) a primeira diferença dos logaritmos.

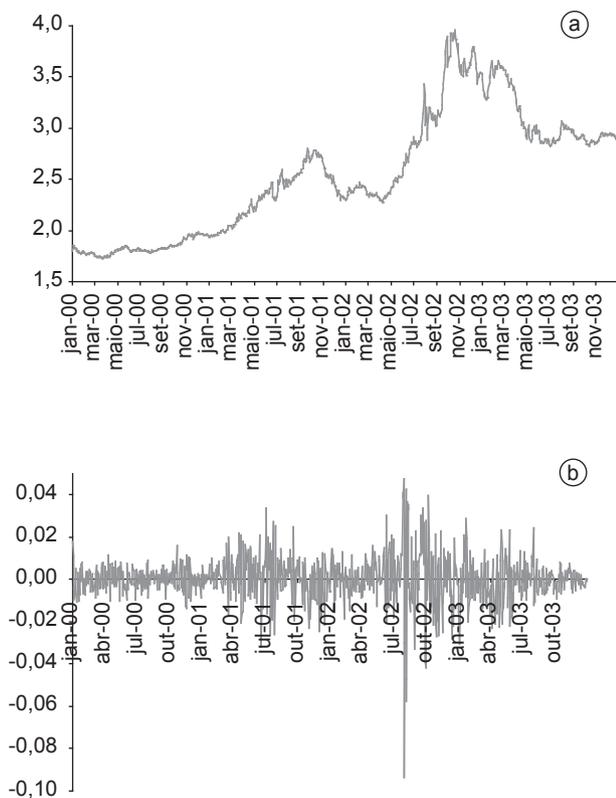


Figura 7. a) Taxa de câmbio brasileira com frequência diária; e b) a primeira diferença dos logaritmos.

da taxa de câmbio durante o período observado. Foram utilizados os primeiros 80% dos dados para estimação dos modelos e os últimos 20% para validação e previsão um-passo-à-frente fora-da-amostra. A Tabela 1 traz o número de observações utilizadas na fase de estimação e validação (previsão) para cada série.

A Tabela 2 traz a estatística descritiva das séries analisadas. O resultado do teste Jarque-Bera mostra que a hipótese de normalidade na distribuição dos retornos foi rejeitada em todas as séries em função da presença de elevada curtose e assimetria positiva (com exceção da série diária). A rejeição da hipótese nula do teste Jarque-Bera indica uma distribuição com uma incidência possivelmente maior de valores extremos, o que dificulta a previsão dessas séries temporais por parte dos modelos considerados neste estudo. O teste Dickey-Fuller Aumentado, por sua vez, mostrou que a hipótese nula de presença de raízes unitárias foi rejeitada em todas as séries, indicando que todas são estacionárias. O teste para presença de heterocedasticidade condicional auto-regressiva (ARCH), baseado no multiplicador de Lagrange, indicou a presença de ARCH em todas as séries.

O teste Brock-Dechert-Scheinkman (BDS) (BROCK et al., 1996) foi também realizado para verificar uma possível dependência não-linear nas séries temporais analisadas. A rejeição da hipótese nula de independência pode ser interpretada como uma evidência favorável ao

Tabela 1. Distribuição do número de observações de cada série histórica de acordo com as etapas de estimação e validação fora-da-amostra.

Série	15 minutos	60 minutos	120 minutos	Diária	Semanal
Estimação	4.515	1.180	730	804	180
Validação	1.128	294	183	201	20
Total	5.643	1.474	913	1.005	200

Tabela 2. Estatística descritiva para as séries de retornos da taxa de câmbio Real/Dólar, de acordo com a frequência observada.

	15 minutos	60 minutos	120 minutos	Diário	Semanal
Média	0,00008	0,0003	0,0004628	0,00047	0,002361
Desvio-padrão	0,00350	0,0067	0,0087056	0,010562	0,025464
Assimetria	2,07798	0,9037	0,9797082	-0,720195	0,432551
Curtose	101,103	34,521	25,784	12,047	5,704
Jarque-Bera	2.266.981	61.224	19.895	3.515	67,182
Prob.	0,000	0,000	0,000	0,000	0,000
Dickey-Fuller	-13,616	-7,300	-7,699	-14,886	-13,804
Prob.	0,000	0,000	0,000	0,000	0,000
ARCH LM	58,574	23,571	9,757	174,954	25,178
Prob.	0,000	0,000	0,001	0,000	0,000
BDS	0,025	0,032	0,034	0,036	0,032
z-score	16,64*	11,51*	9,57*	11,98*	4,29*
N. Observações	5.643	1.474	913	1.005	200
Período observado	01/01/2002 a 01/01/2003			01/01/2000 a 01/01/2004	

* indica significância estatística a um nível de 95% de confiança.

uso de modelos não-lineares para o ajuste da série. O teste foi aplicado às séries de resíduos obtidas a partir da estimação de um modelo AR(1)-GARCH(1,1). Os resultados indicaram a presença de não-linearidade em todas as séries. Além disso, foi constatado que quanto maior é a frequência da série, maior é a rejeição da hipótese nula de linearidade (com exceção da série diária).

Considerando a possibilidade de que o modelo AR(1)-GARCH(1,1) pode não estar bem ajustado às séries temporais utilizadas, aplicou-se o teste BDS também aos resíduos obtidos a partir de um modelo GARCH assimétrico (EGARCH). Entretanto, os resultados em termos de rejeição da hipótese nula de independência foram semelhantes aos obtidos quando utilizado um modelo AR(1)-GARCH(1,1). Além disso, foi observado o mesmo padrão em relação ao fato de que quanto maior é a frequência da série, maior é a rejeição da hipótese nula de linearidade.

Para este trabalho, utilizou-se o *software* MATLAB® para a implementação das rotinas computacionais para os modelos de redes neurais e sistemas nebulosos. Durante a implementação computacional, foram consideradas as representações matemáticas dinâmicas não-lineares NARX (*Non-linear Auto-Regressive with eXogenous inputs*) e NARMAX (*Non-linear Auto-Regressive Moving Average with eXogenous inputs*) (CHEN; BILLINGS, 1989), que podem ser descritas pelas Equações 13 a 16.

NARX

$$y(t) = f(y(t-1), \dots, y(t-n_y), u(t-1), \dots, u(t-n_u)) + e(t) \quad (13)$$

$$y(t) = f\left(\sum_{i=1}^{n_y} b_{yi}y(t-i) + \sum_{i=1}^{n_u} b_{ui}u(t-i)\right) + e(t) \quad (14)$$

NARMAX

$$y(t) = f(y(t-1), \dots, y(t-n_y), u(t-1), \dots, u(t-n_u), e(t-1), \dots, e(t-n_e)) + e(t) \quad (15)$$

$$y(t) =$$

$$f\left(\sum_{i=1}^{n_y} b_{yi}y(t-i) + \sum_{i=1}^{n_u} b_{ui}u(t-i) + \sum_{i=1}^{n_e} b_{ei}e(t-i)\right) + e(t) \quad (16)$$

Em que $y(t)$, $u(t)$ e $e(t)$ são, respectivamente, as saídas do sistemas, as entradas e o componente média móvel; n_y , n_u e n_e são as defasagens máximas das saídas, das entradas e do componente média móvel; $\{e_t\}$ é assumido ser ruído branco. A função f é uma função não-linear que descreve o comportamento do sistema.

O critério de informação de Akaike (AIC) foi utilizado para a escolha do número de termos e defasagens dos modelos ARMA e ARMA-GARCH simétricos e assimétricos. Especificamente, foi construída uma rotina computacional que estima para até 3 defasagens todas as possíveis diferentes especificações que os modelos

ARMA e ARMA-GARCH podem assumir. Na seção de resultados, apenas a estrutura que conseguiu melhor ajuste será mostrada. No caso dos modelos de redes neurais e sistemas nebulosos, o número de neurônios na camada intermediária, o número de funções de pertinência e o número de termos e defasagens foram obtidos por meio da minimização do erro quadrático médio entre a saída desejada e estimada dentro-da-amostra. Assim como no caso anterior, na seção de resultados, apenas a estrutura que conseguiu melhor ajuste será mostrada.

A avaliação do desempenho de previsão fora-da-amostra dos modelos foi realizada de acordo com os seguintes critérios:

Root mean squared error (RMSE):

$$RMSE = \sqrt{\frac{\sum_{j=1}^m e_{t+j}^2}{m}} \quad (17)$$

Em que e é o erro de previsão calculado para a observação $t + j$ e m é o número de observações.

Índice de Desigualdade de Theil (U-Theil):

$$U - Theil = \frac{\sqrt{\sum_{t=T+1}^{T+m} (\hat{y}_t - y_t)^2 / m}}{\sqrt{\sum_{t=T+1}^{T+m} (\hat{y}_t)^2 / m} + \sqrt{\sum_{t=T+1}^{T+m} (y_t)^2 / m}} \quad (18)$$

Em que \hat{y}_t é o valor previsto (estimado) para o período t , y_t é o valor observado e m é o número de observações. O índice U-Theil procura avaliar o ajustamento da série prevista à série original. Quanto mais próximo a 1, significa que a série estimada pelo modelo tem pouco ajustamento à série original.

Percentual de sinais corretamente previstos (SCP): quanto maior for o número de SCP, melhor será o modelo. Em outras palavras, quanto maior for o número de vezes em que o modelo consegue prever a direção do movimento, ou seja, quando o retorno do ativo será positivo ou negativo, melhor será seu desempenho.

Estatística de falha de previsão Pesaran-Timmermann: A estatística Pesaran-Timmermann (PT) é particularmente interessante quando a distribuição de probabilidade (intervalos de confiança) das previsões é de difícil derivação por métodos analíticos, como no caso de um grande número de modelos dinâmicos não-lineares (Pesaran e Timmermann, 1992). A estatística PT é dada pela Equação 19:

$$PT_n = \frac{\hat{P} - \hat{P}_*}{\{\text{var}(\hat{P}) - \text{var}(\hat{P}_*)\}^{1/2}} \sim N(0,1), \quad (19)$$

Em que \hat{P} é o percentual de vezes em que o sinal de mudança da variável foi previsto corretamente, \hat{P}_* é o valor esperado de \hat{P} , dado por $\hat{P}_* = P_+ P_+ + (1 - P_+)(1 - P_+)$ em que P_+ é o percentual calculado de sinais positivos e

P_+ é o percentual calculado de sinais positivos. $\text{var}(\hat{P})$ e $\text{var}(\hat{P}_*)$ são dados pelas Equações 20 e 21.

$$\text{var}(\hat{P}) = n^{-1} \hat{P}_* (1 - \hat{P}_*) \quad (20)$$

$$\begin{aligned} \text{var}(\hat{P}_*) &= n^{-1} (2\hat{P}_+ - 1)^2 \hat{P}_+ (1 - \hat{P}_+) + \\ &n^{-1} (2\hat{P}_- - 1)^2 \hat{P}_- (1 - \hat{P}_-) + \\ &4n^{-2} \hat{P}_+ \hat{P}_- (1 - \hat{P}_+) (1 - \hat{P}_-) \end{aligned} \quad (21)$$

4 Resultados

As Tabelas de 3 a 7 mostram os resultados da avaliação de previsão das séries temporais financeiras utilizadas neste estudo. O resultado mais importante consiste no fato de que os modelos de redes neurais e sistemas nebulosos utilizados no estudo tiveram um desempenho de previsão superior aos modelos ARMA e ARMA-GARCH em todas as séries analisadas. Para a série com frequência de 15 minutos, a melhor performance, dentre os modelos ARMA e ARMA-GARCH, consistiu em um índice U-Theil de 0,867 e uma previsão correta da direção dos movimentos dos retornos de 44%, sem significância estatística. O melhor modelo dentre as redes neurais e sistemas nebulosos, por sua vez, obteve um índice U-Theil de 0,786 e previu corretamente 54% dos movimentos dos retornos, com uma significância estatística de 99% calculada pela estatística PT.

Os resultados indicam ainda que quanto menor é a frequência da série, maior é a capacidade dos modelos em realizar boas previsões. A análise das Tabelas de 3 a 7 revela que à medida que a frequência observada das séries vai caindo, a qualidade das previsões aumenta. Uma possível explicação para essa constatação reside no fato de que as séries de alta frequência apresentam um nível de não-linearidade maior (Tabela 2), o que torna a previsão uma tarefa mais difícil. Além disso, séries de alta frequência estão mais contaminadas com ruídos advindos da microestrutura do mercado, o que também dificulta sua previsão. Nesse sentido, as piores previsões foram obtidas para as séries de retorno com frequência de 15 minutos (Tabela 3) enquanto que as melhores previsões foram obtidas para a série de retornos com frequência semanal (Tabela 7). Nesse último caso, o modelo de redes neurais MLP-NARMAX obteve um índice U-Theil de 0,469 e previu corretamente 82% da direção do movimento com uma significância estatística de 99%, enquanto que, dentre os modelos ARMA e ARMA-GARCH, o melhor desempenho consistiu em um índice U-Theil de 0,654 e uma previsão correta da direção dos movimentos dos retornos de 67%, com 95% de significância estatística.

Deve ser destacado que o tamanho do erro de previsão (RMSE) mostrou-se um critério pouco útil para escolher qual é o modelo de melhor desempenho. No caso, por

Tabela 3. Resultados da previsão para a série temporal da taxa de câmbio brasileira com frequência de 15 minutos.

Modelo	RMSE	U-Theil	SCP (%)	PT
AR	0,005	0,931	44	-3,08
MA	0,005	0,930	44	-3,01
ARMA	0,005	0,894	42	-5,35
ARMA-GARCH	0,005	0,867	44	-4,74
MLP-NARX	0,005	0,783	53	2,19*
MLP-NARMAX	0,005	0,786	54	3,62*
RBF-NARX	0,005	0,828	51	1,91
RBF-NARMAX	0,005	0,815	48	1,56
TS-NARX	0,03	0,957	54	1,14
TS-NARMAX	0,04	0,940	51	0,03

* indica que a estatística PT é significante a 95% de confiança.

Tabela 4. Resultados da previsão para a série temporal da taxa de câmbio brasileira com frequência de 60 minutos.

Modelo	RMSE	U-Theil	SCP (%)	PT
AR	0,01	0,843	51	0,67
MA	0,01	0,830	50	0,19
ARMA	0,01	0,817	51	0,62
ARMA-GARCH	0,01	0,877	49	-0,47
MLP-NARX	0,001	0,652	59	3,56*
MLP-NARMAX	0,001	0,640	60	3,31*
RBF-NARX	0,001	0,740	59	3,10*
RBF-NARMAX	0,001	0,725	60	3,21*
TS-NARX	0,07	0,953	54	2,69*
TS-NARMAX	0,07	0,900	55	2,03*

* indica que a estatística PT é significante a 95% de confiança.

Tabela 5. Resultados da previsão para a série temporal da taxa de câmbio brasileira com frequência de 120 minutos.

Modelo	RMSE	U-Theil	SCP (%)	PT
AR	0,013	0,782	55	1,46
MA	0,013	0,762	54	1,01
ARMA	0,013	0,735	49	-0,20
ARMA-GARCH	0,013	0,844	53	0,70
MLP-NARX	0,01	0,637	63	3,74*
MLP-NARMAX	0,01	0,617	61	3,19*
RBF-NARX	0,01	0,753	58	2,42*
RBF-NARMAX	0,01	0,714	58	2,32*
TS-NARX	0,09	0,938	60	3,79*
TS-NARMAX	0,09	0,837	55	2,17*

* indica que a estatística PT é significante a 95% de confiança.

exemplo, das séries com frequência de 15 minutos e semanal, o RMSE dos modelos ARMA e ARMA-GARCH foi maior que o RMSE dos modelos de redes neurais e sistemas nebulosos. Esse resultado corrobora o estudo de Leitch e Tanner (1991), os quais concluíram que, além do tamanho erro de previsão, outros critérios devem ser levados em conta na escolha de qual é o melhor modelo de previsão.

A Tabela 8 traz um resumo dos resultados da avaliação de previsão para o modelo de melhor desempenho em

cada série. O modelo de redes neurais MLP com representação matemática NARMAX foi o que obteve os melhores resultados de previsão nas séries temporais analisadas.

A Tabela 9 traz os resultados da estratégia de negociação estabelecida com base nas previsões geradas pelos modelos. São comparados o retorno acumulado, retorno médio, desvio-padrão, índice de Sharpe e índice de Sharpe ajustado da estratégia de negociação baseada

Tabela 6. Resultados da previsão para a série temporal da taxa de câmbio brasileira com frequência diária.

Modelo	RMSE	U-Theil	SCP (%)	PT
AR	0,007	0,781	52	1,96*
MA	0,008	0,787	51	1,49
ARMA	0,008	0,771	47	0,19
ARMA-GARCH	0,008	0,867	54	1,55
MLP-NARX	0,001	0,700	61	3,85*
MLP-NARMAX	0,001	0,701	58	3,21*
RBF-NARX	0,001	0,821	60	3,14*
RBF-NARMAX	0,01	0,732	58	2,42*
TS-NARX	0,10	0,953	57	1,17
TS-NARMAX	0,10	0,772	59	2,34*

* indica que a estatística PT é significante a 95% de confiança.

Tabela 7. Resultados da previsão para a série temporal da taxa de câmbio brasileira com frequência semanal.

Modelo	RMSE	U-Theil	SCP (%)	PT
AR	0,013	0,681	57	1,78
MA	0,013	0,654	67	2,33*
ARMA	0,013	0,624	62	2,06*
ARMA-GARCH	0,013	0,658	57	0,74
MLP-NARX	0,02	0,458	75	3,27*
MLP-NARMAX	0,02	0,469	82	4,32*
RBF-NARX	0,02	0,698	62	2,11*
RBF-NARMAX	0,13	0,796	65	1,50
TS-NARX	0,13	0,910	62	2,48*
TS-NARMAX	0,13	0,904	60	2,00*

* indica que a estatística PT é significante a 95% de confiança.

Tabela 8. Resumo dos resultados de previsão para o modelo de melhor desempenho em cada série.

Série	15 minutos	60 minutos	120 minutos	Diária	Semanal
Melhor modelo	MLP-NARMAX	MLP-NARMAX	MLP-NARX	MLP-NARMAX	MLP-NARMAX
RMSE	0,005	0,001	0,01	0,001	0,02
U-theil	0,786	0,640	0,617	0,700	0,469
SCP	54%	60%	61%	61%	82%
PT	3,62*	3,31*	3,19*	3,85*	4,32*

* indica que a estatística PT é significante a 95% de confiança.

no modelo de melhor ajuste para cada série e também na estratégia ingênua do tipo *buy-and-hold*. O índice de Sharpe ajustado, proposto por Cerny (2004), consiste em uma correção do índice de Sharpe tradicional para distribuições de retorno que não cumprem com os supostos de normalidade. Seu cálculo toma como base um(a) investidor(a) com função de utilidade quadrática. Para cada série de retornos X gerada pela estratégia de negociação, o índice de Sharpe ajustado será diferente to índice de Sharpe tradicional sempre que as condições da Equação 22 forem violadas (*bliss point conditions*):

$$\begin{aligned} \max(X) \cdot E(X) < E(X^2), \text{ caso } E(X) > 0 \\ \min(X) \cdot E(X) > E(X^2), \text{ caso } E(X) < 0 \end{aligned} \quad (22)$$

Em que $E(X)$ e $E(X^2)$ indicam a média da série de retornos e a média das observações ao quadrado, respectivamente. Finalmente, é assumido que os custos de transação da estratégia estejam embutidos no *spread* entre a compra e a venda e bem como a possibilidade de se fazer venda a descoberto.

A estratégia de negociação está baseada nas seguintes regras:

- Regra 1: Se o valor previsto para o próximo período for positivo, computa-se uma compra;
- Regra 2: Se o valor previsto para o próximo período for negativo, computa-se uma venda.

Os resultados da estratégia de negociação indicam que as estratégias baseadas em modelos de redes neurais obtiveram retornos (acumulados e médios) superiores àqueles

Tabela 9. Resultados da estratégia de negociação.

Estratégia/ Modelo	Retorno acumulado (%)	Retorno médio (%)	Desvio-padrão (%)	Índice de Sharpe	Índice de Sharpe ajustado
15 minutos					
AR	12,5	0,012	0,49	0,0239	0,0239
MLP-NARMAX	48,1	0,036	0,49	0,0737	0,0925
<i>Buy-and-hold</i>	13	0,012	0,49	0,0243	0,0243
60 minutos					
ARMA	36,5	0,022	1,001	0,0218	0,0218
MLP-NARMAX	92,7	0,224	0,982	0,2291	0,2321
<i>Buy-and-hold</i>	13,6	0,048	0,048	0,0483	0,0483
120 minutos					
MA	25,1	0,110	1,330	0,0826	0,0826
MLP-NARMAX	80,2	0,420	1,270	0,3320	0,3205
<i>Buy-and-hold</i>	12,4	0,073	1,332	0,0549	0,0549
Diária					
AR	36,9	0,092	0,770	0,1198	0,1198
MLP-NARMAX	50	0,202	0,744	0,2726	0,2726
<i>Buy-and-hold</i>	-18,1	-0,085	0,768	-0,1114	-0,1114
Semanal					
ARMA-GARCH	6,3	0,435	1,081	0,4120	0,4120
MLP-NARMAX	55,2	1,192	1,823	0,6602	0,4548
<i>Buy-and-hold</i>	-6,2	-0,440	2,140	-0,2084	-0,2084

obtidos em estratégias baseadas em modelos ARMA e ARMA-GARCH para todas as séries temporais analisadas. Além disso, as duas medidas de performance ajustadas ao risco utilizadas indicaram que as estratégias de negociação baseadas em modelos de redes neurais obtiveram uma melhor relação risco-retorno em todos os casos.

5 Conclusões e sugestões para trabalhos futuros

Este trabalho teve como objetivo a comparação do desempenho de previsão dos modelos de redes neurais MLP, redes neurais RBF e os sistemas nebulosos TS em relação aos modelos ARMA e ARMA-GARCH, os quais estão amplamente estabelecidos na literatura relacionada à econometria financeira. Foram utilizadas séries de retorno da taxa de câmbio real/dólar com frequências de 15 minutos, 60 minutos, 120 minutos, diária e semanal.

Os resultados indicam que os modelos de redes neurais obtiveram um desempenho de previsão superior aos demais modelos considerados, corroborando evidências empíricas anteriores que atestam a robustez dessa abordagem quando aplicada à aproximação e previsão de séries financeiras. Além disso, verificou-

se que a consistência das previsões está diretamente relacionada à frequência com que a série é observada, possivelmente devido à presença de não-linearidades e de ruídos advindos da microestrutura do mercado, como, por exemplo, o *spread* entre compra e venda e o fluxo de ordens de compra e venda, que afetam a dinâmica das séries de alta frequência. Com relação à estratégia de negociação estabelecida com base nas previsões geradas pelos modelos, constatou-se que as estratégias baseadas em modelos de redes neurais obtiveram retornos superiores às demais estratégias. Entretanto, quando o risco da estratégia é levado em conta, o resultado torna-se misto.

Deve-se destacar que a otimização da estrutura, precisão, complexidade computacional e ordens dos vetores de entrada dos modelos baseados em inteligência computacional devem ser analisadas em maior profundidade em trabalhos futuros. Na área de avaliação do desempenho de previsão, alguns estudos têm sugerido que a avaliação baseada em previsões pontuais não é capaz de discriminar corretamente os ganhos advindos do uso de modelos mais sofisticados (CLEMENTS et al., 2004). Neste sentido, a utilização de critérios de avaliação de desempenho mais robustos, como, por exemplo, previsão de densidades, tem recebido grande destaque na literatura especializada recente.

Can we forecast Brazilian exchange rates? Empirical evidences using computational intelligence and econometric models

Abstract

Computational intelligence approaches, such as artificial neural networks and fuzzy systems, have become popular tools in approximating complicated nonlinear systems and time series forecasting. In Finance applications, there is evidence that these computational intelligence models are able to provide a more accurate forecast given their capacity for capturing nonlinearities and other stylized facts of financial time series. Thus, this paper investigates the hypothesis that the mathematical models of multilayer perception, radial basis function neural networks (NN), and the Takagi-Sugeno (TS) fuzzy systems are able to provide a more accurate out-of-sample forecast than the traditional AutoRegressive Moving Average (ARMA) and ARMA Generalized AutoRegressive Conditional Heteroskedasticity (ARMA-GARCH) models. Using a series of Brazilian exchange rate (R\$/US\$) returns with 15 minutes, 60 minutes, 120 minutes, daily and weekly basis, the one-step-ahead forecast performance is compared. The results indicate that forecast performance is strongly related to the series' frequency, possibly due to nonlinearities effects. Besides, the forecasting evaluation shows that NN models perform better than the ARMA and ARMA-GARCH ones. In the trade strategy based on forecasts, NN models achieved higher returns when compared to a buy-and-hold strategy and to the other models considered in this study.

Keywords: Forecasting. Nonlinear models. Linear models. Time series. Neural networks. Fuzzy systems.

Referências bibliográficas

- ABHYANKAR, A.; COPELAND, L.; WONG, W. Nonlinear dynamics in real-time equity markets indices: evidence from United Kingdom. **The Economic Journal**, v. 105, n. 431, p. 864-880, 1995.
- BARR, J.; SARACENO, F. Cournot competition, organisation and learning. **Journal of Economic Dynamics and Control**, v. 29, n. 1-2, p. 277-295, 2005.
- BOLLERSLEV, T.; CHOU, R.; KRONER, K. ARCH modeling in finance: a review of the theory and empirical work. **Journal of Econometrics**, v. 52, n.1-2, p. 5-59, 1992.
- BOX, G.; JENKINS, G. **Time series analysis: forecasting and control**. San Francisco: Holden Day, 1976.
- BROCK, W.; DECHERT, D.; LEBARON, B.; SCHEINKMAN, J. A test for independence based on a correlation dimension. **Econometric Review**, v. 15, n. 3, p. 197-235, 1996.
- CAMPBELL, J.; LO, A.; MACKINLAY, A. **The econometrics of financial markets**. 1 ed. Princeton: Princeton university press, 1997.
- CERNY, A. **Mathematical techniques in finance: tools for incomplete markets**. Princeton: University Press, 2004.
- CHEN, S.; BILLINGS, S. Representation of nonlinear systems: the NARMAX model. **International Journal of Control**, v. 49, p. 1012-1032, 1989.
- CLEMENTS, M.; FRANSES, P.; SWANSON, N. Forecasting economic and financial time-series with non-linear models. **International Journal of Forecasting**, v. 20, p. 169-183, 2004.
- EAKINS, S.; STANSELL, S. Can value-based stock selection criteria yield superior risk-adjusted returns: an application of neural networks. **International Review of Financial Analysis**, v. 12, n. 1, p. 83-97, 2003.
- GUSTAFSON, D. E.; KESSEL, W. C. Fuzzy Clustering with a Fuzzy Covariance Matrix. In: IEEE CONFERENCE ON CONTROL AND DECISION, 1979. **Proceedings...**, San Diego, CA, USA, p. 761-766, 1979.
- KAASHOEK, J.; VAN DIJK, H. Neural network pruning applied to real exchange rate analysis. **Journal of Forecasting**, v. 21, p. 559-577, 2002.
- LEITCH, G.; TANNER, E. Economic forecast evaluation: profits versus the conventional error measures. **American Economic Review**, v. 81, n. 3, p. 580-590, 1991.
- LEÓN, T.; LIERN, V.; VERCHER, E. Viability of infeasible portfolio selection problems: a fuzzy approach. **European Journal of Operational Research**, v. 139, n. 1, p. 178-189, 2002.
- LEVENBERG, K. A method for the solution of certain non-linear problems in least squares. **Quart. Applied Mathematics**, v. 2, n. 2, p. 164-168, 1944.
- MALHOTRA, R.; MALHOTRA, D. Differentiating between good credits and bad credits using neuro-fuzzy systems. **European Journal of Operational Research**, v. 136, n. 1, p. 190-211, 2002.
- MARINKE, R.; ARAUJO, E.; COELHO, L. S.; MATIKO, I. Particle swarm optimization (PSO) applied to fuzzy modeling in a thermal-vacuum system. In: INTERNATIONAL CONFERENCE ON HYBRID INTELLIGENT SYSTEMS (HIS), 5, 2005, Brazil. **Proceedings...**, Brazil, p. 67-72, 2005.
- MARQUARDT, D. W. An algorithm for least-squares estimation of nonlinear parameters. **SIAM J. of Applied Mathematics**, v. 11, n. 2, p. 431-441, 1963.
- MEADE, N. A comparison of the accuracy of short term foreign exchange forecasting methods. **International Journal of forecasting**, n. 18, p. 67-83, 2002.
- MORELLI, M.; MONTAGNA, G.; NICROSINI, O.; TRECCANI, M.; FARINA, M.; AMATO, P. Pricing financial derivatives with neural networks. **Physica A**, v. 388, n.1-2, p. 160-165, 2004.

- NELSON, D. Conditional Heteroskedasticity in Asset Returns: A New Approach. **Econometrica**, v. 59, n. 2, p. 347-370, 1991.
- OCAL, N. Nonlinear models for UK macroeconomic time series. **Studies in nonlinear dynamics and econometrics**, v. 4, n. 3, p. 123-135, 2000.
- PÉREZ-RODRÍGUEZ, J.; TORRA, S.; ANDRADA-FÉLIX, J. STAR and ANN models: forecasting performance on the Spanish 'Ibex-35' stock index. **Journal of Empirical Finance**, n. 12, p. 490-509, 2005.
- PESARAN, M., TIMMERMANN, A. A simple nonparametric test of predictive performance. **Journal of Business & Economic Statistics**, v. 10, n. 4, p. 561-565, 1992.
- SHAPIRO, A. F. The merging of neural networks, fuzzy logic, and genetic algorithms. **Insurance: Mathematics and Economics**, v. 31, n. 1, p. 1151-31, 2002.
- TAKAGI, T.; SUGENO, M. Fuzzy identification of systems and its application to modeling and control. **IEEE Transactions on Systems, Man, and Cybernetics**, v. 15, n. 1, p. 116-132, 1985.
- TSAIH, R.; HSU, Y.; LAI, C. Forecasting S&P 500 stock index futures with a hybrid AI System. **Decision Support Systems**, v. 23, n. 2, p. 161-174, 1998.
- TSENG, F. M.; TZENG, G. H. A fuzzy seasonal ARIMA model for forecasting. **Fuzzy Sets and Systems**, v. 126, n. 3, p. 367-376, 2002.
- ZAKOIAN, J. M. Threshold heteroskedastic models. **Journal of Economic Dynamics and Control**, v. 18, n. 5, p. 931-955, 2002.

Sobre os autores

Leandro dos Santos Coelho

Laboratório de Automação e Sistemas, PPGEPS, CCET, Pontifícia Universidade Católica do Paraná – PUCPR,
Rua Imaculada Conceição, 1155, CEP 80215-901, Curitiba, PR, Brasil,
e-mail: leandro.coelho@pucpr.br

André Alves Portela Santos

Universidad Carlos III de Madrid,
C/ Madrid, 126 – 28903, Getafe (Madrid), Espanha,
e-mail: andre.alves@uc3m.es

Newton Carneiro Affonso da Costa Jr.

Universidade Federal de Santa Catarina – UFSC,
CP 476, CEP 88040-900, Florianópolis, SC, Brasil,
e-mail: newton@cse.ufsc.br

Agradecimentos: Os autores agradecem aos revisores anônimos quanto a sugestões no texto.

Recebido em 27/5/2006
Aceito em 12/9/2008