

PROGRAMAÇÃO GENÉTICA MULTIGÊNICA APLICADA À PREVISÃO DA TAXA DE CÂMBIO REAL POR DÓLAR

Adriano Soares Koshiyama

PUC-Rio - Rua Marquês de São Vicente, 225, Gávea - Rio de Janeiro/RJ
adriano@ele.puc-rio.br

Carlos Magno Catharino Olsson Valle

PUC-Rio - Rua Marquês de São Vicente, 225, Gávea - Rio de Janeiro/RJ
cmcat46@gmail.com

Douglas Mota Dias

PUC-Rio - Rua Marquês de São Vicente, 225, Gávea - Rio de Janeiro/RJ
douglasm@ele.puc-rio.br

Resumo: 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 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 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 deve oscilar em torno de 1,87 a 1,92 R\$/US\$ nos meses de maio, junho e julho de 2012.

Palavras-Chave: Programação Genética; Indivíduo Multigênico; Séries Temporais; Taxa de Câmbio.

Abstract: 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.

Keywords: Genetic Programming; Multigenic Individual; Time Series; Exchange Rate.

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 desenvolvimento pleno de uma sociedade. Além de evitar crises na balança comercial, um câmbio relativamente desvalorizado pode fornecer estímulos para aumentos de investimento e poupança, impactando o processo de acumulação de capital (KRUGMAN e 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 e 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 e 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 autorregressivos (VAR), ou processos autorregressivos 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 modelar temporalmente um sistema, enfrentando quebras estruturais na série, ausência de estacionariedade, ou comportamentos relativos à distribuição de probabilidade da série, tal sistema deve encarar como uma mudança de padrões que estavam vigentes e, portanto, há a ocorrência de um novo comportamento a ser modelado (CHEN e 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 exigem menos pressupostos em sua estimação e previsão, são de extrema importância para a formulação de políticas externas, estabilização e expansão econômica. O objetivo deste trabalho é aplicar programação genética na modelagem da série de taxa de câmbio 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\$) 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 autorregressivos 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 4 componentes que estão presentes intrinsecamente em uma observação (Y_t). Destes, temos a tendência, que denota o comportamento mais geral da variável no tempo; ciclo, que traduz as oscilações de longo prazo na série; sazonalidade, oscilações de curto prazo periódicas, normalmente por origens naturais; e o ruído, movimentos irregulares e por causas desconhecidas (MONTGOMERY et al., 1990). Existem duas formas de decomposição serial: aditiva, usada quando as oscilações na série são fracas (séries homocedásticas) e a multiplicativa, mais apropriada a séries com alta volatilidade.

Modelos baseados na decomposição estrutural são os de Holt-Winters, pois são apropriados para dados em que se verificam a ocorrência de tendência linear, além da componente de sazonalidade (MONTGOMERY et al., 1990). Sua aplicação é, também, válida para séries não estacionárias.

Morettin e Tolói (2004) comentam que as vantagens desse modelo são o fácil entendimento, a aplicação não dispendiosa e por ser adequada para séries com padrão de comportamento mais geral. E que as desvantagens principais passam pelas dificuldades de se determinar os valores mais apropriados das constantes de suavização (α, β, γ) e a impossibilidade de se estudar as propriedades estatísticas, tais como média e variância da previsão e, conseqüentemente, a construção de um intervalo de confiança mais apropriado. Existem dois tipos de modelos, o aditivo e o multiplicativo, que funcionam semelhantemente à decomposição serial. As equações de estimação e previsão estão presentes na tabela 1.

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

| Holt-Winters aditivo | Holt-Winters multiplicativo |
|---|--|
| $L_t = \alpha(Y_t - S_{t-s}) + (1 - \alpha)(L_{t-1} + b_{t-1})$ | $L_t = \alpha\left(\frac{Y_t}{S_{t-s}}\right) + (1 - \alpha)(L_{t-1} + b_{t-1})$ (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) |

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 a frente.

2.2. Modelos Box-Jenkins

A técnica de decomposição serial, apesar de elucidar o comportamento estrutural da série temporal, possui limitações, principalmente na dificuldade de tratar o ruído e buscar melhores explicações para sua geração. Para esse fim, o uso de modelos estocásticos se torna indispensável para o estudo conciso e objetivo da série analisada. As classes de modelos mais comumente utilizadas em análise de séries temporais são as dos modelos autorregressivos integrados e de médias móveis - ARIMA (SOUZA e CAMARGO, 1996).

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) (integrado de ordem d), e são definidos:

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

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

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

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

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

$$W_t = \nabla^d \nabla_S^D Y_t;$$

$$\Phi(B) = (1 - \Phi_S B^S - \Phi_2 S B^2 S - \dots - \Phi_P S B^P S)$$

$$\Theta(B) = (1 - \Theta_S B^S - \Theta_2 S B^2 S - \dots - \Theta_Q S B^Q S)$$

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

O modelo (5) é um AR de ordem p , onde o objetivo é descrever o comportamento da observação no instante t , sendo explicada pelas observações no tempo $t - p$, ou seja, o modelo faz uma autorregressão 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.

Estes 3 modelos apresentados são baseados no pressuposto de estacionaridade da série temporal. Essa estacionaridade é tomada de forma flexível, ou de 2a ordem, onde se presume que os 2 primeiros momentos da série (média e variância) se mantém constantes, além da existência de autocorrelação serial de defasagem (*lag*) k .

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)x(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 leva em conta certos pressupostos, que especificam os procedimentos para estimação e inferência dos parâmetros. Esses pressupostos estão

ligados em sua maioria com os resíduos (ε_t), tais como: ausência de correlação serial entre as perturbações estocásticas ($Cov(\varepsilon_t, \varepsilon_{t-k}) = 0$, para qualquer tk), 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 4 passos: identificação, estimação, diagnóstico e previsão (MORETTIN e 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 e 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 e BERA, 1987). A última etapa consiste em realizar a previsão.

2.3. Programação Genética

Programação genética (PG) (KOZA, 1992) é um método de aprendizado computacional, pertencente à Computação Evolutiva, biologicamente inspirado e que evolui (sintetiza) programas de computador para desenvolver alguma tarefa. A PG realiza tal tarefa inicializando aleatoriamente uma população de programas de computador (representadas por estruturas em árvores), e então, aplicando operadores de recombinação (mutação e cruzamento) nos melhores indivíduos dessa população, de forma a elaborar uma nova população de programas que herdarão configurações semelhantes, em certo nível, à 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, através de uma base de dados, ou um sistema, especificamente por PG, tal processo é normalmente conhecido como regressão simbólica. Diferentemente da análise de regressão (que no caso a forma funcional é definida pelo usuário), a PG evolui automaticamente ambas, tanto a estrutura do modelo, quanto os seus parâmetros. Ultimamente, a regressão simbólica tem obtido

resultados satisfatórios academicamente (KORDON, 2006) e em aplicações industriais (ALFARO-CID et al., 2009).

Tipicamente, regressão simbólica é realizada pela utilização da PG para a evolução de uma população de estruturas em árvores, onde em cada uma é codificada uma equação matemática que prediz/descreve um vetor ($nx1$) de saídas y , usando uma matriz (nxm) de entradas correspondentes, onde n é o número de observações da variável resposta e m é o número de variáveis explicativas. Baseado nessa estrutura se realiza a regressão simbólica de gene unitário.

Em comparação, a regressão simbólica multigênica, ou PG multigênica, é uma combinação linear das saídas de uma quantidade i -ésima qualquer de árvores de programas, onde cada árvore pode ser considerada um "gene" em um cromossomo. A figura 1 apresenta um exemplo de modelo multigênico.

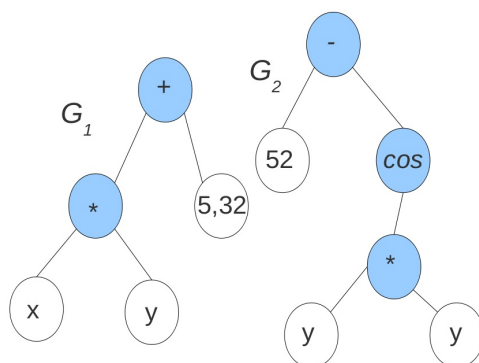


Figura 1: Exemplo de árvore de um indivíduo multigênico, onde: $f(x, y) = \beta_0 + \beta_1 * (xy + 5, 32) + \beta_2 * (\cos(y^2) - 52)$.

Tal estrutura do modelo, apresentado na figura 1, contém termos não-lineares (exemplo: o operador unitário cosseno, além da potência), embora também lineares em respeito aos 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}), e com ambas as ferramentas, demarcar o nível máximo de complexidade que uma população a ser evoluída possuirá. 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.

Para cada modelo os coeficientes lineares (β_i) são estimados a partir da base de dados de treinamento, e então o modelo proposto é validado por uma partição da base de dados de treinamento e, em seguida, confrontado fora da amostra de treinamento, por um conjunto de testes. Nesse ponto serão verificadas as métricas de ajuste e comparados os diferentes modelos usados.

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 os programas que geram, ou seja, o quanto as funções que estão gerando estão próximas de modelar a série de dados. A partir dessa avaliação, os indivíduos são imersos em processo de seleção, onde se costuma utilizar o método do torneio (POLI et al., 2008) para tal seleção. Os operadores de cruzamento e mutação são possivelmente aplicados aos indivíduos selecionados via torneio.

O cruzamento é um meio de aquisição e destruição de material genético durante um processo evolutivo. 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 normalmente também usada em algoritmos

genéticos.

Mas há outra possibilidade de se realizar o crossover entre indivíduos multigênicos, além de somente combinar partições de árvores. Seja o i -ésimo gene de um indivíduo indexado por G_i , e considere que esse indivíduo possua 3 genes, logo $(G_1G_2G_3)$ e um segundo indivíduo qualquer possua $(G_4G_5G_6G_7)$ (considere que G_{max} é 5). Dois pontos aleatórios são criados em cada indivíduo, logo, no primeiro indivíduo $(G_1 < G_2 > G_3)$ e no segundo $(G_4 < G_5G_6G_7 >)$, onde $< \dots >$ representa o ponto duplo de corte. Os genes enclausurados pelos pontos de cruzamento são trocados, e dois novos indivíduos: $(G_1G_5G_6G_7G_3)$ e (G_4G_2) são criados. Nos casos em que há a quebra de restrição de G_{max} , quando 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 em um 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. Após esse processo, de cruzamento e mutação, é verificado se o critério de parada foi atendido. Se sim, o algoritmo retorna o melhor indivíduo; senão, reinicia-se todo o processo até atingir o critério de parada especificado pelo operador (número de gerações, ou aptidão mínima de interesse etc.). O mecanismo multigênico da PG pode ser formulado por um modelo clássico de regressão:

$$Y_i = \beta_0 + \beta_1G_1 + \beta_2G_2 + \dots + \beta_iG_i + \varepsilon_i, \text{ onde } i = 1, 2, \dots, n \quad (10)$$

onde Y_i é a série de dados sendo aproximada por uma função elaborada por PG, $\beta_0, \beta_1, \beta_2, \dots, \beta_i$ são parâmetros lineares estimados por mínimos quadrados ordinários, e G_i são os genes de cada indivíduo na população. Logo, quanto mais genes um indivíduo possui, mais parâmetros ele possui na equação. Logo a PG multigênica combina a regressão linear com a sua habilidade de capturar estruturas não-lineares e o seu comportamento adaptativo ao longo do tempo. Em Hinchliffe et al. (1996) foi demonstrado que a PG multigênica pode ser mais acurada e computacionalmente eficiente do que a PG padrão para aplicação em regressão simbólica. A função de avaliação da PG, soma dos quadrados dos resíduos (SQR), é dada pela equação:

$$\min : SQR = \sum (Y_i - \hat{Y}_i)^2 \quad (11)$$

Onde \hat{Y}_i é o valor gerado pela equação (10), e Y_i é o valor observado na série. Ao buscar a minimização da soma da distância ao quadrado, entre o valor observado e o ajustado por um modelo pertencente a um indivíduo qualquer, tenta-se novas combinações funcionais, até que a condição de parada apresente o indivíduo que melhor minimizou essa função. Assim obtemos uma forma funcional que melhor se adequa às observações em estudo.

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:

$$SMAPE = \frac{\sum_{t=1}^T |Y_t - \hat{Y}_t|}{\frac{(\sum_{t=1}^T (Y_t + \hat{Y}_t))/2}{T}} \quad (11)$$

$$U^2 = \frac{\sqrt{\sum_{t=1}^{T-1} (Y_{t+1} - \hat{Y}_{t+1})^2}}{\sqrt{\sum_{t=1}^{T-1} (Y_{t+1} - Y_t)^2}} \quad (12)$$

O SMAPE (erro médio simétrico absoluto percentual) (11), descreve o percentual de erro do modelo em relação às 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 0 (MAKRIDAKIS e HIBON, 2000). Estas duas métricas são comumente conhecidas como *in-sample*, pois avaliam a qualidade do ajuste do modelo com todas as observações disponíveis.

O U^2 -Theil (11) 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 (MORETTIN e TOLÓI, 2004). Desse modo, medidas do U^2 -Theil cujo valor seja superior a 1 implicam no fato do modelo de previsão empregado ser 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 utilizado a biblioteca PGTIPS (SEARSON, 2009) implementada em MATLAB. As configurações da PG implementada na biblioteca PGTIPS são apresentadas na tabela 2.

Tabela 2: Principais configurações do PG multigênico.

| Característica do Algoritmo | Parâmetro |
|--|----------------|
| Altura máxima da árvore | 8 |
| Máximo de nós por árvore | 100 nós |
| Tamanho da População | 200 indivíduos |
| Número limite de gerações | 100 gerações |
| Quantidade limite permitida de genes por indivíduo | 4 genes |
| Indivíduos participando no torneio | 6 indivíduos |
| Taxa de cruzamento | 85% |
| Probabilidade de mutação | 2% |
| Ramificação gerada pela mutação | 5 nós |
| Taxa de elitismo | 3% |

Foram omitidos 6 meses de dados, ou seja, na modelagem somente foi considerado o período de janeiro de 2000 a outubro de 2011. Este período omitido serviu como uma forma de avaliar as previsões geradas pelos modelos fora da amostra (*out-of-sample*). O restante da série foi subdividida em 3 partes: conjunto de treinamento (60%), validação (20%) e teste (20%). As variáveis (conjunto de terminais) e operações matemáticas possíveis (conjunto de funções) a serem realizadas pela PG estão relacionados na tabela 3.

Tabela 3: Operações e variáveis do PG multigênico.

| Variáveis | Operadores |
|--|--|
| Defasagens mensais (de 1 a 12 meses) da série temporal da taxa de câmbio | Soma, subtração, produto, seno, cosseno, tangente hiperbólica, logaritmo protegido e exponencial |

3. Resultados e Discussões

A figura 2 apresenta o processo de evolução da população da PG multigênica. O melhor indivíduo obteve um SQR de 0,0998, enquanto que a média da mesma geração foi de 0,1052, com desvio padrão de 0,0053. Portanto, pode-se inferir que houve diferença

significativa ($p\text{-valor} < 5\%$) entre o melhor indivíduo e a média da população, logo, o melhor indivíduo possui características funcionais que lhe fizeram se sobrepôr a maior parte da população.

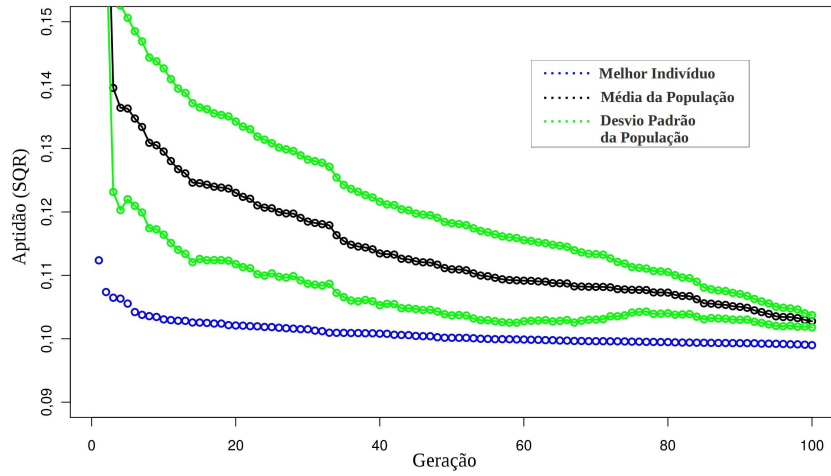


Figura 2: Avaliação dos indivíduos durante o processo de evolução da PG.

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

$$\hat{Y}_t = 0,0008 + 1,1366Y_{t-1} - 0,0665Y_{t-3} + 0,0645Y_{t-4} - 0,0662Y_{t-5} - 0,3549Y_{t-6} + 0,2351Y_{t-8} + 0,0565Y_{t-9} - 0,0533Y_{t-12} \quad (14)$$

verifica-se que a função reproduzida pelo melhor indivíduo utiliza 3 operadores, a soma, a subtração e o produto. Além disso, o modelo relaciona o período sazonal da série (Y_{t-12}) e uma autorregressão a 7 meses anteriores (Y_{t-1} , Y_{t-4} , Y_{t-5} , Y_{t-6} , Y_{t-8} , Y_{t-9}). A figura 3 apresenta os valores observados da série da taxa de câmbio, os valores ajustados pelo modelo de PG multigênico, e a previsão 6 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.

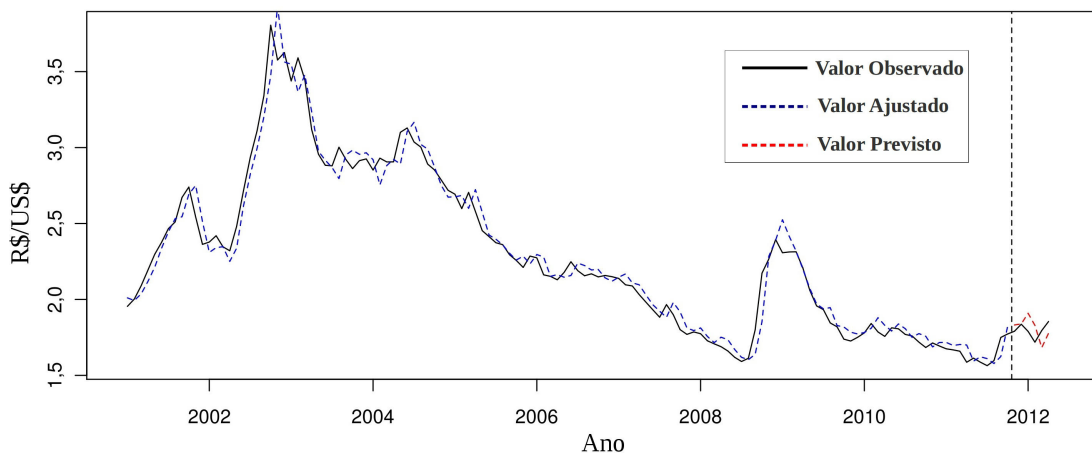


Figura 3: Valores observados, ajustados e preditos pelo modelo PG multigênico.

Os resultados dispostos apresentam, em sua maioria, a superioridade do modelo PG multigênico frente aos demais métodos (tabelas 4 e 5). Apesar de pequenas as diferenças

Tabela 4: Métricas de qualidade do ajuste dos modelos estimados.

| Modelos | MAPE <i>in-sample</i> (%) | SMAPE (%) | U ² -Theil |
|----------------|---------------------------|-----------|-----------------------|
| Holt-Winters | 2,99 | 2,98 | 0,81 |
| ARIMA (7,1,0) | 2,58 | 2,59 | 0,88 |
| PG multigênico | 2,29 | 2,31 | 0,71 |

Tabela 5: MAPE *out-of-sample* (%) 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 multigênico | 1,21 | 3,28 | 3,63 |

absolutas, tais resultados devem ser encarados de forma mais relativa. O MAPE (%) e SMAPE (%) do PG multigênico foi 12% e 24% melhor do que o ARIMA (7,1,0) e Holt-Winters, respectivamente, e no U²-Theil mostrou vantagens similares perante aos demais modelos. O MAPE *out-of-sample* (%) 2 passos-à-frente apresentou erro percentual menor do que o Holt-Winters e ARIMA (7,1,0). Nos demais casos, o erro do PG multigênico aumentou de forma consideravelmente rápida, 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 multigênico obteve sucesso, pois em nenhuma defasagem testada o *p-valor* do teste de Ljung-Box não apresentou significância estatística (*p-valor* > 5%) de autocorrelação dos resíduos.

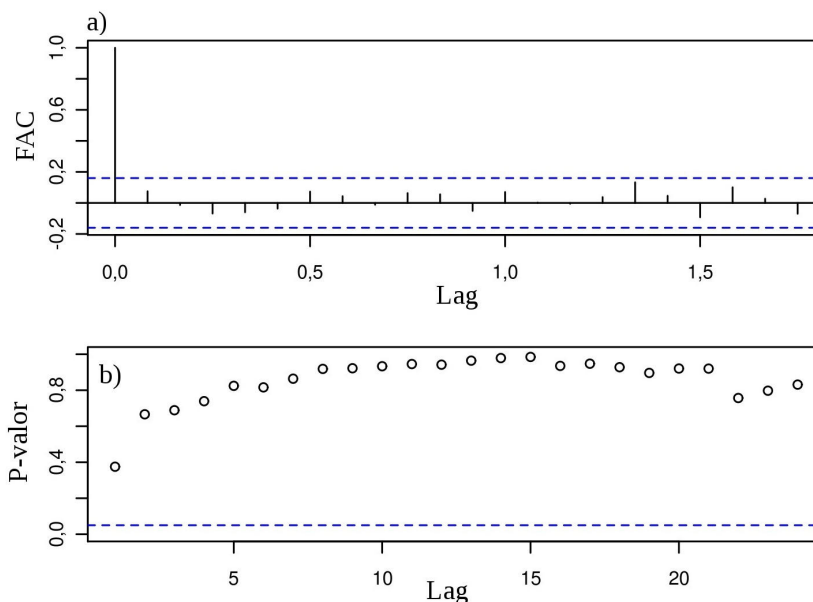


Figura 4: A) Função de autocorrelação (FAC) dos resíduos do modelo. B) Teste de Ljung-Box.

A tabela 6 apresenta as previsões geradas pela modelagem PG multigênica, incluindo os dados não considerados previamente, para a previsão dos meses de maio, junho e julho.

da taxa cambial R\$/US\$.

Tabela 6: Previsões do modelo PG multigênico.

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

4. Conclusões

Devido às duras crises internacionais como a Asiática (1997) e da Rússia (1998-1999), após janeiro de 99 o governo abandona o sistema de âncora cambial e, com a flutuação cambial, foi possível a criação de importantes bases de dados e a necessidade de se desenvolver cada vez mais 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 ajuste fornecida pela abordagem frente aos métodos tradicionais de previsão.

Apesar de promissores, os resultados possivelmente seriam mais significativos se houvesse um 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 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, etc.) e de técnicas para melhorar o processo de busca pelos parâmetros, variáveis e operadores da função que minimiza a soma dos quadrados dos erros.

5. Referências Bibliográficas

ALBUQUERQUE, C.R. *Taxa de câmbio e inflação no Brasil: um estudo econométrico*. Tese de Doutorado. 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. 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 SP 500. In: *The 4th JAFEE International Conference on Investments and Derivatives (JIC'97)*, Royal Park Hotel, Tokyo, 1997.

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*, p. 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. *Proceedings of the 2006 IEEE Congress on Evolutionary Computation*, p. 7854-7861, 2006.
- 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.
- MAKRIDAKIS, S.; HIBON, M. The M3-Competition: results, conclusions and implications. *International Journal of Forecasting*, v.16, n.1, p. 451-476, 2000.
- MONTGOMERY, D. C., JOHNSON, L. A. e GARDINER, J. S. *Forecasting and Time Series Analysis*. New York, McGraw-Hill, 1990.
- MORETTIN, P. A.; TOLÓI, C. M. C. *Análise de Séries Temporais*. São Paulo: Edgar Blucher, 2004.
- POLI, R.; LANGDON, W.B.; McPHEE, N.F. *A field guide to genetic programming*. Publicado: <http://lulu.com> e disponível em: <http://www.PG-field-guide.org.uk>, 2008.
- R DEVELOPMENT CORE TEAM. *R: A language and environment for statistical computing*. R Foundation for Statistical Computing: Vienna, 2012.
- 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, 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.