



Revista ADM.MADE

Revista do Mestrado em Administração e
Desenvolvimento Empresarial - Universidade
Estácio de Sá

Revista ADM.MADE, Rio de Janeiro, ano 12, v.16, n.2, p.16-31, maio/agosto, 2012

Revista do Mestrado em Administração e Desenvolvimento Empresarial da Universidade

Estácio de Sá – Rio de Janeiro (MADE/UNESA). ISSN: 2237-5139

Conteúdo publicado de acesso livre e irrestrito, sob licença Creative Commons 3.0.

Editora responsável: Isabel de Sá Affonso da Costa

Organizador do número temático: Marco Aurélio Carino Bouzada

Previsão da Taxa de Câmbio Real por Dólar através de Programação Genética Multigênica

Adriano Soares Koshiyama¹

Carlos Magno Catharino Olsson Valle²

Douglas Mota Dias³

Uma versão preliminar deste artigo foi apresentada no XV SPOLM - Simpósio de Pesquisa Operacional e Logística da Marinha, agosto de 2012, Rio de Janeiro – RJ. Disponível em <https://www.casnav.mar.mil.br/spolm/index.html>.

Artigo recebido em 13/11/2012 e aprovado em 07/12/2012. Artigo convidado à submissão e avaliado em *double blind review*.

¹ Economista pela Universidade Federal Rural do Rio de Janeiro (UFRRJ). Mestrando em Engenharia Elétrica na área de Métodos de Apoio à Decisão pela Pontifícia Universidade Católica do Rio de Janeiro (DEE/PUC-Rio). Endereço: PUC-Rio - Rua Marquês de São Vicente, 225, Gávea - Rio de Janeiro/RJ. Email: adriano@ele.puc-rio.br.

² Graduando em Engenharia de Controle e Automação pela Pontifícia Universidade Católica do Rio de Janeiro (DEE/PUC-Rio). Endereço: PUC-Rio - Rua Marquês de São Vicente, 225, Gávea - Rio de Janeiro/RJ. Email: cmcat46@gmail.com.

³ Doutor em Engenharia Elétrica pela Pontifícia Universidade Católica do Rio de Janeiro (PUC-Rio). Pesquisador de pós-doutorado pela Pontifícia Universidade Católica do Rio de Janeiro (DEE/PUC-Rio). Endereço: PUC-Rio - Rua Marquês de São Vicente, 225, Gávea - Rio de Janeiro/RJ. Email: douglasm@ele.puc-rio.br.

Previsão da Taxa de Câmbio Real por Dólar através de Programação Genética Multigênica

A importância da taxa de câmbio tem sido frequentemente abordada por economistas, devido aos impactos positivos e negativos de sua variação no processo de desenvolvimento econômico de um país. Um modelo de previsão que capte e que se adapte a essas variações repentinas é de elevada significância para a formulação concisa de um planejamento ou uma decisão estratégica. Utilizando a programação genética multigênica, este trabalho tem, por objetivo, examinar o comportamento futuro da taxa de câmbio do real por dólar americano (R\$/US\$), a partir de dados de janeiro de 2000 a abril de 2012. Os modelos Holt-Winters e ARIMA também foram ajustados para efeitos de comparação. Os resultados demonstraram a potencialidade da programação genética multigênica perante aos demais modelos ajustados, e a acurácia em previsões de curto-prazo. Segundo o método, a taxa de câmbio R\$/US\$ deve oscilar em torno de 1,87 a 1,92 nos meses de maio, junho e julho de 2012.

Palavras-chave: programação genética (PG); indivíduo multigênico; séries temporais; taxa de câmbio.

Keywords: genetic programming (GP); multigenic individual; time series; exchange rate.

Forecasting The Exchange Rate Real Per US Dollar By Multi-Gene Genetic Programming

The importance of the exchange rate has been frequently discussed by economists, due to positive and negative impacts of its variation in the process of economic development of a country. A prediction model that captures and adapts to sudden changes is highly significant for the concise formulation of a plan, or a strategic decision. Using multi-gene genetic programming, this study aims to examine the future behavior of the exchange rate Real per US dollar (R\$/US\$), based on data from January 2000 to April 2012. Holt-Winters and ARIMA models were also been adjusted for the purpose of comparison. The results demonstrated the potential of multi-gene genetic programming against the others adjusted models and its accuracy in forecasting short-term. According to the method, the exchange rate should fluctuate around 1.87 to 1.92 R\$/US\$ in May, June and July 2012.

1. Introdução

A política cambial, da qual a competitividade da taxa de câmbio é elemento fundamental, representa um ferramental estratégico a ser utilizado por governos visando ao crescimento e ao desenvolvimento pleno de uma sociedade. Além de evitar crises na balança comercial, um câmbio relativamente desvalorizado pode fornecer estímulos para aumentos de investimento e de poupança, impactando o processo de acumulação de capital (KRUGMAN; OBSTEFELD, 2001). Quanto a questões produtivas, um câmbio competitivo pode gerar impactos positivos para o resto da economia, a partir do estímulo à produção de manufaturas para exportação, além de evitar processos de desindustrialização ou agropecuária predatória (BRESSER-PEREIRA, 2006).

Além das demandas de competitividade, o câmbio é notavelmente um instrumento de estabilização econômica. O último plano de estabilização econômica brasileiro, denominado Plano Real, foi introduzido em duas etapas: primeiramente com a criação de

um indexador (a URV) a partir de dezembro de 1993, e somente em julho de 1994 ocorreu a reforma do padrão monetário, com subsequente ancoragem da economia ao câmbio (FAVA; ALVES, 1997). A segunda etapa introduziu o regime de metas para inflação e a mudança do regime cambial, que, no início de 1999, passou a ser flutuante (GREMAUD et al., 2005). A mudança em si trouxe elevada volatilidade do câmbio e contribuiu para aumentar o grau de incerteza no processo de tomada de decisões de investimento nos muitos setores da economia (RAMOS; BRITTO, 2004).

Diversos autores (SOUZA et al., 2006; ALBUQUERQUE, 2005; ROSSI, 1996) apresentam abordagens relativamente semelhantes para a modelagem da taxa de câmbio, tanto empregando modelos da família Box-Jenkins de previsão, quanto vetores auto-regressivos (VAR), ou processos auto-regressivos com médias móveis de memória longa (ARFIMA). Estas abordagens presumem que as séries sejam, no mínimo, estacionárias de 2ª ordem, também pouco ruidosas, ausentes de mudanças estruturais e que seguem uma distribuição normal de probabilidade.

Então, a opção por um modelo adaptativo - que tenha êxito em detectar e em modelar temporalmente um sistema, enfrentando quebras estruturais na série, ausência de estacionariedade, ou comportamentos relativos à distribuição de probabilidade da série - possibilita modelar a dinâmica a cada novo momento, capturando peculiaridades e novos fatos estilizados a todo instante (CHEN; YEH, 1997). A programação genética (PG) (KOZA, 1992) possibilita encontrar tais parâmetros, variáveis e forma funcional que melhor se adequam à série em estudo.

A significativa diferença entre as abordagens clássica e a PG reside na sensibilidade quanto às perturbações em um processo estocástico. Os modelos ajustados pela PG possuíram maior poder de adaptação em sistemas com perturbações, devido à sua flexibilidade funcional no processo de estimação. Seja em séries com comportamentos lineares ou não lineares, a PG tem a capacidade de encontrar uma forma funcional que melhor se ajuste à série analisada. Logo, a PG pode ser encarada como uma família de inúmeros modelos, que buscam se adequar a um dado processo estocástico.

Portanto, estudos que auxiliam na compreensão do comportamento da taxa de câmbio, buscando métodos que se adequam e que exigem menos pressupostos em suas estimação e previsão, são de extrema importância para a formulação de políticas externas, e para a estabilização e a expansão econômicas. O objetivo deste trabalho é aplicar programação genética na modelagem da série de taxa de câmbio do real por dólar americano (R\$/US\$) e realizar comparações dos seus resultados com os modelos Holt-Winters (MONTGOMERY et al., 1990) e da família Box e Jenkins (1976).

2. Metodologia

Este trabalho se baseia no estudo da estrutura serial da taxa de câmbio real (R\$/US\$) por dólar americano (US\$) mensal, do período de janeiro de 2000 a abril de 2012 (BCB, 2012). A primeira seção da metodologia inicia apresentando o conceito de uma série temporal e os modelos Holt-Winters. A segunda seção aborda resumidamente os modelos autoregressivos

integrados de médias móveis (ARIMA) da família Box-Jenkins de previsão. Após a apresentação das técnicas de previsão clássicas que servem como base de comparação, apresenta-se, na terceira seção, a PG multigênica, seus princípios, métodos de convergência e evolução. E, por fim, a plataforma de implementação e as métricas estatísticas de comparação dos modelos.

2.1. Séries temporais, decomposição serial e modelos Holt-Winters

Uma série temporal é uma sequência de observações obtidas em intervalos regulares de tempo, durante um período específico. Ao analisar uma série temporal, primeiramente deve-se modelar o fenômeno estudado para, então, descrever o comportamento estrutural, realizar estimativas e definir relações causais dos possíveis fatores que influenciam a série.

Uma das formas mais simples de analisar uma série temporal é através da decomposição serial, que objetiva mensurar quatro componentes que estão presentes intrinsecamente em uma observação (Y_t). Destes, temos: 1) a tendência, que denota o comportamento mais geral da variável no tempo; 2) o ciclo, que traduz as oscilações de longo prazo na série; sazonalidade, oscilações de curto prazo periódicas, normalmente por origens naturais; e 3) o ruído, que são movimentos irregulares e por causas desconhecidas (MONTGOMERY et al., 1990).

Modelos baseados na decomposição estrutural são os de Holt-Winters, pois são apropriados para dados em que se verifica a ocorrência de tendência linear, e a componente de sazonalidade, além de sua aplicação ser também válida para séries não estacionárias (MORETTIN; TOLÓI, 2004). Existem dois tipos de modelos, o Aditivo e o Multiplicativo, que funcionam semelhantemente à decomposição serial. As equações de estimação e de previsão são dispostas na Tabela 1.

Tabela 1 - Equações dos modelos Holt-Winters aditivo e multiplicativo

| Holt-Winters Aditivo (HWA) | Holt-Winters Multiplicativo (HWM) |
|---|--|
| $L_t = \alpha(Y_t - S_{t-s}) + (1 - \alpha)(L_{t-1} + b_{t-1})$ | $L_t = \alpha\left(\frac{Y_t}{S_{t-s}}\right) + (1 - \alpha)(L_{t-1} + b_{t-1})$ (1) |
| $b_t = \beta(L_t - L_{t-1}) + (1 - \beta)b_{t-1}$ | $b_t = \beta(L_t - L_{t-1}) + (1 - \beta)b_{t-1}$ (2) |
| $S_t = \gamma(Y_t - L_t) + (1 - \gamma)S_{t-s}$ | $S_t = \gamma\left(\frac{Y_t}{L_t}\right) + (1 - \gamma)S_{t-s}$ (3) |
| $F_{t+k} = (L_t - b_{t-k})S_{t-s-k}$ | $F_{t+k} = (L_t - b_{t-k})S_{t-s-k}$ (4) |

Fonte: Morettin e Tolói (2004).

O parâmetro de nível (1) do modelo Holt-Winters funciona como um intercepto, o valor inicial e constante da equação de previsão. A tendência (2) aponta o grau e direção da declividade do modelo e os fatores sazonais (3) modelam os fenômenos com periodicidade regular na série. A equação de previsão (4) relaciona todos os parâmetros estimados para a previsão k passos à frente.

2.2. Box e Jenkins: ARIMA

Os modelos ARIMA formam uma família de modelos propostos por Box e Jenkins (1976). Essa família é adequada para descrever séries não estacionárias, ou seja, séries que não possuem média constante no período analisado. O modelo ARIMA pode ser classificado de diferentes maneiras, conforme os parâmetros AR(p) (autorregressivo de ordem p), MA(q) (média móvel de ordem q) e I(d) (integrador de ordem d), e são definidos:

$$\phi(B)Y_t = \varepsilon_t, \text{ onde } \phi(B) = (1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p) \quad (5)$$

$$Y_t = \theta(B)\varepsilon_t, \text{ onde } \theta(B) = (1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q) \quad (6)$$

$$\phi(B)Y_t = \theta(B)\varepsilon_t \quad (7)$$

$$\phi(B)W_t = \theta(B)\varepsilon_t, \text{ onde } W_t = \nabla^d Y_t \quad (8)$$

$$\Phi(B^S)\phi(B)W_t = \Theta(B^S)\theta(B)\varepsilon_t, \text{ onde:} \quad (9)$$

$$W_t = \nabla^d \nabla_S^P Y_t$$

$$\Phi(B) = (1 - \Phi_S B^S - \Phi_{2S} B^{2S} - \dots - \Phi_{PS} B^{PS})$$

$$\Theta(B) = (1 - \Theta_S B^S - \Theta_{2S} B^{2S} - \dots - \Theta_{PS} B^{PS})$$

e: $t = 1, 2, \dots, T$; p, q, d, P, Q, D, B e $S \in \mathbb{N} < \infty$

os ϕ, θ, Φ e Θ são parâmetros estimados por máxima verossimilhança, enquanto que B simboliza o operador de atraso e ∇ o operador de diferença com d atrasos. O modelo (5) é um AR de ordem p , onde o objetivo é descrever o comportamento da observação no instante t , sendo explicado pelas observações no tempo $t-p$ - ou seja, o modelo faz uma autoregressão com seus valores passados. O (6) é um MA de ordem q . Esse modelo resulta da combinação linear dos ruídos ocorridos nos instantes t passados, visando a prever o comportamento da série nos instantes $t+k$. O modelo (7) é a combinação de um AR (p) com MA (q), formando um ARMA (p, q), combinando linearmente valores e ruídos defasados p e q períodos, respectivamente, na especificação do modelo.

O (8) é propriamente o ARIMA (p, d, q). Esta família é útil para modelar uma série temporal em que haja necessidade de aplicar transformações por meio de diferenças para torná-la estacionária. O número de diferenças (d) necessárias para tornar uma série estacionária é denominada ordem de integração e é representada pela letra I.

Por fim, o (9) é o SARIMA (p, d, q) \times (P, D, Q)^S. O objetivo desse modelo é captar a componente periódica sazonal que se repete a cada período S . Logo, a dependência é, em um modelo baseado em dados mensais, não somente em Y_{t-1} , mas também Y_{t-12} , Y_{t-24} , etc.

A família de modelos Box-Jenkins levam em conta certos pressupostos, que especificam os procedimentos para estimação e inferência dos parâmetros (BOX et al., 1994). Esses pressupostos estão ligados em sua maioria com ε_t , tais como ausência de correlação serial entre as perturbações estocásticas ($\text{Cov}(\varepsilon_t, \varepsilon_{t-k}) = 0$, para qualquer $k \neq 0$), homogeneidade das variâncias dos erros e ε_t seguir uma distribuição normal, com média zero e variância constante e finita $\varepsilon_t \sim N(0, \sigma^2 < \infty)$.

2.2.1. Etapas da metodologia de Box-Jenkins

Esta família de modelos se apoia em uma metodologia consistente, baseada em quatro passos: identificação, estimação, diagnóstico e previsão (MORETTIN; TOLÓI, 2004). Através desses passos é possível especificar corretamente o modelo para o objetivo desejado.

A identificação busca a ordem dos parâmetros do (S)ARIMA $(p,d,q) \times (P,D,Q)^S$. Uma das maneiras de realizar a identificação da ordem p , P , q e Q é através da função de autocorrelação (FAC) e da função de autocorrelação parcial (FACp). A necessidade de diferenciar a série devido à ausência de estacionaridade pode ser atendida subjetivamente, observando a existência de tendência no gráfico plotado da série temporal, ou através do comportamento vagaroso na queda da FAC ao longo das defasagens. Objetivamente, são feitos alguns testes para identificação de raiz unitária, como o Dickey-Fuller Aumentado (SAID; DICKEY, 1984), onde a hipótese alternativa é a estacionaridade da série.

A estimação é a própria construção do modelo, através da estimação dos parâmetros por mínimos quadrados ordinários ou pelo método máxima verossimilhança. Alguns autores indicam que, em pequenas amostras, o método da máxima verossimilhança produz melhores resultados em termo de eficiência dos estimadores (HAMILTON, 1994).

O diagnóstico é baseado na análise dos resíduos do modelo. A ideia principal é que os resíduos devem apresentar ausência de autocorrelação, indicando que a especificação do modelo foi correta. Esta é feita, ora através dos gráficos dos resíduos em relação ao tempo, ora pela FAC dos resíduos, onde se busca a existência de não autocorrelação dos resíduos em alguma defasagem (exceto a defasagem 0). O teste de Ljung-Box (ENDERS, 1995) busca testar a existência significativa de autocorrelação dos resíduos em alguma defasagem k . A inexistência de autocorrelação dos resíduos em nenhuma defasagem implica a ausência de correlação temporal na estrutura dos resíduos do modelo. Por fim, testamos a hipótese dos resíduos seguirem uma distribuição Gaussiana, através do teste de Jarque-Bera (JARQUE; BERA, 1987). A última etapa consiste em realizar a previsão.

2.3. Programação genética multigênica

Programação genética (PG) (KOZA, 1992; LANGDON; POLI, 2002) é um método pertencente à Computação Evolutiva, biologicamente inspirado e que sintetiza programas de computador para desenvolver alguma tarefa. A PG realiza tal tarefa inicializando aleatoriamente uma população de programas de computador (representadas principalmente por estruturas em árvores), e então, aplicando operadores de recombinação (mutação e cruzamento) nos melhores indivíduos dessa população, esta elabora uma nova

população de programas que herdarão configurações similares, em certo nível, às da população anterior. Esse processo se reinicializa até que, na população, haja programas que consigam otimalidade (ou subotimalidade) para a tarefa desejada.

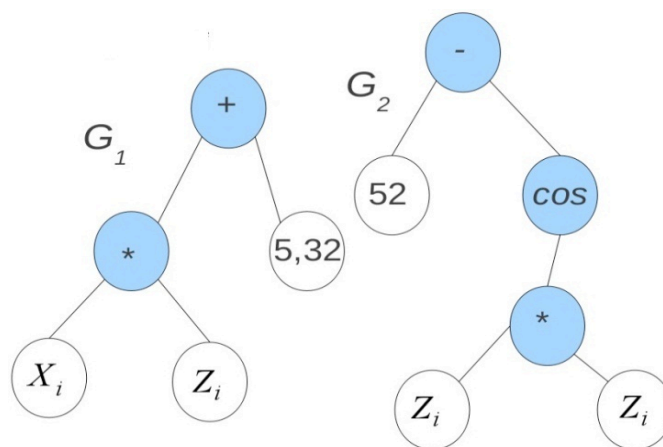
Quando a tarefa é ajustar um modelo matemático empírico, especificamente por PG, tal processo é denominado regressão simbólica. Diferentemente da análise de regressão (cuja forma funcional, no caso, é definida pelo usuário), a PG sintetiza automaticamente ambas, tanto a estrutura do modelo, quanto os seus parâmetros. Recentemente, a regressão simbólica tem obtido resultados satisfatórios academicamente (KORDON, 2006; LANGDON; POLI, 2002) e em aplicações industriais (DIAS, 2005; ALFARO-CID et al., 2009).

Tipicamente, a PG usa uma população de estruturas em árvores, onde em cada uma é codificada uma equação matemática que prediz e descreve um vetor de saídas Y_i , usando uma matriz (X_{ij}) de variáveis explicativas. Baseado nessa estrutura se realiza a regressão simbólica tradicional (gene unitário no caso da multigênica).

Em comparação, a regressão simbólica multigênica, ou PG multigênica (HINCHLIFFE et al., 1996; SEARSON et al., 2007) perfaz uma combinação linear de uma quantidade de n saídas quaisquer de programas, onde cada programa pode ser considerado um "gene" em um indivíduo. A Figura 1 apresenta um exemplo de modelo multigênico.

Figura 1 - Exemplo de árvore de um indivíduo multigênico, onde:

$$Y_i = \beta_0 + \beta_1(X_i + 5,32) + \beta_2(\cos(Z_i^2) - 52)$$



Fonte: Experimentos conduzidos pelos autores.

A estrutura do modelo apresentado na Figura 1 apresenta um indivíduo que possui dois genes (duas árvores). A saída Y_i é resultado da combinação linear de cada uma das árvores a partir dos parâmetros β_0 , β_1 e β_2 . Na prática, o usuário especifica o número máximo de genes (G_{max}) que um indivíduo pode possuir, além de definir a altura máxima da

árvore (D_{max}). Searson et al. (2007) sugerem, após exaustivos testes, que restringir a altura máxima da árvore em até 4 a 5 nós sobrepostos costuma resultar em modelos relativamente compactos.

A população inicial na PG multigênica é elaborada de forma aleatória, criando-se indivíduos com genes entre 1 e o máximo permitido (G_{max}). Após a inicialização, os indivíduos são avaliados conforme quão próximas de modelar a série de dados estão as funções que estão gerando. A partir dessa avaliação, os indivíduos são imersos em processo de seleção, onde se costuma utilizar o método do torneio (POLI et al., 2008). Os operadores de cruzamento e de mutação são possivelmente aplicados aos indivíduos selecionados via torneio.

O cruzamento mais comum é o de um ponto, onde uma determinada parte de uma árvore (gene) é trocado por uma parte de outra árvore de outro indivíduo na população. Essa forma é a mais comum e também utilizada na PG tradicional. Na PG multigênia há outra possibilidade de se realizar o cruzamento entre indivíduos. Seja o i -ésimo gene de um indivíduo indexado por G_i , e considere que esse indivíduo possua três genes [$G_1 G_2 G_3$], que um segundo indivíduo qualquer possua [$G_4 G_5 G_6 G_7$], e que G_{max} é de 5. Dois pontos aleatórios são aleatoriamente criados em cada indivíduo; logo, no primeiro indivíduo [$G_1 < G_2 > G_3$] e no segundo [$G_4 < G_5 G_6 G_7 >$], onde $< \dots >$ representa o ponto duplo de corte. Os genes enclausurados pelos pontos de cruzamento são trocados, e dois novos indivíduos: [$G_1 G_5 G_6 G_7 G_3$] e [$G_4 G_2$] são criados. Nos casos em que um indivíduo possui mais genes que G_{max} , o algoritmo seleciona aleatoriamente um gene e o elimina até que $G_i < G_{max}$.

O efeito da mutação no sistema acarreta aumento na composição da árvore. Ou seja, dado que um G_i de um indivíduo foi selecionado, a operação de mutação amplia as direções, implementando novas variáveis e operações em seu sistema. Também há a possibilidade de reprodução direta, onde simplesmente o indivíduo é copiado integralmente para a próxima população. Após esse processo é verificado se o critério de parada foi atendido. Caso positivo, o algoritmo retorna à última população de soluções; se não, reinicia-se todo o processo até atingir o critério de parada especificado pelo operador (número de gerações, aptidão mínima de interesse etc.).

Portanto, dispendo dessas definições, a saída de um indivíduo é dada pelo modelo:

$$\hat{Y}_i = \beta_0 + \sum_{k=1}^K \beta_k G_{ik} \quad (10)$$

onde \hat{Y}_i é a saída de um indivíduo qualquer da população, β_k são parâmetros lineares e G_i são as saídas dos genes que o indivíduo possui. Os β_k desse método podem ser estimados por mínimos quadrados, levando em conta que o cálculo do erro $\varepsilon_i = Y_i - \hat{Y}_i$ é factível. A função de avaliação desse indivíduo pode ser formulada em termos de erro médio quadrático (MSE):

(11)

$$MSE = \sum_{i=1}^n \frac{(Y_i - \hat{Y}_i)^2}{n}$$

O indivíduo que minimizar o MSE é o que melhor produziu e combinou os G_i . Ressalta-se que outras métricas podem ser usadas, tais como a raiz quadrada do erro médio quadrático (RMSE) e o erro médio absoluto percentual (MAPE). Quando há a ocorrência de empates em termos de avaliação entre indivíduos que são os respectivos melhores da população, aplica-se a pressão lexicográfica parcimoniosa (LUKE; PAINAT, 2002), que define o melhor indivíduo da população em uma determinada geração como sendo o que possuir o menor número de nós.

2.4. Métricas de ajuste e plataforma computacional

Quando há um conjunto de diferentes modelos, é necessário realizar testes e usar critérios para mensurar o ajuste de cada modelo em relação aos dados e selecionar o melhor modelo de previsão. Alguns critérios são definidos a seguir:

$$SMAPE = \sum_{t=1}^T \frac{\frac{|Y_t - \hat{Y}_t|}{|Y_t + \hat{Y}_t|/2}}{T} \quad (12)$$

$$U^2 = \sqrt{\sum_{t=1}^T \frac{(Y_t - \hat{Y}_t)^2}{(Y_t - Y_{t-1})^2}} \quad (13)$$

onde: $t = 1, 2, \dots, T$ e $k = 1, 2, \dots, K$

O SMAPE (erro médio simétrico absoluto percentual) (12) descreve o percentual de erro que o modelo está cometendo nas suas previsões. O uso do SMAPE visa a minimizar alguns erros que o tradicional MAPE (erro médio absoluto percentual) comete quando o valor da previsão é maior ou menor que o observado, ou quando os valores das previsões são próximos de zero (MAKRIDAKIS e HIBON, 2000).

O U^2 -Theil (13) é chamado de coeficiente de desigualdade, pois mensura as discrepâncias cometidas pela previsão um passo à frente. Atua como um parâmetro de avaliação. Quanto mais próximo de zero, melhor o desempenho do modelo (SOUZA et al., 2007). Desse modo, medidas do U^2 -Theil cujo valor é superior a 1 implicam que o modelo de previsão empregado é menos eficiente que o método ingênuo.

As análises estatísticas e de modelagem clássica foram desenvolvidas no pacote estatístico R (R DEVELOPMENT CORE TEAM, 2012) e para a PG multigênica foi utilizada a biblioteca GPTIPS (SEARSON, 2009) implementada em MATLAB (MATLAB, 2010). As configurações da PG são apresentadas na Tabela 2.

Tabela 2 - Principais configurações da PG multigênica

| Característica do Algoritmo | Parâmetro |
|--|--------------|
| Altura máxima da árvore | 5 |
| Máximo de nós por árvore | 100 nós |
| Número limite de gerações | 100 gerações |
| Quantidade limite permitida de genes por indivíduo | 4 genes |
| Indivíduos participando no torneio | 2 indivíduos |
| Probabilidade de cruzamento | 85% |
| Probabilidade de mutação | 10% |
| Probabilidade de reprodução direta | 5% |
| Ramificação gerada pela mutação | 5 nós |
| Taxa de elitismo | 2% |

Fonte: Experimentos conduzidos pelos autores.

Foram omitidos os seis últimos meses da séries, considerando apenas o período de janeiro de 2000 a outubro de 2011 para treinamento dos modelos, subdividida em três partes: conjunto de treinamento (60%), validação (20%) e teste (20%). O período omitido serviu como forma de avaliar as previsões geradas pelos modelos fora da amostra. Os terminais (variáveis) de entrada e funções matemáticas possíveis a serem usadas pela PG foram relacionados na Tabela 3.

Tabela 3 - Operações e variáveis do PG multigênica

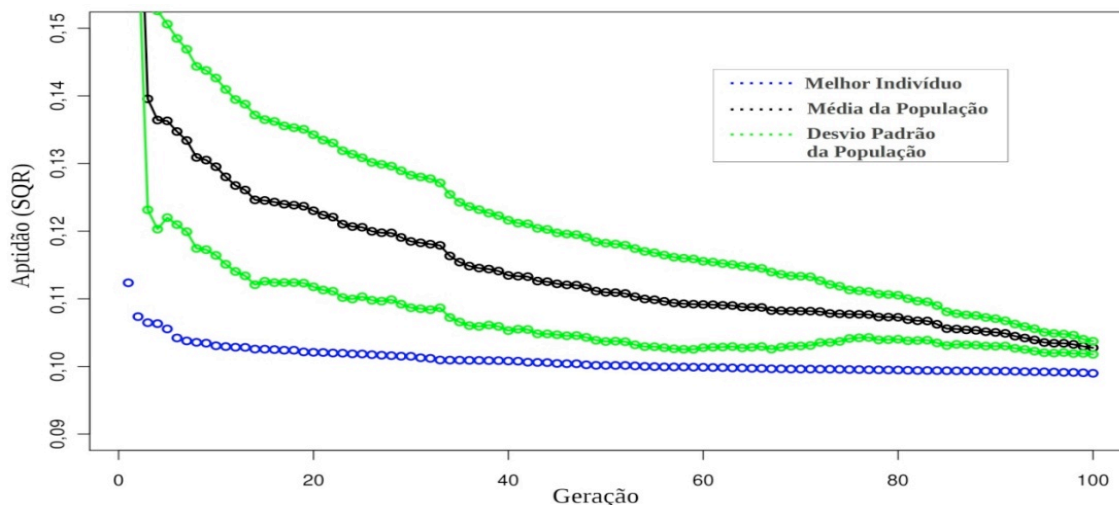
| Terminais | Funções |
|--|---|
| Defasagens mensais (de 1 a 12 meses) da série temporal da taxa de câmbio e constantes no intervalo [-10 10]. | Soma, subtração, produto, seno, cosseno, tangente hiperbólica, logaritmo protegido e exponencial. |

Fonte: Experimentos conduzidos pelos autores.

3. Resultados e Discussões

A Figura 2 apresenta o processo de evolução da população da PG multigênica. O melhor indivíduo obteve um MSE de 0,0998, enquanto que o MSE médio da última geração foi de 0,1052, com desvio padrão de 0,0053. Portanto, pode-se inferir que houve diferença significativa (p -valor < 5%) entre o melhor indivíduo e a média da população; logo, o melhor indivíduo possui características funcionais que o fizeram sobrepôr a maior parte da população.

Figura 2 - Avaliação dos indivíduos durante o processo de evolução da PG.

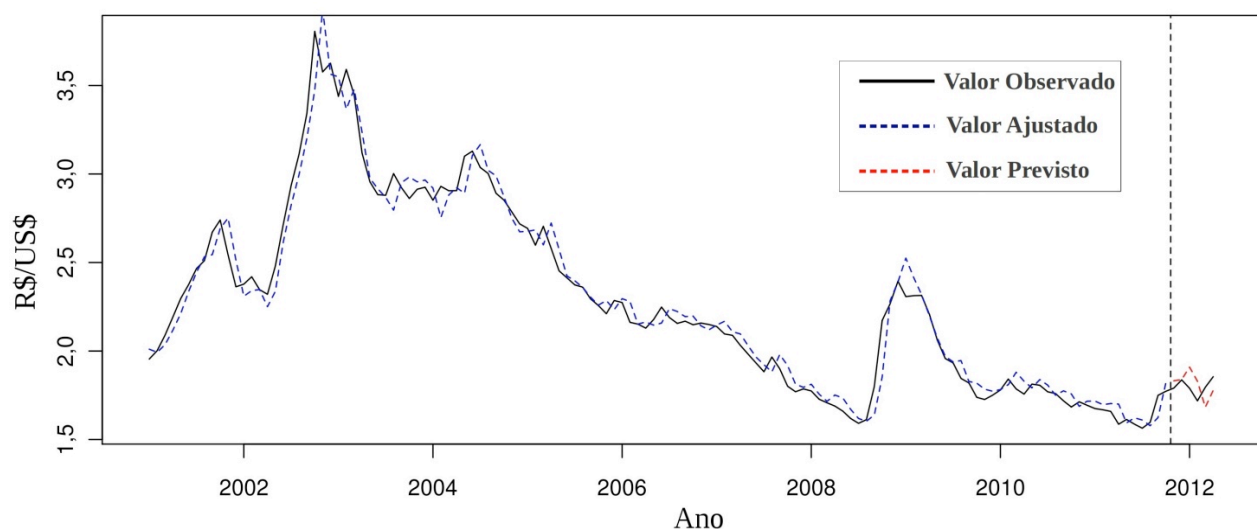


Fonte: Experimentos conduzidos pelos autores.

A função estimada pelo melhor indivíduo em todas as gerações está expressa na equação abaixo:

$$\widehat{Y}_t = 1,14Y_{t-1} - 0,07Y_{t-2} + 0,06Y_{t-4} - 0,06Y_{t-5} - 0,35Y_{t-6} + 0,24Y_{t-8} + 0,05Y_{t-9} + 0,05Y_{t-12} + 0,01 \quad (14)$$

Verifica-se que a função reproduzida pelo melhor indivíduo utiliza três operadores: a soma, a subtração e o produto. Além disso, o modelo relaciona o período sazonal da série (t-12) e a uma autoregressão a sete meses anteriores (t-1, t-4, t-5, t-6, t-8, t-9). A Figura 3 apresenta os valores observados da série da taxa de câmbio, os valores ajustados pelo modelo de PG multigênica, e a previsão seis passos à frente. As Tabelas 4 e 5 apresentam as métricas dos demais modelos ajustados, para efeitos de comparação com a PG multigênica. Os modelos Holt-Winters e ARIMA foram estimados seguindo critérios apresentados na metodologia.

Figura 3 - Valores observados, ajustados e preditos pelo modelo PG multigênica.

Fonte: Experimentos conduzidos pelos autores.

Tabela 4 - Métricas de qualidade do ajuste na fase de treinamento dos modelos estimados

| Modelos | MAPE (%) | SMAPE (%) | U2-Theil |
|----------------|----------|-----------|----------|
| Holt-Winters | 2,99 | 2,98 | 0,81 |
| ARIMA (7,1,0) | 2,58 | 2,59 | 0,88 |
| PG multigênica | 2,29 | 2,31 | 0,71 |

Fonte: Experimentos conduzidos pelos autores.

Tabela 5 - MAPE fora da fase de treinamento (%) dos modelos estimados, dado um horizonte de previsão

| Modelos | 2 passos à frente | 4 passos à frente | 6 passos à frente |
|----------------|-------------------|-------------------|-------------------|
| Holt-Winters | 3,12 | 3,21 | 3,24 |
| ARIMA (7,1,0) | 2,52 | 3,53 | 3,37 |
| PG multigênica | 1,21 | 3,28 | 3,63 |

Fonte: Experimentos conduzidos pelos autores.

Os resultados dispostos apresentam, em sua maioria, a superioridade do modelo GP multigênica frente aos demais métodos. Apesar de pequenas diferenças absolutas, tais resultados devem ser encarados de forma mais relativa. O MAPE (%) e o SMAPE (%) do PG multigênica foram, respectivamente, 12% e 24% melhores do que o ARIMA (7,1,0) e Holt-Winters, respectivamente, e no U²-Theil mostraram vantagens similares perante os demais modelos. O MAPE fora da amostra (%) dois passos à frente apresentou erro percentual

menor do que o Holt-Winters e o ARIMA (7,1,0). Nos demais casos, o erro do PG multigênica aumentou consideravelmente rápido, revelando a sua necessidade de incorporar novos dados para readaptar a função.

Após a verificação dos critérios de qualidade do ajuste do modelo, é necessário verificar se o modelo obteve êxito em captar toda a estrutura de autocorrelação das observações, e, assim, gerar resíduos na modalidade de ruído branco. As Figuras 4a e 4b apresentam a função de autocorrelação, e o teste de Ljung-Box, e os resultados demonstram que o modelo estimado por PG multigênica obteve sucesso, pois em nenhuma defasagem testada o p-valor do teste de Ljung-Box não apresentou significância estatística (p-valor > 5%) de presença de correlação serial nos resíduos.

Figura 4. A - Função de autocorrelação (FAC) dos resíduos do modelo

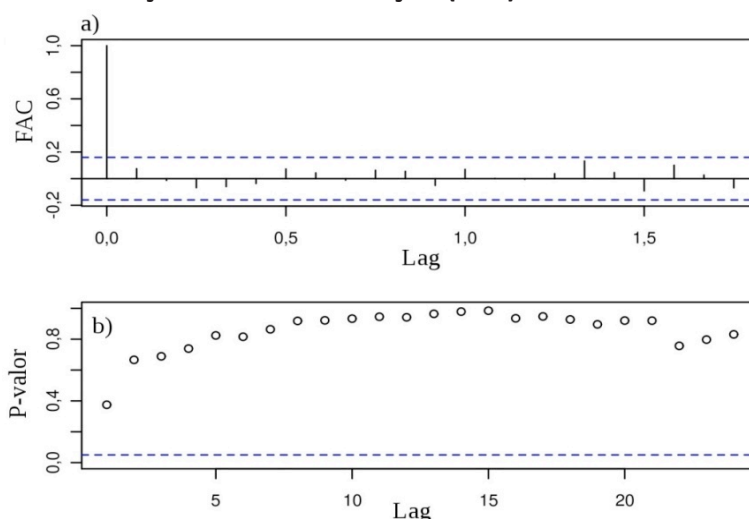


Figura 4.B - Teste de Ljung-Box

Fonte: Experimentos conduzidos pelos autores.

A Tabela 6 apresenta as previsões geradas pela modelagem PG multigênica, incluindo os dados não considerados, para a previsão dos meses de maio, junho e julho da taxa cambial R\$/US\$.

Tabela 6 - Previsões do modelo PG multigênica.

| Horizonte | Previsão |
|--------------|----------|
| Maio – 2012 | 1,87 |
| Junho – 2012 | 1,92 |
| Julho - 2012 | 1,90 |

Fonte: Elaboração própria.

4. Conclusões

A taxa de câmbio R\$/US\$ foi, no período de fevereiro de 1994 a janeiro 1999, administrada pelo sistema de bandas cambiais implementadas pelos governos Itamar Franco e Fernando Henrique Cardoso. A administração fez-se necessária pela vinculação do câmbio à estabilidade econômica e monetária do plano Real nos períodos iniciais. Devido às duras crises internacionais como a asiática (1997) e da Rússia (1998-1999), após janeiro de 1999 o governo abandonou o sistema de âncora cambial e, com a flutuação da taxa de câmbio, foi possível a criação de importantes bases de dados, assim como houve a necessidade de cada vez mais se desenvolverem estudos empíricos para compreender o comportamento da taxa de câmbio brasileira.

Através da programação genética (PG) multigênica foi possível identificar a forma funcional, selecionar as variáveis mais pertinentes e estimar um modelo de previsão, baseado em critérios evolutivos para a série da taxa de câmbio real por dólar (R\$/US\$). Ao final do trabalho, apresentamos alguns resultados que afirmavam a qualidade da previsão e do ajuste fornecida pela abordagem frente aos métodos tradicionais de previsão.

Apesar de promissores os resultados, tais possivelmente seriam mais significativos se houvesse tratamento estatístico mais amplo, com o fornecimento de intervalos de confiança para as previsões geradas pelo modelo. O comportamento evolutivo da PG multigênica possibilita a alteração das características do modelo a cada nova observação; logo, há a possibilidade de que, a todo instante de tempo, haja uma reestimação dos parâmetros e das variáveis do modelo, tornando-o dinâmico e de previsões de curto prazo.

Sugestões para trabalhos futuros residem em propor a inclusão de alguns operadores estocásticos (média, variância) e de técnicas que abordem a inclusão da inspiração quântica para melhorar o processo de busca pelos parâmetros, pelas variáveis e pelos operadores da função que minimiza a soma dos quadrados dos erros.

Referências

- ALBUQUERQUE, C.R. **Taxa de câmbio e inflação no Brasil**: um estudo econométrico. Tese (Doutorado em Economia) - Programa de Pós-Graduação em Economia, Universidade Federal do Rio Grande do Sul, Porto Alegre, 2005.
- ALFARO-CID, E.; ESPARCIA-ALCÁZAR, A.I.; MOYA, P.; FEMENIA-FERRER, B.; SHARMAN, K.; MERELO, J.J. Modeling pheromone dispensers using genetic programming. In: **Lecture Notes in Computer Science**. Berlin: Springer, 2009.
- BCB – BANCO CENTRAL DO BRASIL. Taxa de câmbio R\$/US\$ mensal: período 2000-2012. Brasília, 2011.
- BOX, G. E. P.; JENKINS, G. M. **Time series analysis, forecasting and control**. San Francisco: Holden-Day, 1976.
- BRESSER-PEREIRA, L. C. Exchange rate, fix, float or manage it?. In: VERNENGO, M. **Financial integration or dollarization**: no panacea. Cheltenham: Edward Elgar, 2006.
- CHEN, S.H.; YEH, C.H. Using genetic programming to model volatility in financial time series: the cases of Nikkei 225 and S&P 500. In: JAFEE International Conference on Investments and Derivatives - JIC, 4., 1997, Tokyo. **Proceedings...** Tokyo: JAFEE, 1997.

- DIAS, D. M. **Síntese automática de programas para microcontroladores digitais por programação genética**. Dissertação (Mestrado em Engenharia Elétrica). Pontifícia Universidade Católica do Rio de Janeiro. Rio de Janeiro, 2005.
- ENDERS, W. **Applied econometric time series**. New York: John Wiley & Sons, 1995.
- FAVA, V. L.; ALVES, D. C. O. Indicadores de movimentação econômica, Plano Real e análise de intervenção. **Revista Brasileira de Economia**, v. 51, n.1, p.133-143, 1997.
- GREMAUD, A.; VASCONCELOS, M. A. S.; TONETO Jr., R. **Economia brasileira contemporânea**. Rio de Janeiro: Atlas, 2005.
- HAMILTON, J. **Time series analysis**. Princeton: Princeton University Press, 1994.
- HINCHLIFFE, M.P.; WILLIS, M.J.; HIDDEN, H.; THAM, M.T.; MCKAY, B.; BARTON, G.W. Modelling chemical process systems using a multi-gene genetic programming algorithm. In: **Genetic Programming: Proceedings of the First Annual Conference** (late breaking papers), 56-65. Michigan: The MIT Press, 1996.
- JARQUE, C. M.; BERA, A. K. A test for normality of observations and regression residuals. **International Statistical Review**, v. 55, n. 2, p.163-172, 1987.
- KORDON, A.K. Future trends in soft computing industrial applications. In: IEEE International Congress on Fuzzy Systems, 2006, Vancouver. **Proceedings...** New York: IEEE, 2006. p. 7854-7861.
- KOZA, J.R. **Genetic programming: on the programming of computers by means of natural selection**. Michigan: The MIT Press, 1992.
- KRUGMAN, P. R.; OBSTFELD, M. **Economia internacional: teoria e política**. São Paulo: Makron Books, 2001.
- LANGDON, W. B.; POLI, R. **Foundations of genetic programming**. Berlin: Springer-Verlag, 2002.
- LUKE, S.; PANAIT, L. Lexicographic parsimony pressure. In: LANGDON, W. B.; CANTU-PAZ, E.; MATHIAS, K.; ROY, R.; DAVIS, D.; POLI, R.; BALAKRISHNAN, K.; HONAVAR, V.; RUDOLPH, G.; WEGENER, J.; BULL, L.; POTTER, M.A.; SCHULTZ, A.C.; MILLER, J.F.; BURKE, E.; JONOSKA, N. (Ed.), **GECCO 2002: Proceedings of the Genetic and Evolutionary Computation Conference**, (pp. 829-836). New York: Morgan Kaufmann Publishers, (2002).
- MAKRIDAKIS, S.; HIBON, M. The M3-Competition: results, conclusions and implications. **International Journal of Forecasting**, v.16, n.1, p. 451-476, 2000.
- MATLAB. version 7.10.0 (R2010a). The MathWorks Inc.: Massachusetts, 2010.
- MONTGOMERY, D. C.; JOHNSON, L. A.; GARDINER, J. S. **Forecasting and time series analysis**. New York: McGraw-Hill, 1990.
- MORETTIN, P. A. **Econometria financeira: um curso em séries temporais financeiras**. São Paulo: ABE, 2006.
- MORETTIN, P. A.; TOLÓI, C. M. C. **Análise de séries temporais**. São Paulo: Edgar Blucher, 2004.
- PHILLIPS, P. C. B.; PERRON P. Testing for Unit Roots in Time Series Regression. **Biometrika**, v.75, n.1, p. 335-346, 1988.
- POLI, R.; LANGDON, W.B.; MCPHEE, N.F. **A field guide to genetic programming**. Publicado por <http://lulu.com> e disponível de graça em: <http://www.gp-field-guide.org.uk>, 2008.
- R DEVELOPMENT CORE TEAM. R: **A language and environment for statistical computing**. R Foundation for Statistical Computing: Vienna, 2011. Disponível em: <http://www.R-project.org>.
- RAMOS, L.; BRITTO, M. **O funcionamento do mercado e trabalho metropolitano brasileiro no período 1991-2002: tendências, fatos estilizados e mudanças estruturais**. Rio de Janeiro: IPEA, 2004.
- ROSSI, J. W. Determinação da taxa de câmbio: testes empíricos para o Brasil. **Pesquisa e Planejamento Econômico**. v. 21, n. 2, p. 397-412, 1996.

SAID, S. E.; DICKEY D. Testing for unit roots in autoregressive moving-average models with unknown order. **Biometrika**, v. 71, n. 1, p. 599-607, 1984.

SEARSON, D.P.; WILLIS, M.J.; MONTAGUE, G.A. Co-evolution of non-linear PLS model components. **Journal of Chemometrics**, v. 2, n.1, p. 592-603, 2007.

SEARSON, D. GPTIPS: Genetic programming & symbolic regression for MATLAB. Disponível em: <http://gptips.sourceforge.net>, 2009.

SOUZA, G. P.; SAMOBYL, R. W; MIRANDA, R. G. **Modelos simplificados de previsão empresarial**. Rio de Janeiro: Editora Ciência Moderna, 2007.

SOUZA, R. C.; CAMARGO, M. E. **Análise e previsão de séries temporais: os modelos ARIMA**. Santa Maria: Sedigraf, 1996.

SOUZA, S.R.S.; TABAK, B.M.; CAJUEIRO, D.O. **Investigação da memória de longo prazo da taxa de câmbio no Brasil**. Banco Central do Brasil, Trabalhos para Discussão n. 113, p. 1-36, 2006.