


# APLICAÇÃO DE SINGULAR SPECTRUM ANALYSIS E DO ARIMA-GARCH PARA PREVISÃO DA TAXA DE CÂMBIO EUR/USD

RAFAEL J. ABREU<sup>1</sup>

 <https://orcid.org/0000-0001-7102-5940>

RAFAEL M. SOUZA<sup>2</sup>

 <https://orcid.org/0000-0001-8798-867X>

JOICE G. OLIVEIRA<sup>3</sup>

 <https://orcid.org/0000-0003-1384-0976>

**Para citar este artigo:** Abreu, R. J., Souza, R. M., & Oliveira, J. G. (2019). Aplicação de *singular spectrum analysis* e do ARIMA-GARCH para previsão da taxa de câmbio EUR/USD. *Revista de Administração Mackenzie*, 20(4). doi:10.1590/1678-6971/eRAMF190146

**Submissão:** 31 ago., 2018. **Aceite:** 17 dez., 2018.

<sup>1</sup> Ipanema Comercial Exportadora e Importadora Ltda., Alfenas, MG, Brasil.

<sup>2</sup> Universidade Federal de Minas Gerais (UFMG), Belo Horizonte, MG, Brasil.

<sup>3</sup> Universidade Federal de Minas Gerais (UFMG), Belo Horizonte, MG, Brasil.



This is an open-access article distributed under the terms of the Creative Commons Attribution License.

This paper may be copied, distributed, displayed, transmitted or adapted if provided, in a clear and explicit way, the name of the journal, the edition, the year and the pages on which the paper was originally published, but not suggesting that RAM endorses paper reuse. This licensing term should be made explicit in cases of reuse or distribution to third parties. It is not allowed the use for commercial purposes.

Este artigo pode ser copiado, distribuído, exibido, transmitido ou adaptado desde que citados, de forma clara e explícita, o nome da revista, a edição, o ano e as páginas nas quais o artigo foi publicado originalmente, mas sem sugerir que a RAM endosse a reutilização do artigo. Esse termo de licenciamento deve ser explicitado para os casos de reutilização ou distribuição para terceiros. Não é permitido o uso para fins comerciais.



## RESUMO

**Objetivo:** O objetivo deste artigo foi modelar a série de minuto das taxas de câmbio do par EUR/USD por meio dos métodos *singular spectrum analysis* (SSA) e ARIMA-GARCH, e avaliar qual gera previsões melhores para um horizonte de cinco minutos.

**Originalidade/valor:** Apesar de o SSA se mostrar uma técnica bem-sucedida em outros ramos da ciência, suas aplicações em finanças ainda são recentes. Além disso, a modelagem da taxa de câmbio é um problema complexo que envolve conceitos e propriedades estatísticas. No entanto, apesar da complexidade, a análise de tal série é de suma importância para diversos agentes que atuam direta ou indiretamente na economia e no mercado financeiro.

**Design/metodologia/abordagem:** Estimaram-se modelos de séries temporais com as técnicas ARIMA-GARCH e SSA, sendo consideradas três amostras do fechamento da cotação *ask* do câmbio: tendências de alta, de baixa e sem tendência dominante.

**Resultados:** As previsões realizadas pelo SSA foram as que mais se aproximaram das observações originais para os três casos. Para as medidas de qualidade, o SSA obteve melhores resultados para as amostras com tendências de alta e de baixa, enquanto, para a amostra sem tendência dominante, os resultados apontaram que a técnica ARIMA-GARCH atingiu resultados mais satisfatórios. Portanto, concluiu-se que as previsões do SSA, no que diz respeito às taxas de câmbio no período em questão, são mais adequadas que aquelas obtidas pelo modelo ARIMA-GARCH, independentemente do movimento do mercado.

## PALAVRAS-CHAVE

Mercado de câmbio. Taxas de câmbio. Dólar. Euro. Previsão de séries temporais.



## 1. INTRODUÇÃO

Nas últimas duas décadas pôde-se observar que, com o intenso avanço da globalização, o fluxo de comércio internacional atingiu patamares antes inimagináveis, nos quais transações envolvendo diferentes pares de moedas se tornaram fundamentais para os diversos agentes da economia. Segundo a World Trade Organization (2014), somente em 2014 as exportações de mercadorias entre seus países-membros totalizaram o montante de 18 trilhões de dólares. Para a instituição, o crescimento do comércio é condição necessária ao desenvolvimento da economia global e ao combate à pobreza. Muitos fatores têm influenciado esse crescimento, como os tratados de cooperação internacional, a qual, possibilitada pelo progresso da tecnologia de informação, resultou em novas abordagens estatísticas para mensurar as transações em termos de valor adicionado.

Nesse contexto, para Atkočiūnas, Mačiulis, Klimavičienė e Kalendienė (2010), taxas de câmbio têm papel central porque permitem que preços de bens e serviços produzidos em diferentes países sejam facilmente comparados, e porque governos, companhias e indivíduos que participam do comércio em escala global estão expostos ao risco das variações cambiais. Portanto, prever as taxas de câmbio é um objeto de pesquisa relevante.

O mercado de *foreign exchange* (FOREX) abrange transações *spot* e *forward*, *swaps* cambiais, opções cambiais e outros derivativos com exposição a mais de uma moeda. De acordo com o Bank for International Settlements (2013), em 2013 o mercado de FOREX movimentou em média 5,3 trilhões de dólares por dia, mais de um trilhão do que em 2010 e mais de dois trilhões do que em 2007, sendo, portanto, o nicho do mercado financeiro mais líquido do mundo. O dólar americano foi a moeda predominante, utilizada em 87% das transações daquele ano, e o euro, por sua vez, foi a segunda moeda mais negociada com um *share* de 33%.

Apesar de o comércio com diferentes pares de moeda não ser um advento recente, foi o *boom* tecnológico que lhe proporcionou tal prodigalidade. A possibilidade de obtenção de dados em tempo real aliada ao aumento da capacidade de processamento deles fez com que, ao longo dos últimos anos, se tornasse cada vez mais comum o uso de modelos estatísticos e matemáticos sofisticados quando do estudo de fenômenos do mercado financeiro. Tais modelos têm sido desenvolvidos e aplicados, por exemplo, para testar se o *capital asset pricing model* (CAPM) ou a *arbitrage pricing theory* (APT) são paradigmas que melhor determinam o retorno sobre ativos de risco; para



explicar as variáveis determinantes na avaliação de títulos de créditos; para obter o *hedge* ótimo e minimizar riscos; para testar se a divulgação da mudança na política de distribuição de dividendos afeta ou não o preço das ações de uma empresa; e para medir e prever a volatilidade e preços de ações, bem como a taxa de câmbio entre pares de moedas (Brooks, 2008).

Por causa de sua importância, vários trabalhos têm sido elaborados com o objetivo de melhor descrever o comportamento das taxas cambiais e de modelá-las, bem como prevê-las ou prever suas volatilidades. Yao e Tan (2000), por exemplo, utilizaram indicadores de análise técnica para alimentar redes neurais e prever a taxa de câmbio entre o dólar americano e o iene japonês, o marco alemão, a libra esterlina, o franco suíço e o dólar australiano. Dacorogna, Müller, Pictet e Vries (2001) demonstraram como o uso de dados de alta frequência do FOREX melhora a eficiência das estimativas de valor em risco. Abraham (2005), por sua vez, comparou métodos de *soft computing* (inteligência artificial, computação evolucionária, lógica probabilística e lógica difusa) e *hard computing* (análise numérica e lógica binária) para previsão das taxas de retorno médio mensais no FOREX. Lai, Yu e Wang (2005) apresentaram uma rede neural com sistema de suporte à decisão (*decision support system* – DSS) para prever mudanças de tendência nas taxas de câmbio. Alamili (2011) comparou as previsões das taxas de câmbio obtidas por *support vector machine* e por redes neurais. Ravi, Lal e Kiran (2012) apresentaram a aplicação dos seguintes métodos de inteligência computacional para previsão de taxas de FOREX: redes neurais com *wavelets*, *splines* de regressão adaptativa multivariada, *support vector machine*, sistema dinâmico evolutivo e programação genética. Chaudhuri e Ghosh (2016) utilizaram redes neurais alimentadas por variáveis macroeconômicas para prever o preço da rupia indiana ante o dólar americano.

Existem condições necessárias para a utilização da técnica de séries temporais em modelagens de taxas de câmbio, o que implica a limitação de seu uso prático, visto que as séries financeiras frequentemente não se restringem aos pressupostos dessa técnica. Recentemente, tem-se admitido que métodos baseados na decomposição em valores singulares (*singular value decomposition* – SVD) são capazes de produzir redução de ruídos em séries não estacionárias (Hassani & Thomakos, 2010).

Dessa forma, neste trabalho propõe-se a seguinte questão de pesquisa: Entre o *singular spectrum analysis* (SSA) e o ARIMA-GARCH, que método de modelagem da taxa de câmbio EUR/USD apresenta melhores resultados quando da utilização de dados de alta frequência para previsão de informações futuras?



Assim, o presente estudo tem como objetivo geral avaliar que modelagem da taxa de câmbio do par EUR/USD gera previsões mais satisfatórias quando da utilização de dados de alta frequência. Para isso, são adotados os seguintes passos: 1. modelagem da série de um minuto das taxas de câmbio do par EUR/USD por meio do método de estimação SSA (não paramétrico) e do ARIMA-GARCH (paramétrico); 2. previsões por meio dos dois modelos<sup>4</sup>; e 3. avaliação de qual modelo é o mais adequado, por meio do *root mean square error* (RMSE), do *mean absolute percentage error* (MAPE) e do *Theil's inequality coefficient* (TIC).

Apesar de o SSA se mostrar uma técnica bem-sucedida em outros ramos da ciência, como meteorologia, oceanografia, biomedicina, tratamento digital de imagens e processamento de sinais digitais, suas aplicações em finanças ainda são recentes. Além disso, a modelagem da taxa de câmbio é um problema complexo que envolve conceitos e propriedades estatísticas, como não estacionariedade, ausência de autocorrelações (*randomwalk*), volatilidades instáveis, distribuições com curtose e assimetria elevada, sazonalidade e ruído. No entanto, apesar da complexidade, a análise de tal série é de suma importância para diversos agentes que atuam direta ou indiretamente na economia e no mercado financeiro.

Espera-se com este trabalho, portanto, aplicar o SSA no âmbito do mercado de FOREX e verificar quais as vantagens ou desvantagens desse método se comparado ao modelo clássico ARIMA-GARCH.

Com intuito de alcançar os objetivos aqui traçados, além desta introdução, o artigo está organizado da seguinte forma: a seção 2 faz uma revisão da literatura; a seção 3 apresenta a metodologia e a base de dados em que, inicialmente, são realizadas a descrição da coleta e a análise dos dados, e, em seguida, indicam-se os aspectos matemáticos de cada modelo; a seção 4, por sua vez, discute os resultados obtidos; e a seção 5, por fim, aponta as considerações finais.

## 2. REVISÃO DE LITERATURA

### 2.1 Métodos clássicos de séries temporais

De acordo com Makridakis e Hibon (1997), depois de refletir sobre o problema das correlações espúrias, o estatístico britânico George Udny Yule

<sup>4</sup> O SSA não é um modelo propriamente dito, mas sim um método de estimação espectral. Para fins didáticos, no entanto, usar-se-ão neste trabalho os termos “modelo SSA” e “modelagem SSA” para referir-se às estimações espectrais realizadas.



introduziu na década de 1920 os modelos autorregressivos (AR). Em seguida, o economista Eugen Slutsky propôs o modelo de médias móveis (*moving average* – MA) ao sugerir a possibilidade da presença de erros de medição em dados econômicos e que componentes cíclicos de tais dados podem ser causados por eventos aleatórios. Por sua vez, em 1938, Herman Wold demonstrou que séries temporais estacionárias podem ser modeladas pela combinação de elementos autorregressivos e de médias móveis, dando origem, assim, ao modelo ARMA.

Para Makridakis e Hibon (1997), a utilização do modelo de Wold se tornou possível em meados da década de 1960 quando computadores (e por consequência a capacidade de realizar os cálculos requeridos e otimizar parâmetros) se popularizaram. Ainda segundo os autores, foi nesse contexto que, em 1970 (edição original, posteriormente publicada em 1976), Box e Jenkins introduziram a metodologia para modelos ARIMA (em que “I” representa o parâmetro de integração necessário para transformar séries não estacionárias em estacionárias) que foi amplamente utilizada no meio acadêmico desde então.

O método proposto por Box e Jenkins começou a ser questionado quando outras técnicas menos sofisticadas se mostraram mais precisas. Groff (1973), por exemplo, realizou previsões de 63 séries de vendas utilizando, para tanto, a metodologia de Box e Jenkins e alisamento exponencial. O autor concluiu que os erros de previsão dos melhores modelos estimados pelo método Box-Jenkins foram superiores àqueles gerados pelo alisamento exponencial. Além disso, Makridakis, Hibone e Moser (1979) aplicaram vários métodos de previsão a 111 séries temporais e concluíram que, para todas, métodos mais simples obtiveram melhores resultados que os modelos ARIMA.

Em 1982, Robert Engle publicou na revista científica *Econometrica* o artigo “Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation”. Em seu trabalho, Engle mostra não ser plausível o pressuposto de variância constante dos modelos ARIMA e introduz o modelo ARCH. Tal modelo foi o predecessor de inúmeros outros que tratam da variância condicional, sendo o GARCH um dos mais amplamente utilizados.

Como o próprio nome sugere, o GARCH, proposto por Bollerslev (1986), é uma generalização do modelo de Heteroscedasticidade Condicional Autorregressiva introduzido por Engle. Os modelos GARCH clássicos expressam a variância condicional como uma função linear do quadrado dos valores passados da série, o que é útil na modelagem de diferentes fenômenos econômicos e financeiros (Bollerslev, 1986). A generalização do modelo





original abrange os principais fatores que caracterizam as séries financeiras, como o comportamento aleatório dos retornos, o agrupamento de volatilidade e assimetrias, e, portanto, obteve sucesso no que diz respeito à sua aplicabilidade nesse ramo da ciência. Segundo Asai, McAleer e Yu (2006), uma ampla gama de modelos multivariados GARCH tem sido desenvolvida, analisada e aplicada quando da caracterização da volatilidade inerente de séries temporais do campo das finanças. Para Bauwens, Laurent e Rombouts (2006), desde a publicação de Engle, os modelos ARIMA foram adaptados a fim de incorporar a variância condicional, e os modelos da família ARCH têm sido comumente utilizados para descrever e prever mudanças na volatilidade de séries financeiras.

## 2.2 Métodos contemporâneos de séries temporais

A não estacionariedade da média ou variância e as quebras estruturais das séries financeiras apresentam-se como desafios à modelagem e previsão dessas séries, visto que tais características não são intrínsecas aos dados, mas provêm de fatores como choques exógenos, mudanças políticas e tecnológicas, mudanças das preferências dos consumidores e da assimetria informacional. Modelos clássicos falham ao lidarem com esses problemas, e, por isso, técnicas novas têm sido desenvolvidas para contorná-los. Entre elas, podem-se citar aplicações do modelo de estado de espaços e o filtro de Kalman, que consideram a variabilidade de coeficientes (como o beta do CAPM) ao longo do tempo; modelos semi e não paramétricos; modelos com cointegração e correção de interceptos, entre outros (Hassani & Thomakos, 2010).

Modelos estruturais ou autorregressivos de séries temporais assumem a normalidade, linearidade e estacionariedade dos dados ou resíduos. Isso limita seu uso prático, visto que as séries financeiras frequentemente não se restringem a esses pressupostos. Além do mais, outro elemento que pode limitar a capacidade preditiva dos modelos é a presença de ruído. Para Hassani e Thomakos (2010), em geral, existem duas abordagens principais no que diz respeito à previsão de séries temporais com ruído: na primeira, ignora-se a presença dele e ajusta-se um modelo que melhor extraia as dinâmicas determinísticas subjacentes; por sua vez, na segunda abordagem, que é mais efetiva, tenta-se decompor a série em subcomponentes, identificar o componente de ruído e extraí-lo, para só então prever novas informações da série filtrada.

Dentro da segunda abordagem, pode-se citar o SSA, que é uma técnica não paramétrica que incorpora elementos clássicos da análise de séries temporais, geometria multivariada, sistemas dinâmicos e processamento de sinais (Hassani, 2007), e que pode ser aplicada em processos estatísticos



arbitrários, lineares ou não lineares, estacionários ou não, gaussianos ou não gaussianos (Hassani, 2010). O SSA possui duas propriedades que justificam seu uso na análise de séries temporais financeiras: não faz pressuposições acerca dos dados e, diferentemente de outros métodos, pode ser aplicada em amostras pequenas (Hassani, 2007).

O SSA tem como característica o fato de ser uma técnica aplicável a séries de tempo com poucas observações. Isso possibilitou seu uso em uma gama de estudos, em que os dados são pouco frequentes ou a coleta em alta frequência destes é inviável. Hassani e Zhigljavsky (2009), por exemplo, demonstraram que a utilização do SSA foi satisfatória para a previsão de séries macroeconômicas tanto pequenas como grandes.

Apesar de o SSA ter se tornado uma ferramenta amplamente utilizada na análise e previsão de séries climáticas (Vautard & Ghil, 1989), meteorológicas (Ghil & Vautard, 1991), hidrológicas (Menezes et al., 2014), biomédicas (Sanei & Hassani, 2015), de tratamento digital de imagens (Rodríguez-Aragón & Zhigljavsky, 2010), de processamento de sinais digitais (Vautard, Yiou, & Ghil, 1992) e também em outros campos do conhecimento, como ciências sociais e físicas (Hassani, 2010), sua aplicação em séries econômicas e financeiras ainda é recente. Hassani e Thomakos (2010) revisaram o desenvolvimento do uso do SSA na análise desse tipo de dados.

Outra característica importante e que deve ser ressaltada é que, embora alguns conceitos probabilísticos e estatísticos sejam empregados nos métodos que se baseiam no SSA, eles não assumem pressupostos de estacionariedade da série analisada e da normalidade dos resíduos (Hassani & Thomakos, 2010).

Segundo Golyandina, Nekrutkin e Zhigljavsky (2001), o SSA pode ser modificado de diversas maneiras, resultando em variações como *Single and Double Centering SSA*, *Toeplitz SSA* e *Sequential SSA*. Uma modificação comumente encontrada nos estudos e nas publicações acerca do SSA é o *multi-channel singular spectrum analysis* (MSSA). Essa última modificação se aplica aos casos em que a série de tempo se constitui de muitas variáveis correlacionadas entre si (Vautard et al., 1992)

Diferentemente do SSA, a aplicação do modelo ARIMA-GARCH depende do tamanho da amostra e de pressupostos como a distribuição dela. Segundo Ng e Lam (2006), diferentes tamanhos de uma série temporal guiam a estimação do modelo para diferentes pontos de soluções ótimas locais. Em seu trabalho, os autores utilizaram amostras de tamanho que variavam de 200 a 3.000 observações do índice Nasdaq de retornos diários e concluíram que a amostra com mil observações apresentou o melhor resultado, apesar de o dia do início das observações interferirem nele.





Em diversas áreas do conhecimento, a técnica SSA tem possibilitado melhores previsões que outros modelos usualmente utilizados para séries temporais (por exemplo, Esquivel, Senna, & Gomes, 2013; Menezes et al., 2014; Royer, Wilhelm, & Patias, 2015). Com base nisso, é considerada a seguinte hipótese de pesquisa:  $H_1$ : O SSA consegue prever com maior precisão a taxa de câmbio EUR/USD se comparada ao modelo ARIMA-GARCH.

## 3. METODOLOGIA E BASE DE DADOS

### 3.1 Base de dados

Para a consecução dos objetivos propostos, realizou-se pesquisa descritiva dividida em duas partes: a primeira consiste na modelagem da taxa de câmbio EUR/USD pelo SSA e pelo modelo ARIMA(p,d,q)-GARCH(r,s). A segunda parte consiste na previsão de dados futuros por meio dos modelos ótimos obtidos na etapa anterior. Os resultados da segunda etapa são utilizados para medir a precisão das estimativas e averiguar qual técnica gera previsões mais satisfatórias.

Para cada conjunto amostral, foram estimados os modelos SSA e ARIMA(p,d,q)-GARCH(r,s). Para tanto, utilizou-se o *software* R, versão 3.3.0 para Windows®, bem como o ambiente de desenvolvimento integrado RStudio®, versão 0.99.484. Para a modelagem da série por meio do SSA, utilizou-se o pacote “Rssa” desenvolvido por Anton Korobeynikov, Alex Shlemov, Konstantin Usevich e Nina Golyandina, e disponibilizado no repositório oficial do *software* estatístico. Para a modelagem da série por meio do ARIMA-GARCH, utilizou-se o pacote “rugarch” desenvolvido por Alexios Ghalanos, também disponibilizado no repositório oficial.

Utilizaram-se três amostras do preço de fechamento *ask* do par EUR/USD com frequência de um minuto. Os dados foram coletados por meio do *software* Marketscope 2.0® da corretora FXCM. A escolha se deu pelo fato de a corretora fornecer cotações por meio do *electronic communication network* (ECN), ou seja, diretamente dos provedores de liquidez (maiores participantes do mercado), e não interferir nelas.

Diferentemente das bolsas de valores, o mercado de FOREX, além de apresentar extrema liquidez, pode ser operado 24 horas por dia a partir das 17 horas (UTC) dos domingos (quando abre a seção asiática, com liquidez proveniente do Japão – sendo Tóquio o polo mais importante –, da China, Austrália, Nova Zelândia, Rússia e outros) até as 17 horas (UTC) das sextas-feiras (quando fecha a seção americana – sendo Nova York o polo mais

importante). Assim sendo, não foi necessário se preocupar com *gaps* (inclusive nos finais de semana).

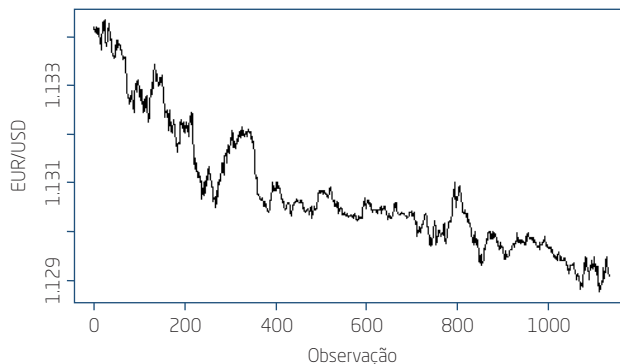
Os dados foram selecionados de modo que se pudessem obter os três movimentos típicos dos mercados financeiros: tendência de alta (*uptrend* ou *bulltrend*), tendência de baixa (*downtrend* ou *beartrend*) e ausência de tendência (*range*). A análise das previsões para ambos os modelos foi feita para cada uma das amostras.

Segundo Murphy (1999), “tendência” é a direção em que o mercado se move, no entanto, esse movimento não se dá em linha reta, mas em uma sequência de zigue-zagues, os quais formam uma série sucessiva de picos e fundos. Assim, para Murphy (1999), pode-se definir a tendência de alta como uma série sucessiva de picos e fundos cada vez mais altos; a tendência de baixa, por sua vez, como uma série sucessiva de picos e fundos cada vez mais baixos; e a ausência de tendência por uma série de picos e fundos horizontais.

A Figura 3.1.1 mostra a evolução do fechamento da taxa de câmbio minuto a minuto da primeira amostra. O intervalo de tempo considerado foi o seguinte: de 23 a 24 de agosto de 2016, das 8h02 (UTC) às 3h02 (UTC). As figuras 3.1.2 e 3.1.3 apresentam, respectivamente, as amostras 2 e 3 utilizadas para estimação dos modelos em tendência de alta (das 15h21 UTC do dia 17 de agosto de 2016 às 16h05 UTC do dia 18 de agosto de 2016) e em *range* (das 9h41 UTC do dia 16 agosto de 2016 às 13h54 UTC do dia 17 de agosto de 2016).

(Figura 3.1.1)

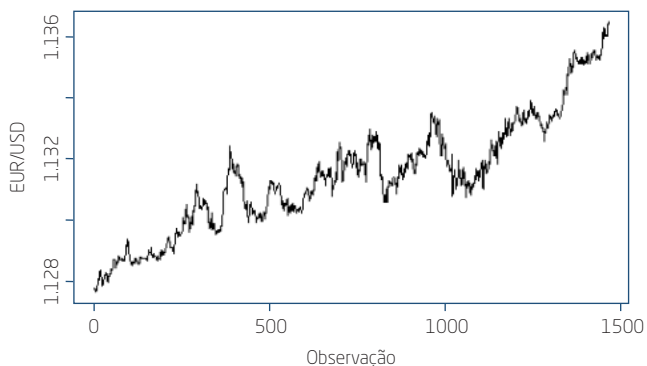
### AMOSTRA 1 - EVOLUÇÃO DO FECHAMENTO DA TAXA DE CÂMBIO DO PAR EUR/USD EM TENDÊNCIA DE BAIXA



Fonte: Elaborada pelos autores com dados obtidos pelo Marketscope 2.0®.

(Figura 3.1.2)

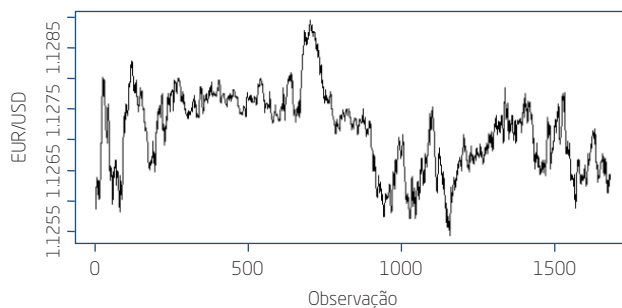
### AMOSTRA 2 - EVOLUÇÃO DO FECHAMENTO DA TAXA DE CâMBIO DO PAR EUR/USD EM TENDÊNCIA DE ALTA



Fonte: Elaborada pelos autores com dados obtidos pelo Marketscope 2.0®.

(Figura 3.1.3)

### AMOSTRA 3 - EVOLUÇÃO DO FECHAMENTO DA TAXA DE CâMBIO DO PAR EUR/USD SEM TENDÊNCIA DOMINANTE



Fonte: Elaborada pelos autores com dados obtidos pelo Marketscope 2.0®.

Deve-se ressaltar, porém, que, diferentemente do mercado acionário, em sua maior parte e principalmente no que diz respeito ao par EUR/USD, a frequência e os impactos das notícias no FOREX se dão em escala global e não local. Tendo em vista esse aspecto do mercado das taxas de câmbio, foram retirados da amostra 34 *outliers* decorridos da divulgação de notícias e/ou relatórios econômicos (*GDP, personal spending, core PCE price index, ISM manufacturing, ADP employment survey, FOMC – fed funds rate, sovereign debt to be rated, PMI*, dentre outros).



Como no FOREX a taxa de câmbio também pode ser entendida como preço, uma vez que, ao se transacionar um par de moedas, o agente vende uma moeda ao mesmo tempo que compra outra, o critério para definir *outlier* foi:

$$ABS(r_t * 100.000) > 50 \rightarrow x_t = outlier \quad (1)$$

em que  $r_t$  é dado por:

$$r_t = \ln \left( \frac{x_t}{x_{t-1}} \right) \quad (2)$$

em que  $x_t$  é a observação  $x$  no instante de tempo  $t$ . O módulo do retorno foi multiplicado por 100.000, visto que a unidade padrão no FOREX (denominada *pip*) é 0,0001, e as diferenças de cotação entre momentos subsequentes da amostra em geral são da ordem de 0,00001.

### 3.2 Modelo ARIMA-GARCH

Conforme Makridakis e Hibon (1997), uma série ARIMA estacionária  $x_t$  pode ser modelada como combinação de seus valores passados, bem como de erros passados:

$$x_t = \varphi_1 x_{t-1} + \varphi_2 x_{t-2} + \dots + \varphi_p x_{t-p} + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \dots - \theta_q \varepsilon_{t-q} \quad (3)$$

$$x_t = \sum_{i=1}^p \varphi_i x_{t-i} - \sum_{i=1}^q \theta_i \varepsilon_{t-i} + \varepsilon_t \quad (4)$$

em que  $p$  é o número de termos AR e  $q$ , o número de termos MA.

O resumo da definição matemática do modelo GARCH apresentada por Bollerslev (1986) é mostrado a seguir. Seja  $\varepsilon_t$  um processo estocástico discreto e  $\psi_t$  o conjunto de todas as informações através do tempo  $t$ . O processo GARCH ( $r, s$ ) é, então, dado por:

$$\varepsilon_t \mid \psi_t \sim N(0, h_t) \quad (5)$$

$$h_t = a_0 + \sum_{i=1}^s a_i \varepsilon_{t-i}^2 + \sum_{i=1}^r \beta_i h_{t-i} = a_0 + A(L) \varepsilon_t^2 + B(L) h_t \quad (6)$$

em que  $p \geq 0$ ;  $q > 0$ ;  $a_0 > 0$ ;  $a_i \geq 0$ ,  $i = 1, \dots, q$ ; e  $\beta_i \geq 0$ ,  $i = 1, \dots, p$ .



A especificação de um modelo ARIMA-GARCH pode, então, ser representada por:

$$\sum_{i=1}^p \varphi_i x_{t-i} - \sum_{i=1}^q \theta_i \varepsilon_{t-i} + \sum_{i=1}^s a_i \varepsilon_{t-i}^2 + \sum_{i=1}^r \beta_i h_{t-i} + \varepsilon_t + a_0 \quad (7)$$

### 3.3 *Singular spectrum analysis (SSA)*

O surgimento do SSA usualmente é associado à publicação dos artigos de Broomhead e King (1986) e Broomhead, Jones, King e Pike (1987). A partir daí muitos trabalhos que tratam dos aspectos metodológicos e das aplicações do SSA foram difundidos: Vautard et al. (1992), Allen e Smith (1996), Danilov e Zhigljavsky (1997) e Golyandina et al. (2001, p. 14).

O método mais básico do SSA consiste na transformação de uma série temporal de valores reais não nulos em uma soma de séries, de modo que cada componente dessa soma possa ser reconhecido como tendência, sazonalidade ou ruído (Hassani, 2007). Para tanto, o algoritmo dessa técnica resume-se em duas etapas: decomposição e reconstrução. A primeira etapa é feita pela transformação dos dados em uma matriz trajetória e sua SVD. A segunda etapa, por sua vez, trata do agrupamento de elementos da matriz trajetória decomposta para formação de subgrupos a partir dos quais a série de tempo será novamente obtida (Hassani, 2010).

Na decomposição, Elsner e Tsonis (1996) demonstraram a possibilidade de se aplicar uma análise estatística multivariada a uma única amostra de série temporal. Segundo os autores, os registros de um sistema dinâmico é o resultado da interação de todas as variáveis, e, portanto, uma única amostra deve conter informações sobre a dinâmica das principais variáveis envolvidas na evolução do sistema. De acordo com Elsner e Tsonis (1996), assume-se que tais variáveis satisfazem um conjunto de  $p$  equações diferenciais de primeira ordem:

$$\begin{aligned} \dot{x}_1 &= f_1(x_1, x_2, \dots, x_p), \\ \dot{x}_2 &= f_2(x_1, x_2, \dots, x_p), \\ &\vdots \\ \dot{x}_p &= f_p(x_1, x_2, \dots, x_p), \end{aligned} \quad (8)$$



em que  $\dot{x}_1$  indica a primeira derivada da variável  $x_1$  em relação ao tempo  $\left(\frac{dx}{dt}\right)$  e assim por diante. Por meio de sucessivas diferenciações, esse sistema pode ser transformado em uma única equação diferencial de ordem  $p$  que representa todo o sistema sem nenhuma perda de informação:

$$x_1^{(p)} = f(x_1, \dot{x}_1, \ddot{x}_1, \dots, x_1^{p-1}), \quad (9)$$

De maneira análoga, é possível fazer a transformação inversa de uma série de tempo discreta e seus sucessivos *lags*. Seja uma série de tempo  $x_t$ , em que  $t = 1, 2, \dots, n$ :

$$x_1, x_2, \dots, x_t \quad (10)$$

E seja  $m$  o número de *lags*, tal que  $2 < m < t$ . Então se pode obter da série original  $n - m + 1$  subconjuntos de tamanho  $m$ , e com elementos subsequentes

$$\begin{aligned} v_1 &= (x_1, x_2, \dots, x_m) \\ v_2 &= (x_2, x_3, \dots, x_{m+1}) \\ &\vdots \\ v_{n-m+1} &= (x_{n-m+1}, x_{n-m+2}, \dots, x_n) \end{aligned} \quad (11)$$

que então são escritos na forma vetorial para obtenção da matriz trajetória  $X$ :

$$X = \begin{bmatrix} v_1 \\ \vdots \\ v_{n-m+1} \end{bmatrix} = \begin{bmatrix} x_1, x_2, \dots, x_m \\ \vdots \\ x_{n-m+1}, x_{n-m+2}, \dots, x_n \end{bmatrix} \quad (12)$$

Portanto, o processo de decomposição se dá pelo mapeamento da série unidimensional  $x_T = (x_1, \dots, x_t)$  em uma matriz  $X_{(t-m+1) \times l}$ .

A matriz trajetória, então, será composta por  $X_1, \dots, X_{n-m+1}$  com  $X_i = (y_i, \dots, y_{i+m-1}) \in R^L$ ;  $m$  representa o tamanho da janela de decomposição (*lags*) e é o único parâmetro a ser definido. A matriz trajetória resultante desse passo é uma matriz de Hankel. Conforme Pan, Rami e Wang (2002), a matriz de Hankel é dada por  $a_{i,j} \in X \rightarrow a_{i,j} = a_{i-1,j+1}$  e pode ser representada da seguinte maneira:



$$\begin{bmatrix} h_0 & h_1 \dots & h_{n-1} \\ h_1 & \ddots & h_n \\ \vdots & \ddots & \vdots \\ h_{n-1} & h_n \dots & h_{2n-2} \end{bmatrix} \quad (13)$$

A matriz trajetória  $X$  simétrica é então transformada em uma soma de matrizes elementares.

Seja  $S = XX^T$ , e  $\lambda_1 \geq \dots \geq \lambda_m$  os autovalores de  $S$  ordenados em ordem decrescentes de magnitude, e  $U_1, \dots, U_m$  o sistema ortogonal de seus autovetores correspondentes aos autovalores. Seja  $d = \text{posto}(X) = \max(i \mid \lambda_i > 0)$ .

Então, se  $V_i = \frac{X^T U_i}{\sqrt{\lambda_i}}$  ( $i = 1, \dots, d$ ), a SVD da matriz trajetória pode ser escrita como  $X = X_1 + \dots + X_d$  em que  $X_i = \sqrt{\lambda_i} U_i V_i^T$ , denominado autotripla.

Já na reconstrução, conforme Golyandina et al. (2001), uma vez que se obteve a matriz trajetória  $X = X_1, \dots, X_{n-m+1}$ , é possível agrupar os componentes dela em subconjuntos e transformar cada matriz resultante desse agrupamento em uma nova série de tempo de tamanho  $m$ . A demonstração matemática desse passo está além do escopo deste trabalho e pode ser encontrada nos trabalhos dos autores citados.

Hassani (2007) propôs a observação de informações suplementares como método para selecionar os componentes da matriz trajetória que serão agrupados e transformados em séries distintas de tendência, sazonalidade e ruído. Para o autor, tais informações são como uma ponte entre a decomposição e a reconstrução; a ideia é dar praticidade ao processo de identificação das autotriplas que constituem cada componente. São consideradas “informações suplementares”, por exemplo, os autovalores e autovetores da matriz  $S$  e a matriz de correlação ponderada. Hassani (2007) mostra que quebras ou saltos entre os autovalores indicam diferentes componentes da série temporal e que, como regra, o ruído produz uma sequência de autovalores que decresce lentamente. Essas características podem ser observadas na figura de autovalores da matriz trajetória  $S$  obtida da série original. Além da figura gerada pelos autovalores, pode-se analisar também a figura gerada pelos elementos de um determinado autovetor e verificar se o autovetor em questão é componente da autotripla que gera de tendência, sazonalidade ou ruído. Para averiguar a separabilidade das autotriplas, ou seja, confirmar se os agrupamentos formados não apresentem correlação entre si, utiliza-se a matriz de correlação ponderada.



### 3.4 Medidas de qualidade

Para testar a capacidade preditiva dos modelos, foram realizadas as seguintes medições:

- *Mean Absolute Percentage Error:*

$$MAPE = \frac{100}{T} \sum_{i=1}^T \left| \frac{x_i - x'_i}{x_i} \right| \quad (14)$$

- *Root Mean Square Error:*

$$RMSE = \sqrt{\frac{1}{T} \sum_{i=1}^T |x'_i - x_i|} \quad (15)$$

- *Theil's Inequality Coefficient:*

$$TIC = \frac{RMSE}{\sqrt{\frac{1}{n} \sum_{i=1}^n x'_i + \frac{1}{n} \sum_{i=1}^n x_i}} \quad (16)$$

em que  $x_i$  é a  $i$ -ésima observação;  $x'_i$  é o  $i$ -ésimo previsão; e  $n$  é o horizonte de previsão. Para todas as medidas, vale a seguinte regra: quanto mais próximo de zero, melhor o ajuste dos modelos.



## 4. RESULTADOS

As modelagens SSA e ARIMA(p,d,q)-GARCH(r,s) foram realizadas para cada uma das três amostras: 1. tendência de baixa, constituída de 1.140 observações; 2. tendência de alta, constituída de 1.471 observações; e 3. ausência de tendência, com 1.686 observações.

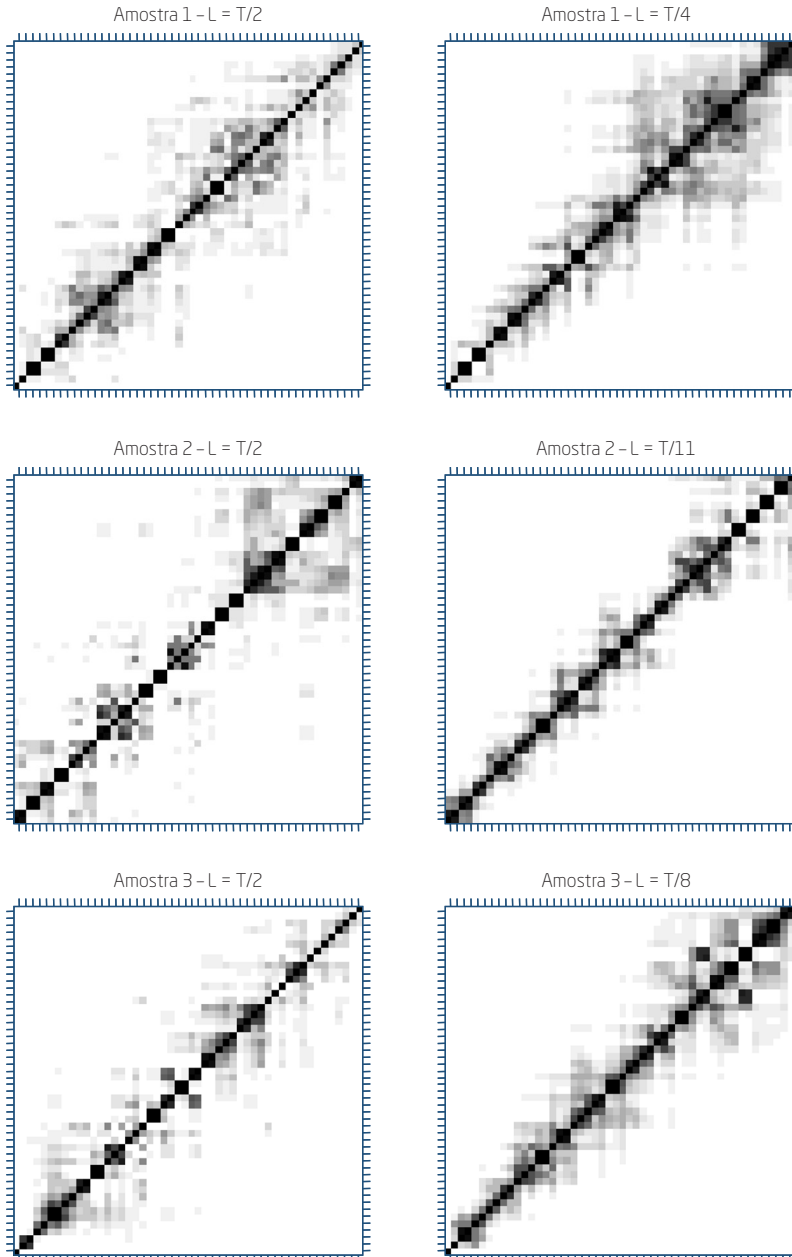
Para realizar as modelagens, só não foram utilizados os cinco últimos dados de cada uma das amostras. Ou seja, a definição dos melhores modelos foi feita de acordo com os critérios próprios de cada técnica em separado e utilizando 1.135 dados da série de tendência de baixa, 1.466 da tendência de alta e 1.681 da amostra com ausência de tendência. No final, foram realizadas previsões fora da amostra para cinco minutos à frente usando os dois modelos ajustados para cada uma das três amostras e comparadas entre si.

### 4.1 Modelagem pelo SSA

Segundo Elsner e Tsonis (1996), a janela de decomposição (mencionada na seção 3.3) é o único parâmetro a ser definido e deve ser um quarto do número de observações. No entanto, isso não é consenso na literatura. Para Golyandina et al. (2001), por exemplo, a janela de decomposição depende de propriedades intrínsecas de cada série a ser modelada. Para Hassani (2007), é possível que  $2 \leq L \leq T$  em que  $L$  é a janela de decomposição, e  $T$ , o número de observações, mas o autor sugere que  $\max\{L\} = \frac{T}{2}$ .

(Figura 4.1.1)

**MATRIZ DE CORRELAÇÃO PONDERADA PARA DIFERENTES  
JANELAS DE DECOMPOSIÇÃO DAS AMOSTRAS**



Fonte: Elaborada pelos autores a partir do pacote RSSA do *software* estatístico R.

O parâmetro em questão exerce grande influência na “separabilidade” da série. Conforme Golyandina et al. (2001), o principal propósito do SSA é a decomposição da série original em uma soma de séries, de modo que cada componente nessa soma possa ser identificado como tendência, componente periódico ou ruído. Para os autores, a decomposição só será bem-sucedida se os componentes aditivos forem separáveis uns dos outros.

A fim de se obter a melhor janela, neste trabalho foram realizadas diversas decomposições pelo método de tentativa e erro, partindo-se do valor máximo sugerido por Hassani (2007),  $L = \frac{T}{2}$ . A “separabilidade” de cada decomposição foi averiguada por meio da matriz de correlação ponderada dos vetores da matriz trajetória. A Figura 4.1.1 mostra algumas das matrizes de correlação ponderada resultantes das diferentes janelas de decomposição das amostras, e a Figura 4.1.2 apresenta os parâmetros ótimos encontrados para cada conjunto amostral. Tais parâmetros foram encontrados pelo método de tentativa e erro.

### (Figura 4.1.2)

#### TAMANHO DA JANELA DE DECOMPOSIÇÃO PARA CADA CONJUNTO AMOSTRAL

Amostra*	T	L
1 - Downtrend	1.135	ROUND/(T/4) = 284
2 - Uptrend	1.466	ROUND/(T/11) = 133
3 - Range	1.681	ROUND/(T/11) = 841

\* Não inclui o horizonte de previsão.

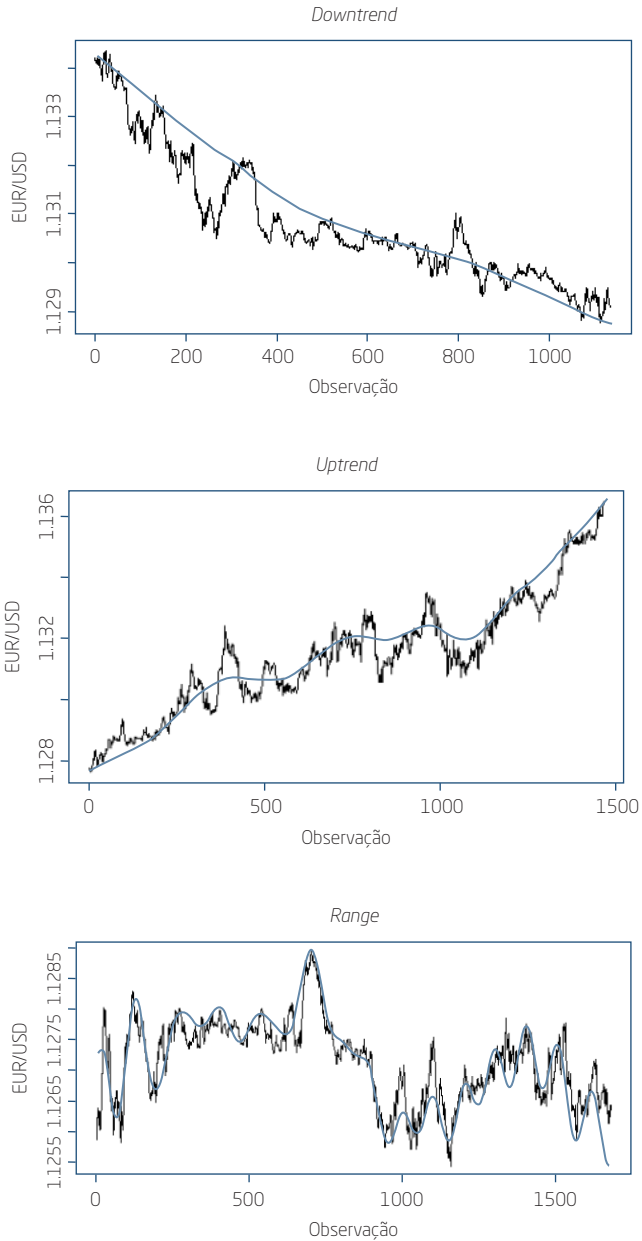
Fonte: Elaborada pelos autores a partir do pacote RSSA do *software* estatístico R.

Uma vez encontrados os parâmetros ótimos das janelas de decomposição das séries, prosseguiu-se para a análise dos autovetores de cada conjunto amostral. A Figura 4.1.3 mostra, em azul, a série reconstruída utilizando apenas o primeiro autovetor e em preto a série original de cada amostra. Para as três amostras, o primeiro autovetor é o mais significativo.

Quando se analisaram em conjunto a matriz de correlações ponderadas e os demais autovetores, adicionaram-se outros componentes que apresentaram características periódicas. Os demais autovetores foram considerados componentes de ruído e excluídos da reconstrução utilizada para realizar as previsões. A Figura 4.1.4 mostra a reconstrução dos componentes de ruído de cada amostra.

(Figura 4.1.3)

## RECONSTRUÇÃO DO CONJUNTO AMOSTRAL PELO AUTOVETOR 1

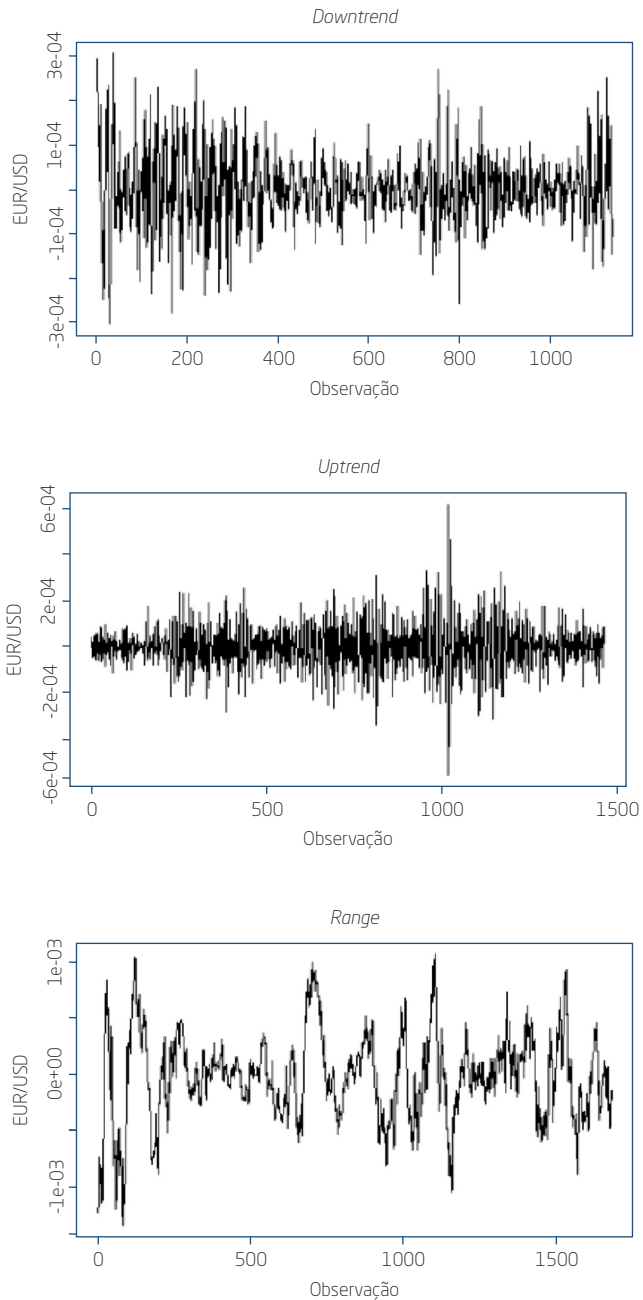


Fonte: Elaborada pelos autores a partir do pacote RSSA do *software* estatístico R.



(Figura 4.1.4)

## RECONSTRUÇÃO DOS COMPONENTES DE RUÍDO DE CADA AMOSTRA



Fonte: Elaborada pelos autores a partir do pacote RSSA do *software* estatístico R.

## 4.2 Modelagem ARIMA-GARCH

Para a modelagem das séries pelo ARIMA(p,d,q)-GARCH(r,s), utilizou-se o método sugerido por Box e Jenkins (1976). Para Makridakis e Hibon (1997), o método pode ser resumido em: 1. verificar se a série é estacionária ou não e, caso não seja, transformá-la em série estacionária em sua média e variância; 2. determinar os parâmetros  $p$  e  $q$  por meio dos coeficientes de autocorrelação e autocorrelação parcial; 3. checar se o modelo é adequado.

A fim de verificar se as séries eram estacionárias, foram utilizados os testes de Dickey-Fuller aumentado e Phillips-Perron. A Figura 4.2.1 apresenta os resultados dos testes para cada série.

Conforme os resultados da Figura 4.2.1, as séries das amostras 1 e 2 não apresentaram estacionariedade. Ambas foram diferenciadas, obtendo-se então  $I(1)$ , e submetidas novamente ao teste. A Figura 4.2.2 mostra os novos resultados.

(Figura 4.2.1)

### TESTES DICKEY-FULLER AUMENTADO E PHILLIPS-PERRON PARA AS TRÊS AMOSTRAS

	Amostra 1 ( <i>downtrend</i> )	Amostra 2 ( <i>uptrend</i> )	Amostra 3 ( <i>range</i> )
Teste ADF			
Nº de observações	1.135	1.466	1.681
Estatística de Dickey-Fuller	-3,45	-2,54	-4,33
Lagorder	10	11,00	11,00
p-valor	0,05	0,35	< 0,01
Hipótese alternativa: estacionariedade			
Teste PP			
Nº de observações	1.135	1.466	1.681
Z (alfa)	-19,23	-21,14	-35,04
Lagorder	7,00	7,00	8,00
p-valor	0,08	0,06	< 0,01
Hipótese alternativa: estacionariedade			

Fonte: Elaborada pelos autores a partir do pacote RSSA do *software* estatístico R.

(Figura 4.2.2)

**TESTES DICKEY-FULLER AUMENTADO E PHILLIPS-PERRON  
PARA AS SÉRIES DIFERENCIADAS**

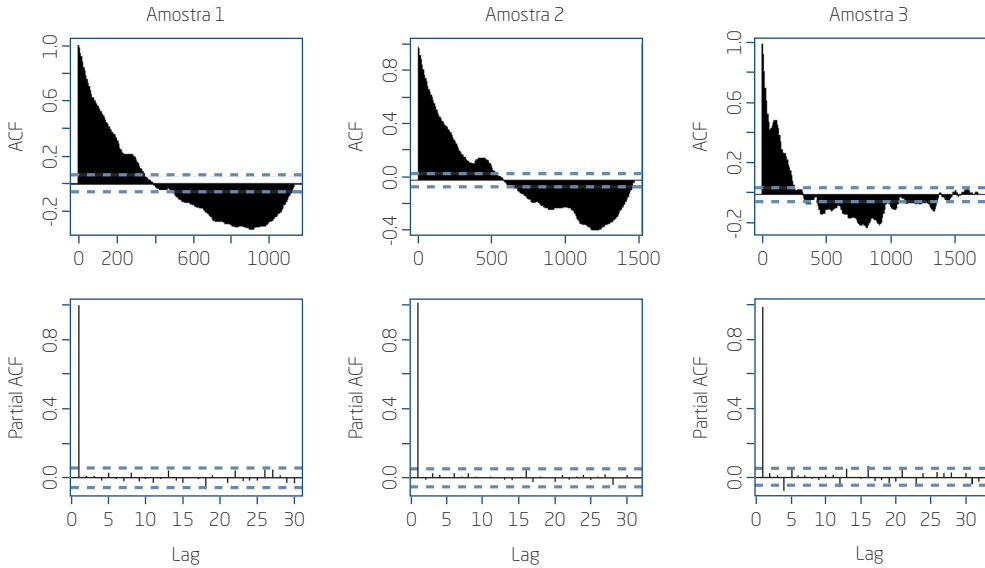
	Amostra 1 ( <i>downtrend</i> )	Amostra 2 ( <i>uptrend</i> )
Teste ADF		
Nº de observações	1.135	1.466
Estatística de Dickey-Fuller	-10,24	-11,86
Lagorder	10	11,00
p-valor	< 0,01	< 0,01
Hipótese alternativa: estacionariedade		
Teste PP		
Nº de observações	1.135	1.466
Estatística de Dickey-Fuller	-1099,3	-1403,60
Lagorder	7,00	7,00
p-valor	< 0,01	< 0,01
Hipótese alternativa: estacionariedade		

Fonte: Elaborada pelos autores a partir do pacote RSSA do *software* estatístico R.

A Figura 4.2.3 apresenta as autocorrelações e autocorrelações parciais de cada amostra. A fim de serem considerados parcimoniosos, foram especificados modelos com a ordem máxima igual a 5 para os parâmetros  $p$  e  $q$  e ordem máxima igual a 1 para os parâmetros  $r$  e  $s$ .

A seleção do modelo ótimo para cada conjunto amostral se deu pelo critério de Akaike. A Figura 4.2.4 mostra os parâmetros estimados dos modelos ótimos e os respectivos critérios AIC.

**(Figura 4.2.3)**  
**ACF E PACF DAS SÉRIES**



Fonte: Elaborada pelos autores a partir do pacote RSSA do *software* estatístico R.

**(Figura 4.2.4)**

**PARÂMETROS E CRITÉRIO AIC DOS MODELOS SELECIONADOS**

Amostra*	ARIMA	GARCH	AIC
1 - Downtrend	(1, 1, 1)	(1, 0)	-15,782
2 - Uptrend	(5, 1, 1)	(1, 0)	-14,909
3 - Range	(4, 0, 1)	(1, 1)	-15,547

\* Não inclui o horizonte de previsão.

Fonte: Elaborada pelos autores a partir do pacote RSSA do *software* estatístico R.

**4.3 Comparação dos resultados**

Uma vez estimados os parâmetros, estabeleceu-se o horizonte de previsão de cinco minutos. As figuras 4.3.1, 4.3.2 e 4.3.3 apresentam as previsões realizadas por cada modelo, bem como o resultado das medidas de qualidade para as amostras 1, 2 e 3, respectivamente.

(Figura 4.3.1)

RESULTADOS DE PREVISÃO PARA A AMOSTRA 1

Amostra 1 - <i>Downtrend</i> Observações: 1.135					
Observações	Previsão ARIMA-GARCH	Previsões SSA	Medidas de qualidade	Modelos	
				ARIMA (1, 1, 1) GARCH (1, 0)	SSA L = T/4
1,12922	1,129090	1,129181			
1,12930	1,129080	1,129194			
1,12928	1,129070	1,129226	MAPE	0,01628%	0,005898%
1,12926	1,129060	1,129274	RMSE	0,000187	0,000078
1,12921	1,129050	1,129330	THEIL	0,000083	0,000034

Fonte: Elaborada pelos autores a partir de dados obtidos pelo Marketscope 2.0®.

(Figura 4.3.2)

RESULTADOS DE PREVISÃO PARA A AMOSTRA 2

Amostra 2 - <i>Uptrend</i> Observações: 1.466					
Observações	Previsão ARIMA-GARCH	Previsões SSA	Medidas de qualidade	Modelos	
				ARIMA (5, 1, 1) GARCH (1, 0)	SSA L = T/11
1,13589	1,136392	1,135931			
1,13606	1,136406	1,135931			
1,13575	1,136390	1,135920	MAPE	0,042457%	0,008645%
1,13586	1,136372	1,135857	RMSE	0,000492	0,000113
1,13596	1,136371	1,135836	THEIL	0,000217	0,000050

Fonte: Elaborada pelos autores a partir de dados obtidos pelo Marketscope 2.0®.

**(Figura 4.3.3)****RESULTADOS DE PREVISÃO PARA A AMOSTRA 3**

Amostra 3 - Range Observações: 1.681					
Observações	Previsão ARIMA-GARCH	Previsões SSA	Medidas de qualidade	Modelos	
				ARIMA (4, 0, 1) GARCH (1, 1)	SSA L = T/2
1,12631	1,126359	1,126352			
1,12639	1,126355	1,126379			
1,12651	1,126348	1,126406	MAPE	0,008221%	0,007227%
1,12648	1,126351	1,126435	RMSE	0,000104	0,000107
1,12626	1,126348	1,126465	THEIL	0,000046	0,000047

Fonte: Elaborada pelos autores a partir de dados obtidos pelo Marketscope 2.0®.

Conforme os resultados mostrados nas figuras, as previsões realizadas pelo SSA foram as que mais se aproximaram das observações originais para os três casos em que a técnica foi aplicada e para todas as medidas de qualidade, exceto na amostra 3, em que o RMSE e o TIC mostraram-se melhores para o ARIMA-GARCH. Tal fato permite concluir que as previsões do SSA, no que diz respeito às taxas de câmbio no período em questão, são mais adequadas que aquelas obtidas pelo modelo ARIMA-GARCH, independentemente do movimento do mercado, ou seja, tanto para tendências de alta e baixa como para quando não houve tendência dominante.

## 5. CONSIDERAÇÕES FINAIS

O objetivo deste trabalho foi avaliar qual a modelagem do fechamento da taxa de câmbio (*ask*) do par EUR/USD gera previsões mais satisfatórias quando da utilização de dados de alta frequência (um minuto). Para alcançar tal objetivo, a pesquisa foi dividida em duas partes: a primeira consistiu na estimação dos parâmetros de cada técnica; e a segunda, na utilização dos modelos estimados para previsão das séries em um horizonte de tempo de cinco minutos.

Os resultados mostraram que as previsões foram mais satisfatórias quando realizadas por meio do SSA. Apenas quando observada a série em que não houve tendência dominante (amostra 3), o modelo ARIMA-GARCH apresentou menores erros de previsão quando observadas as métricas RMSE



e TIC. Ainda assim, obteve um MAPE maior que o SSA. Os resultados obtidos no trabalho estão em consonância com o encontrado na literatura, como em Esquivel et al. (2013), que destacam a maior acurácia do SSA se comparado a outros modelos de séries temporais. Além disso, segundo Hassani (2007), o SSA não faz pressuposições acerca dos dados e, diferentemente de outros métodos, pode ser aplicado em amostras pequenas, constituindo, assim, uma vantagem do método.

Cabe ressaltar que os tipos de previsões decorridas das duas técnicas, pelo SSA e pelo modelo clássico ARIMA-GARCH, são diferentes. No primeiro caso, busca-se identificar e desmembrar as séries em seus componentes de tendência, componentes periódicos e ruídos. Em seguida, elimina-se o componente de ruído e efetuam-se as previsões. No segundo caso, assume-se que o ruído é o erro estocástico com variância condicional e busca-se modelá-lo com os demais componentes autorregressivos e sazonais.

Quanto às limitações da pesquisa, deve-se mencionar o fato de que as amostras utilizadas na modelagem e previsão da taxa de câmbio possuem tamanhos diferentes. Esse pode ser um fator relevante quando da estimação dos parâmetros do modelo ARIMA-GARCH. De acordo com o trabalho de Ng e Lam (2006), o tamanho da amostra pode guiar a estimação para diferentes soluções ótimas locais, e, por esse motivo, as amostras selecionadas são relativamente pequenas.

Outra questão a ser considerada é o período do qual as amostras foram selecionadas. Se as observações fossem provenientes de outra época do ano, as previsões realizadas pelo modelo clássico ainda seriam menos satisfatórias? Essa é uma pergunta pertinente, visto que o mercado FOREX apresenta diferentes volatilidades em diferentes períodos do ano e até mesmo em períodos dentro de um mesmo dia.

Tendo em vista os resultados encontrados e as limitações deste trabalho, sugere-se como tema de pesquisas posteriores a previsão da volatilidade da taxa de câmbio EUR/USD; o impacto da diferença de volatilidade em diferentes períodos nas previsões dos modelos; a previsão dos retornos e/ou volatilidade dos retornos no mercado acionário brasileiro; a utilização do SSA para previsões com amostras dinâmicas (ou seja, rolagem de observações); e a utilização do SSA para especificação de um modelo de *value at risk* não paramétrico.



# APPLYING SINGULAR SPECTRUM ANALYSIS AND ARIMA-GARCH FOR FORECASTING EUR/USD EXCHANGE RATE

## ABSTRACT

**Purpose:** The objective of this article is to model a minute series of exchange rates for the EUR/USD pair using the singular spectrum analysis (SSA) and ARIMA-GARCH methods and evaluate which one offers better forecasts for a five-minute horizon.

**Originality/value:** Despite being a successful technique in other branches of science, the application of SSA in finance is quite new. Furthermore, exchange rate modeling is a complex problem, comprising statistical concepts and properties. However, despite the complexity, the analysis of this series is extremely important for several agents playing, directly or indirectly, a role in economy and the financial market.

**Design/methodology/approach:** Time series models were estimated using the ARIMA-GARCH and SSA techniques, taking into account three samples of ask exchange rate (closing): uptrend, downtrend, and no well-defined trend.

**Findings:** The forecasts carried out by the SSA were the ones closest to the original observations for the three cases. Regarding the quality measurements, SSA obtained the best results for both uptrend and downtrend samples; for the sample with no well-defined trend, the findings indicated that the ARIMA-GARCH technique attained better results. However, it was concluded that the SSA forecasts, regarding exchange rates during the studied period, are more appropriate than the ones obtained by the ARIMA-GARCH model, regardless of market movement.

## KEYWORDS

Exchange market. Exchange rates. Dollar. Euro. Time-series forecast.



## REFERÊNCIAS

- Abraham, A. (2005). Hybrid soft and hard computing based Forex monitoring systems. In N. Nedjah & L. de M. Mourelle (Eds.), *Fuzzy systems engineering* (pp. 113–129). Berlin: Springer. doi:10.1007/11339366\_5
- Alamili, M. (2011). *Exchange rate prediction using support vector machines*. Thesis doctorate, Delft University of Technology, Delft, Netherlands.
- Allen, M. R., & Smith, L. A. (1996). Monte Carlo SSA: Detecting irregular oscillations in the presence of colored noise. *Journal of Climate*, 9(12), 3373–3404. doi:10.1175/1520-0442(1996)009<3373:MCSDIO>2.0.CO;2
- Asai, M., McAleer, M., & Yu, J. (2006). Multivariate stochastic volatility: A review. *Econometric Reviews*, 25(2-3), 145–175. doi:10.1080/07474930600713564
- Atkočiūnas, V., Mačiulis, N., Klimavičienė, A., & Kalendienė, J. (2010). *Short-term currency exchange rate forecasting with econometric models*. Thesis doctorate, ISM Vadybos ir ekonomikos universitetas, Vilnius, Lithuania.
- Bank for International Settlements (2013). *Triennial Central Bank Survey. Foreign exchange turnover in April 2013: Preliminary global results*. Recuperado de <http://www.bis.org/publ/rpfx13fx.pdf>
- Bauwens, L., Laurent, S., & Rombouts, J. V. (2006). Multivariate GARCH models: A survey. *Journal of Applied Econometrics*, 21(1), 79–109. doi:10.1002/jae.842
- Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, 31(3), 307–327. doi:10.1016/03044076(86)90063-1
- Box, G. E., & Jenkins, G. M. (1976). *Time series analysis: Forecasting and control*. San Francisco: Holden-Day.
- Brooks, C. (2008). *Introductory econometrics for finance*. Cambridge: Cambridge University.
- Broomhead, D. S., Jones, R., King, G. P., & Pike, E. R. (1987). Singular system analysis with application to dynamical systems. In E. R. Pike & L. A. Lugiato (Eds.), *Chaos, noise and fractals* (pp. 15–27). Boca Raton, FL: CRC Press.
- Broomhead, D. S., & King, G. P. (1986). Extracting qualitative dynamics from experimental data. *Physica D: Nonlinear Phenomena*, 20(2–3), 217–236. doi:10.1016/0167-2789(86)90031-X



- Chaudhuri, T. D., & Ghosh, I. (2016). Artificial neural network and time series modeling based approach to forecasting the exchange rate in a multivariate framework. *Journal of Insurance and Financial Management*, 1(5), 92–123. doi:arxiv.org/abs/1607.02093
- Dacorogna, M. M., Müller, U. A., Pictet, O. V., & Vries, C. G. (2001). Extremal forex returns in extremely large data sets. *Extremes*, 4(2), 105–127. doi:10.1023/A:1013917009089
- Danilov, D., & Zhigljavsky, A. (1997). *Principal components of time series: The “Caterpillar” method*. St. Petersburg: University of St. Petersburg.
- Elsner, J. B., & Tsonis, A. A. (1996). *Singular spectrum analysis: A new tool in time series analysis*. New York: Plenum Press.
- Esquivel, R. M., Senna, V., & Gomes, G. S. S. (2013). Análise espectral singular: Comparação de previsões em séries temporais. *Revista ADM. MADE*, 16(2), 87–101.
- Ghil, M., & Vautard, R. (1991). Interdecadal oscillations and the warming trend in global temperature time series. *Nature*, 350(6316), 324–327. doi:10.1038/350324a0
- Golyandina, N., Nekrutkin, V., & Zhigljavsky, A. A. (2001). *Analysis of time series structure: SSA and related techniques*. Boca Raton: Chapman & Hall, CRC.
- Groff, G. K. (1973). Empirical comparison of models for short range forecasting. *Management Science*, 20(1), 22–31. doi:10.1287/mnsc.20.1.22
- Hassani, H. (2007). Singular spectrum analysis: Methodology and comparison. *Journal of Data Science*, 5, 239–257.
- Hassani, H. (2010). Singular spectrum analysis based on the minimum variance estimator. *Nonlinear Analysis: Real World Applications*, 11(3), 2065–2077. doi:10.1016/j.nonrwa.2009.05.009
- Hassani, H., & Thomakos, D. (2010). A review on singular spectrum analysis for economic and financial time series. *Statistics and its Interface*, 3(3), 377–397. doi:10.4310/SII.2010.v3.n3.a11
- Hassani, H., & Zhigljavsky, A. (2009). Singular spectrum analysis: methodology and application to economics data. *Journal of Systems Science and Complexity*, 22(3), 372–394. doi:10.1007/s11424-009-9171-9
- Lai, K. K., Yu, L., & Wang, S. (2005). A neural network and web-based decision support system for forex forecasting and trading. In Y. Shi, W. Xu, & Z. Chen (Eds.), *Data Mining and knowledge management* (pp. 243–253). Berlin: Springer. doi:10.1007/978-3-540-30537-8\_27

- Makridakis, S., & Hibon, M. (1997). ARMA models and the Box–Jenkins methodology. *Journal of Forecasting*, 16(3), 147–163. doi:10.1002/(SICI)1099-131X(199705)16:3<147::AID-FOR652>3.0.CO;2-X
- Makridakis, S., Hibon, M., & Moser, C. (1979). Accuracy of forecasting: An empirical investigation. *Journal of the Royal Statistical Society. Series A (General)*, 142(2), 97–145. doi:10.2307/2345077
- Menezes, M. L. D., Cassiano, K. M., Souza, R. M. D., Teixeira Júnior, L. A., Pessanha, J. F. M., & Souza, R. C. (2014). Modelagem e previsão de demanda de energia com filtragem SSA. *Revista da Estatística da Universidade Federal de Ouro Preto*, 3(2), 170–187.
- Murphy, J. J. (1999). *Technical analysis of the financial markets: A comprehensive guide to trading methods and applications*. New York: New York Institute of Finance.
- Ng, H. S., & Lam, K. P. (2006). How does sample size affect GARCH models? *Proceedings of the 2006 Joint Conference on Information Sciences*, Kaohsiung, Taiwan. Recuperado de [https://www.researchgate.net/publication/221556756\\_How\\_does\\_Sample\\_Size\\_Affect\\_GARCH\\_Models](https://www.researchgate.net/publication/221556756_How_does_Sample_Size_Affect_GARCH_Models)
- Pan, V. Y., Rami, Y., & Wang, X. (2002). Structured matrices and Newton's iteration: Unified approach. *Linear Algebra and its Applications*, 343, 233–265. doi:10.1016/S0024-3795(01)00336-6
- Ravi, V., Lal, R., & Kiran, N. R. (2012). Foreign exchange rate prediction using computational intelligence methods. *International Journal of Computer Information Systems and Industrial Management Applications*, 4, 659–670.
- Rodríguez-Aragón, L. J., & Zhigljavsky, A. (2010). Singular spectrum analysis for image processing. *Statistics and Its Interface*, 3(3), 419–426. doi:10.4310/SII.2010.v3.n3.a14
- Royer, J. C., Wilhelm, V. E., & Patias, J. (2015). Previsão de séries temporais de subpressão de barragens com filtragem SSA e regressão múltipla com modelagem ARIMA dos resíduos. *Congresso de Métodos Numéricos em Engenharia*, Lisboa, Portugal.
- Sanei, S., & Hassani, H. (2015). *Singular spectrum analysis of biomedical signals*. Boca Raton: CRC Press.
- Vautard, R., & Ghil, M. (1989). Singular spectrum analysis in nonlinear dynamics, with applications to paleoclimatic time series. *Physica D: Nonlinear Phenomena*, 35(3), 395–424. doi:10.1016/0167-2789(89)90077-8
- Vautard, R., Yiou, P., & Ghil, M. (1992). Singular-spectrum analysis: A toolkit for short, noisy chaotic signals. *Physica D: Nonlinear Phenomena*, 58(1–4), 95–126. doi:10.1016/0167-2789(92)90103-T



World Trade Organization (2014). *International Trade Statistics*. Recuperado de [https://www.wto.org/english/res\\_e/statis\\_e/its2014\\_e/its14\\_toc\\_e.htm](https://www.wto.org/english/res_e/statis_e/its2014_e/its14_toc_e.htm)  
Yao, J., & Tan, C. L. (2000). A case study on using neural networks to perform technical forecasting of forex. *Neurocomputing*, 34(1-4), 79-98. doi:10.1016/S0925-2312(00)00300-3

## NOTAS DOS AUTORES

**Rafael J. Abreu**, Faculdade de Ciências Econômicas (Face), Universidade Federal de Minas Gerais (UFMG); **Rafael M. Souza**, Departamento de Engenharia Elétrica, Pontifícia Universidade Católica do Rio de Janeiro (PUC-Rio); e **Joice G. Oliveira**, Departamento de Administração e Contabilidade, Universidade Federal de Viçosa (UFV).

Rafael J. Abreu é agora analista comercial da empresa Ipanema Comercial Exportadora e Importadora Ltda.; Rafael M. Souza é agora professor adjunto do Departamento de Ciências Contábeis da Universidade Federal de Minas Gerais (UFMG); Joice G. de Oliveira é agora doutoranda em Controladoria e Contabilidade do Departamento de Ciências Contábeis da Universidade Federal de Minas Gerais (UFMG).

Correspondências sobre este artigo devem ser enviadas para Rafael J. Abreu, Rua Joaquim Borges, 97, ap. 101, Jardim Aeroporto, Alfenas, Minas Gerais, Brasil, CEP 37130-810.

E-mail: [abreu.fael@gmail.com](mailto:abreu.fael@gmail.com)

### CONSELHO EDITORIAL

**Editoras-chefe**  
*Janette Brunstein*  
*Silvia Marcia Russi de Domênico*

**Editor Associado**  
*Paulo Ceretta*

**Suporte Técnico**  
*Vitória Batista Santos Silva*

### PRODUÇÃO EDITORIAL

**Coordenação Editorial**  
*Jéssica Dametta*

**Preparação de originais**  
*Carlos Villarruel*

**Revisão**  
*Studio Ayres*

**Diagramação**  
*Emap*

**Projeto Gráfico**  
*Libro*