

---

# Modelagem e Previsão do Preço do Café Brasileiro

---

Carlos Enrique Carrasco-Gutierrez<sup>1</sup>

Fernanda Matos de Moura Almeida<sup>2</sup>

**Resumo:** O Brasil se destaca como maior produtor de café do mundo sendo responsável por 30% do mercado internacional, além de maior exportador e o segundo maior consumidor, atrás somente dos Estados Unidos. Este trabalho tem como finalidade encontrar um modelo econométrico na presença de quebra estrutural, que seja capaz de prever o preço do café no Brasil para auxiliar os produtores e comercializadores nas tomadas de decisões sobre os investimentos de café. Os dados utilizados são os preços do café Arábica recebido pelos produtores Brasileiros no período de 1997 até 2009 em frequência mensal. Variáveis macroeconômicas foram inseridas nos modelos a fim de verificar a influência destas na previsão do preço do café. Os resultados da modelagem identificaram uma quebra estrutural nos preços reais e dependência com algumas variáveis macroeconômicas. Em relação aos resultados de previsão o modelo identificado fornece um percentual de erro médio absoluto de previsão na ordem de 1,6%.

**Palavras-chave:** Previsão, Modelagem, Preço do Café.

**JEL:** C51; C52; C53

1. Doutor em Economia pela EPGE - Fundação Getúlio Vargas - RJ (2008), Doutor em Engenharia Elétrica (métodos de apoio à decisão) pela Pontifícia Universidade Católica do Rio de Janeiro (2006) e Mestre em Engenharia Elétrica pela Pontifícia Universidade Católica do Rio de Janeiro (2002).

2. Mestre em Ciências Contábeis pela Fucape Business School. Possui graduação em Ciências Contábeis - Faculdades Integradas De Caratinga (2004) e pós-graduação em Estratégia Empresarial pela FIJ - Faculdades Integradas De Jacarepaguá (2006).

## *Modelling and Forecasting the Coffee Prices in Brazil*

**Abstract:** *The Brazil stands as the largest coffee producer in the world responsible for 30% of the international market, beyond largest exporter and second largest consumer, behind only the United States. This work aims to find an econometric model able to predict the price of coffee in Brazil in order to help producers and traders in decision making on investments coffee. The data come from Arabica coffee prices received by producers Brazilians from 1997 until 2009 in monthly frequency. Macroeconomic variables were included in the models to verify the influence on the forecast price of coffee. The modeling results have identified a structural break in real prices and dependence with some macroeconomic variables. Regarding the results of forecasting models identified the percentage of average absolute prevision error of around 1.6%.*

**Keywords:** *Forecasting, Modelling, Price of Coffee.*

**JEL:** C51; C52; C53

### **1. Introdução**

A cafeicultura é uma atividade econômica e socialmente importante em nível mundial. Na economia brasileira o café apresenta uma participação expressiva da produção agrícola, refletindo positivamente na balança comercial. A produção de café do Brasil, de acordo com dados da Confederação da Agricultura e Pecuária do Brasil (CNA, 2008) representa 40% da produção total mundial. O Brasil se destaca como o maior produtor de café do mundo, além de maior exportador (MAPA, 2010) e segundo maior consumidor. Margarido & Barros (2000) confirmam a liderança do Brasil na produção e exportação agrícola, mas enfatizam seu baixo poder de determinar os preços desses produtos uma vez que o país está exposto a variações de preços externos. Os autores relatam que o setor agrícola, é o setor mais sensível aos choques de oferta e demanda que consequentemente afetam o preço desses produtos.

Os preços do café no Brasil, até 1990 eram regulados pelo Instituto Brasileiro do Café (IBC) com intervenção do Governo visando aumentar a demanda e reduzir a oferta sempre que a produção se excedia. A partir da referida data o Brasil passou a receber influência do mercado futuro para determinar o preço interno do café proporcionando maior segurança aos envolvidos nesse negócio. Os contratos do mercado futuro representam uma opção de venda antecipada para os produtores (BACHA, 1998; MATTIELLO, 2005). Em virtude do surgimento deste mercado, os produtores podem garantir melhores preços para o café que produzem antes mesmo de colher o seu produto. Entretanto,

essa movimentação com o mercado futuro, ainda gera insegurança entre os produtores visto que o preço do café tem se mostrado instável e as opções ofertadas pelo mercado futuro visam atendimento da demanda pelo produto. Assim, é comum os produtores se sujeitarem a essas negociações sem terem a oportunidade de comparar as ofertas do mercado com alguma estimativa de preço do café.

Se o produtor tem uma previsão do preço que será pago pelo seu café na época da safra, torna-se possível analisar que tipo de investimentos poderá fazer e qual será o melhor período para vender seu café garantindo uma rentabilidade maior. Assim, o conhecimento do comportamento dos preços poderá ser extremamente útil nas suas tomadas de decisões com relação ao planejamento da produção, à manutenção e formação de estoques, podendo desta forma aproveitar de maneira mais eficiente as fases de baixa e de alta nos preços, para a maximização dos lucros. De forma geral a previsão do preço do café, auxiliará o produtor nas tomadas de decisões que são importantes para seu negócio.

Do ponto de vista dos comercializadores de café, a previsão do preço é muito importante no momento de realizar contratos no mercado spot e futuro. Normalmente, o preço no mercado spot reage ou cai antecipadamente em relação ao preço recebido pelo produtor, e em maior magnitude. Portanto, a previsão do preço será um parâmetro de decisão.

Sendo assim, um estudo que apresente uma previsão do preço do café poderá auxiliar os produtores e comerciantes nas tomadas de decisões. Nessa perspectiva, esta pesquisa objetiva encontrar um modelo econométrico de séries temporais, adequado à previsão do preço do café recebido pelos produtores Brasileiros.

Dois tipos de café são os mais importantes no Brasil: o café arábica e o café robusta. O café arábica se destaca pela sua capacidade de consumo *in natura*, pelas suas propriedades organolépticas que fazem com que este tipo de café seja apreciado pelo seu sabor mais suave e menos conteúdo de cafeína e tudo isto seja compensado por um preço maior, enquanto que a vantagem do café Robusta está no maior rendimento e menor vulnerabilidade a doenças e só é consumido via *blend*. É sabido que existem mais de quarenta variedades de café, as variedades arábica e robusta são as únicas relevantes e representam entre 60%-70% e 40%-30% da produção mundial, respectivamente. Este artigo estuda o preço do café arábica produzido pelo Brasil visto a sua representatividade na produção total.

Este trabalho está organizado, além desta introdução, em quatro seções. Na próxima seção é apresentada uma revisão bibliográfica existente a respeito da modelagem e previsão do preço de café. Na terceira seção descreve-se os dados, a metodologia, e apresenta-se os resultados da modelagem e previsão. Por fim, na quarta seção encontram-se as conclusões.

## 2. Fundamentação Teórica

Diferentes trabalhos têm-se dedicado ao estudo do mercado de café, tanto do ponto de vista teórico como prático. Vogelvang (1980, 1988, 1990, 1992), Lord (1991), Bressan (2004) são alguns exemplos. Vogelvang (1988) menciona que o estudo do mercado de café depende dos objetivos a serem seguidos, por exemplo:

- Especificação, estimação e teste de equações estruturais para obter *insight* na estrutura causal do mercado;
- Modelagem parcial ou total do mercado de café para testar hipóteses específicas das relações entre algumas variáveis;
- Modelagem com objetivos de previsão;
- Construção de um modelo para formulação de políticas.

Vogelvang (1980) iniciou um estudo para formação de preço do café estimando modelos econométricos baseado no comportamento microeconômico dos participantes do mercado presente e futuro. Vogelvang (1990) estuda o mercado de café incluindo outras commodities e analisa co-integração entre os preços do café, cacau, chá e açúcar. Vogelvang (1992) utilizando preços trimestrais no período 1960-1982 investiga as relações de longo prazo entre os preços spots de quatro tipos de café por meio de uma análise de cointegração, encontrando provas sobre a existência de duas relações de longo prazo. Otero & Milas (1998) ampliam a amostra até 1998, confirmando os resultados encontrados por Vogelvang (1992).

Outros trabalhos têm-se dedicado à modelagem de séries temporais com o objetivo de obter previsões dos preços. Por exemplo, um grupo de sete commodities: café, cacau, cobre, açúcar, algodão, milho e soja, foi estudado por Lord (1991). No Brasil, Bressan (2004) estimou um modelo econométrico para encontrar as previsões de café, soja e boi gordo a fim de auxiliar nas decisões de compra e venda de contratos futuros, levando em conta dados da BM&F. Ribeiro, Souza & Rogers (2006) analisaram o comportamento do preço do café à vista e futuro em relação aos estoques e exportações. Para tanto, utilizaram um modelo de vetores autoregressivos (VAR) aplicado a séries temporais. Almeida Pinto et al. (2008) estudaram o comportamento dos preços de cacau, café, cana de açúcar, laranja e soja do Brasil, a fim de analisar a sensibilidade dos preços recebidos pelos produtores dessas commodities.

Poucos trabalhos têm se dedicado à previsão do preço do café no Brasil, nosso objetivo está na procura de um modelo econométrico simples capaz de prever o preço do café brasileiro. A ideia neste trabalho é propor um modelo de fácil implementação que atenda os objetivos propostos. Os modelos dinâmicos de defasagem distribuída têm sido muito úteis para este fim do ponto de vista da

facilidade na sua implementação. Existe uma tendência de procurar modelos mais sofisticados que nem sempre os ganhos em termos de previsão compensam o esforço da modelagem. Justificamos o uso dos modelos apresentados neste trabalho em função do resultado de previsão obtido.

Incluímos neste estudo outros fatores de instabilidade do preço do café que estão diretamente relacionados a variáveis como: taxa de câmbio, taxa de juros, crédito rural, PIB do Brasil e PIB dos principais países que importam café do Brasil<sup>1</sup> entre outras variáveis macroeconômicas.

Os resultados desta pesquisa mostram modelos bem especificados apropriados para a previsão do preço do café, sendo o modelo que apresenta melhores resultados de previsão, aquele que considera a variável taxa de câmbio e algumas defasagens do preço na sua especificação. O modelo identificado fornece erros percentuais médios absolutos de previsão na ordem de 1,6% quando comparados com dados reais.

### **3. Metodologias e Resultados**

#### **3.1 Dados**

O período da análise deste trabalho estende-se de Janeiro de 1997 a Dezembro de 2009, com periodicidade mensal. Os preços de café arábica foram extraídos da Associação Brasileira da Indústria do Café (ABIC, 2009)<sup>2</sup>. Os dados de taxa de câmbio, taxa de juros, crédito rural, deflator (IGP-M), PIB do Brasil e os índices de PIB dos principais países que importam café do Brasil foram obtidos do IpeaData. As séries foram deflacionadas considerando o Índice Geral de Preços de Mercado (IGPM) usando-se como base a média de 1995 = 100.

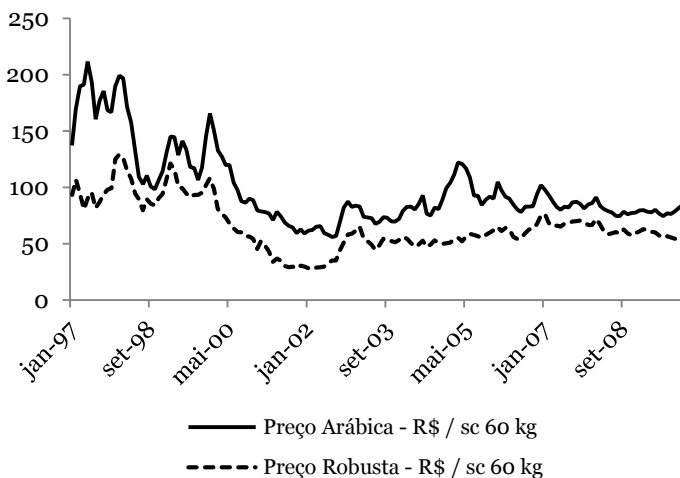
A Figura 1 mostra as séries dos preços recebidos pelos produtores para dois tipos de café: o café arábica tipo 6 bebida dura e o robusta tipo 7. Pode ser observada uma mudança de comportamento da série de preços após o ano 1999 a partir da qual os preços estão mais estáveis. O preço do café no Brasil sempre foi marcado por altas e baixas em decorrência, entre outros fatores, do aumento (redução) da produtividade refletindo em aumento (redução) do estoque de café<sup>3</sup>.

1 Entre 1976 e 2004, os principais países importadores de café foram os Estados Unidos, Alemanha, Itália, Japão e Bélgica.

2 <http://www.abic.com.br/publico/cgi/cgilua.exe/sys/start.htm?sid=47#139>

3 Foram feitos testes de sazonalidade para a série arábica mostrando pouca influência no comportamento dos dados. Optamos por usar a série original.

FIGURA 1. MÉDIA MENSAL DOS PREÇOS DE CAFÉ RECEBIDOS PELOS PRODUTORES (1995 = 100).



Fonte: ABIC (2009)

\*Notas: Preço do café arábica e conilon - preço real.\*Arábica tipo 6 bebida dura e café robusta tipo 7  
 Preço recebido pelo produtor, período Janeiro 1997 a Dezembro 2009.

TABELA 1. ESTATÍSTICAS DO TESTE DE RAIZ UNITÁRIA

Séries	ADF	PP
LogPreço Arábica	2,673208 (0,081)	-2,035135 (0,2717)
LogTaxa de câmbio	-2,171475 (0,2176)	-2,055919 (0,2630)
LogCrédito rural	0,530543 (0,9873)	0,404767 (0,9826)
Taxa de juros	-4,961506* (0,000)	-4,559129* (0,0002)
Log PIB Brasil	0,761638 (0,9931)	-0,193490 (0,9355)
Log PIB USA	-2,343872 (0,1597)	-3,166075 (0,0240)

Nota: A hipótese nula dos dois testes é a presença de raiz unitária. Valores com logaritmo, com exceção da taxa de juros. Valor crítico da estatística ADF ao nível de significância de 5%. -2,880211. (Exogenous Constant). Valor crítico da estatística PP ao nível de significância de 5%. (Exogenous Constant). Os valores em parênteses são as probabilidades (P-valor).

### 3.2 Teste de Raiz Unitária

Antes de qualquer análise econométrica com dados de séries temporais é necessário investigar as propriedades de estacionaridade das séries envolvidas. Se ao realizarmos o teste de raiz unitária identificamos que a série contém uma raiz unitária (ordem de integração de 1,  $I(1)$ ) transformaremos esta variável em estacionária por meio da diferenciação, antes da sua inclusão no modelo

econométrico<sup>4</sup>. A Tabela 1 mostra os resultados dos testes de Augmented Dickey-Fuller (ADF) e Phillips-Perron (PP) para todas as séries envolvidas<sup>5</sup>.

Todas as estatísticas mostram que o preço de café arábica, crédito rural, taxa de câmbio, PIB Brasil e PIB Americano apresentam raiz unitária. A única variável estacionária é a taxa de juros<sup>6</sup>.

### 3.3 Teste de Quebra Estrutural

É conhecida a fraqueza dos resultados dos testes de ADF e PP na presença de potenciais quebras estruturais mostrando evidências de não estacionariedade. Em outras palavras, para uma série encontrada como não estacionária, pode existir a possibilidade de que de fato seja estacionária em torno da quebra estrutural,  $I(0)$ , mas erroneamente classificada como  $I(1)$ <sup>7</sup>.

Perron (1989) desenvolveu um procedimento para testar a raiz unitária permitindo a incorporação de uma quebra estrutural exógena. Isto é, a quebra estrutural é conhecida. Por outro lado, há uma vasta literatura sobre quebras endógenas no qual a data da quebra é determinada seguindo algum critério de identificação através do uso de algum tipo de *outlier*<sup>8</sup>.

Zivot & Andrews (1992), Perron & Vogelsang (1992), Perron (1997), e Clemente-Montanes-Reyes (1998) propuseram testes de quebra estrutural no qual o ponto de quebra era determinado endogenamente. Alguns autores têm mostrado as vantagens e desvantagens destes testes. Shrestha & Chowdhury (2005) argumentam que em caso da existência de quebra estrutural a potência do teste Perron-Vogelsang é superior ao de Zivot-Andrews. Enders (2004) argumenta que Perron & Vogelsang (1992) é mais apropriado se a data da quebra é incerta. Baum (2004) justifica os testes de Perron & Vogelsang (1992) e Clemente-Montanes-Reyes (1998) como os mais preferidos.

Os testes de quebra estrutural endógena de Perron & Vogelsang e Clemente-Montanes-Reyes oferecem dois tipos de modelos:

- a) O modelo *outlier* aditivo (OA) o qual captura mudanças abruptas na média das séries.
- b) O modelo *outlier* inovacional (OI) o qual captura a mudança gradual na média das séries.

4 As estimativas de modelos econométricos com séries econômicas temporais não estacionárias produzem resultados espúrios (Granger & Newbold, 1974)

5 Embora tenhamos considerado o PIB dos maiores países exportadores de café, através de uma primeira análise de correlação (resultados no anexo) mostraram que o PIB Americano com o PIB dos outros países estão altamente correlacionados. Consideramos na análise apenas o PIB USA.

6 Consideramos como variável determinística a constante em todos os testes. Os resultados são ao nível de significância de 5%.

7 De fato Perron mostra que ao considerar quebras estruturais nas séries os testes tradicionais têm menos habilidade em rejeitar a hipótese de raiz unitária que na verdade é falsa.

8 *Outliers* são observações anormais ou excepcionais que estão distantes do resto dos dados.

Para o estudo em questão usamos o teste de Clemente-Montanes-Reyes (1998), cuja hipótese nula é a presença de raiz unitária contra a hipótese alternativa que a série seja estacionária com quebra estrutural. Na Tabela 2 mostra-se os resultados do teste de quebra estrutural endógena para a série de preços do café arábica.

TABELA 2. TESTE DE RAIZ UNITÁRIA DE CLEMENTE - MONTANES- REYES

Modelo	Lag K	Estatística	Valores Críticos	Ponto de quebra
OA	10	-3,593*	-3,56	1999:10
OI	11	-6,173*	-4,27	1999:11

Nota : \* significa rejeição da hipótese nula ao nível de significância de 5%. O teste considera apenas uma quebra estrutural endógena.

Como a estatística de teste, em módulo, é maior que o valor crítico rejeita-se a hipótese nula de raiz unitária com quebra, portanto, a série de preços é estacionária com uma quebra. Esta informação deverá ser incorporada na fase de identificação e estimação do modelo na escolha da melhor especificação.

### 3.4 Modelo Econométrico

Neste trabalho usamos o modelo dinâmico de defasagens distribuídas na qual incorporamos a informação da quebra estrutural identificada. Este modelo é também conhecido como modelo Auto-regressivo de Defasagens Distribuídas (ADD) por levar em conta, como variáveis pré-determinadas, as defasagens da variável endógena. Também estimamos alguns modelos univariados a modo de comparar os desempenhos de previsão com os modelos propostos. Dentre os modelos univariados podemos destacar os modelos ARMA e ARIMA muito conhecidos na literatura. A seguir mostramos as especificações dos modelos univariados, o modelo de defasagem distribuída e finalmente o modelo dinâmico de defasagens distribuídas que generaliza os dois anteriores.

#### i) Modelos Univariados

O modelo univariado considera apenas a variável de interesse para a modelagem. Diz-se que o processo  $y_t$  segue um modelo  $ARMA(p, q)$  se:

$$y_t = \alpha + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \dots + \phi_p y_{t-p} + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \dots + \theta_q \varepsilon_{t-q} + \varepsilon_t$$

que pode ser escrito como  $y_t \phi(L) = \varepsilon_t \theta(L)$ , em que  $\phi(L)$  e  $\theta(L)$  são dois polinômios definidos como:

$$\phi(L) = 1 - \phi_1 L - \phi_2 L^2 - \dots - \phi_p L^p$$

e

$$\theta(L) = 1 + \theta_1 L + \theta_2 L^2 + \dots + \theta_q L^q$$

Se  $p = 1$  e  $q = 0$  temos o modelo autoregressivo de ordem 1 conhecido como AR(1); se  $p = 0$  e  $q = 1$  temos o modelo univariado de médias móveis conhecido como MA(1). O processo  $y_t$  segue um modelo *ARIMA*( $p, d, q$ ) se é definido como  $\Delta^d y_t \phi(L) = \varepsilon_t \theta(L)$  em que  $d$  é a ordem de integração do processo  $y_t$ . Nesse caso, se  $d = 1$ ,  $p = 1$  e  $q = 1$ , então  $y_t$  tem uma raiz unitária que resulta no modelo *ARIMA*(1, 1, 1); se  $d = 1$ ,  $p = 1$  e  $q = 0$  temos o modelo *ARIMA*(1, 1, 0) que pode ser escrito como  $(1 - L)y_t(1 - \phi_1 L) = \varepsilon_t$ .

### ii) Modelo de Defasagem Distribuída

Este modelo apresenta a variável endógena  $y_t$  como uma função de valores correntes e passados de uma ou mais variáveis exógenas. Considere o modelo com apenas uma variável exógena  $x_t$ .

$$y_t = \alpha + \phi_1 x_t + \phi_2 x_{t-1} + \dots + \phi_p x_{t-p} + \varepsilon_t$$

Observe que a reação de  $y_t$  não é explicada somente pelo ocorrido em  $x_t$  no tempo  $t$ , mas também pelo histórico de observação de  $x_t$  nos instantes  $t - 1, t - 2, \dots$ . Isto é,  $x_t$  tem influência sobre  $y_t$  até  $p$  períodos de tempo anteriores a  $t$ .

### iii) Modelo Dinâmico de Defasagens Distribuídas

Este modelo de séries temporais é mais geral que os anteriores e apresenta a variável endógena  $y_t$  como uma função de seus valores defasados e de variáveis exógenas  $x_t$ , de valores correntes e defasados por um ou mais períodos. O termo dinâmico é em referência o que as defasagens da variável endógena entram como variáveis exógenas. No modelo dinâmico de defasagem distribuída pelo fato deste descrever melhor a evolução da economia e suas reações encadeadas ao longo do tempo apresenta-se apropriado para a previsão do preço do café. No caso de uma única variável explicativa  $x_t$  o modelo é dado por:

$$y_t = \alpha + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \dots + \phi_p y_{t-p} + \theta_0 x_t + \theta_1 x_{t-1} + \theta_2 x_{t-2} + \dots + \theta_q x_{t-q} + \varepsilon_t$$

$$y_t = c + \sum_{i=1}^p \phi_i y_{t-i} + \sum_{j=0}^q \theta_j x_{t-j} + \varepsilon_t$$

A identificação dos modelos (a1)<sup>9</sup>, (a2) e (a3) consiste em encontrar os valores dos parâmetros  $p$  e  $q$ . Para estabelecer estes parâmetros utilizam-se os critérios de informação,  $I(p, q)$  que são funções destes parâmetros. Na prática estima-se o modelo dinâmico de defasagens distribuídas para vários valores de  $p$  e  $q$  e escolhe-se aquele que apresentar o menor valor do critério de informação. Para nossa análise usamos dois critérios de informação muito conhecidos na literatura, Akaike,  $AIC(p, q)$  e Schwarz  $SC(p, q)$ .

### 3.5 Identificação, Estimação e Teste

Considerando os resultados dos testes de raiz unitária com e sem quebra estrutural o modelo a ser identificado levará as séries de preço e taxa de juros em nível e as outras séries em primeira diferença. O modelo dinâmico de defasagem distribuída a ser identificado será da forma:

$$\ln P_t = c + DT_t + \sum_{i=0}^p \beta_i \ln P_{t-i} + \sum_{j=0}^q \theta_j X_{t-j} + \varepsilon_t$$

em que  $X_t = [txjur_t, \Delta txcam_t, \Delta \ln cred_t, \Delta \ln pibusa_t, \Delta \ln pibbras_t]$  é um vetor que contém variáveis exógenas,  $\theta_j$  é um vetor de parâmetros associados às variáveis exógenas,  $DT_t$  é uma variável dummy que representa a mudança na inclinação, devido à quebra estrutural, definida como  $DT_t = TB - t$  se  $t < TB$  e zero caso contrário;  $TB$  é a data da quebra e consideramos o valor de 1999:11 do resultado OI<sup>10</sup>;  $txjur$  é a taxa de juros,  $txcam$  é a taxa de câmbio,  $cred$  é o crédito rural,  $pibusa$  é o PIB dos Estados Unidos e  $pibbras$  é o PIB do Brasil. O símbolo  $\Delta$  é o operador defasagem definido como  $\Delta = 1 - L$ .

Os modelos foram identificados e estimados com informação correspondente ao período de 1997 até 2007, deixando os dois últimos anos (2008 e 2009) para analisar o desempenho de previsão dos modelos.

A identificação dos modelos foi feita a partir de um conjunto de especificações determinadas pelos parâmetros  $p$  e  $q$ . O número de defasagens de cada uma das variáveis independentes ( $q$ ), e defasagens do preço ( $p$ ), no modelo (3), foram determinados através da minimização do critério de informação Akaike ( $AIC$ ) e Schwarz ( $SC$ ) a partir de um máximo de 3 lag. Dentre um conjunto grande de modelos foram selecionados cinco com os menores valores destes critérios. Na Tabela B do Apêndice mostra-se os valores destes critérios e algumas estatísticas usuais de identificação. Os modelos identificados são mostrados a seguir:

9 O modelo univariado também pode ser identificado através da metodologia Box & Jenkins.

10 Acreditando que a mudança na série seja gradual consideramos o resultado de OI. De qualquer forma a diferença com o OA é apenas de um mês não tendo efeito significativo na escolha do modelo.

*Modelo 1:*

$$\ln P_t = c + DT_t + \beta_1 \Delta txcam_t + \beta_2 \Delta \ln cred_t + \beta_3 txjur_t + \beta_4 \ln P_{t-1} + \beta_5 \ln P_{t-2} + \varepsilon_t$$

*Modelo 2:*

$$\ln P_t = c + DT_t + \beta_1 \Delta txcam_t + \beta_2 \Delta \ln cred_t + \beta_3 \ln P_{t-1} + \beta_4 \ln P_{t-2} + \varepsilon_t$$

*Modelo 3:*

$$\ln P_t = c + DT_t + \beta_1 \Delta txcam_t + \beta_2 \Delta txcam_{t-1} + \beta_3 \ln P_{t-1} + \beta_4 \ln P_{t-2} + \varepsilon_t$$

*Modelo 4:*

$$\ln P_t = c + DT_t + \beta_1 \Delta txcam_{t-1} + \beta_2 \Delta \ln P_{t-1} + \beta_3 \ln P_{t-2} + \varepsilon_t$$

*Modelo 5:*

$$\ln P_t = c + DT_t + \beta_1 \Delta txcam_t + \beta_2 \Delta \ln cred_t + \beta_3 \Delta \ln cred_{t-1} + \beta_4 \ln P_{t-1} + \beta_5 \ln P_{t-2} + \varepsilon_t$$

Pode-se observar uma dependência dos modelos com as variáveis taxa de câmbio, taxa de juros, crédito rural e algumas defasagens do preço, porém, não incluem as séries do PIB.

Para determinar a validade destas especificações realizamos os testes de diagnóstico. Os resultados são apresentados na Tabela 3. O teste LM tem como hipótese nula a ausência de autocorrelação até a ordem de defasagem  $h$  e não indica, em nenhum caso, a presença de correlação serial ao nível de significância de 5%<sup>11</sup>. O teste de heterocedasticidade de White não encontra evidências de heterocedasticidade ao nível de significância de 5%<sup>11</sup> para todos os modelos, e finalmente o teste de Normalidade de Jarque-Bera não rejeita a hipótese nula de normalidade das séries ao nível de significância de 5%. Em resumo, todos os modelos mostraram ser corretamente especificados.

Por meio desses testes, infere-se que todos os modelos estimados captaram informações relevantes para garantir bons resultados de previsão do preço do café arábica.

<sup>11</sup> Foi considerada uma defasagem da ordem  $h = 15$ .

TABELA 3. ESTATÍSTICAS DO TESTE DE ESPECIFICAÇÃO DOS MODELOS 1997-2007

Modelo	White heteroscedasticidade	LM-test =nR <sup>2</sup> Autocorrelação	Jarque-Bera Normalidade
1	12.72558	18.96322	0,538403
2	8.575141	21.04119	0,864818
3	9.767362	17.59021	0,967921
4	11.50808	19.08901	1,057178
5	9.393401	19.81256	0,685092

### 3.6 Desempenhos de previsão

A escolha do modelo final baseia-se no critério de desempenho de previsão fora da amostra, e para tal são comparadas as previsões para dois e três anos à frente usando as observações que não foram utilizadas na fase de identificação, estimação e teste dos modelos<sup>12</sup>.

As Tabelas 4 e 5 mostram as estatísticas de desempenho de previsão dos modelos identificados. Adicionalmente, os resultados de cinco modelos univariados: AR(1), AR(2), ARMA(1,1), ARIMA(1,1,0) e ARIMA(1,1,1) são apresentados. Os modelos AR(1), AR(2), ARMA(1,1) foram estimados sobre a hipótese de estacionariedade das séries em nível, segundo os resultados do teste de raiz unitária com quebra estrutural. A especificação ARIMA seguiu os resultados da presença de raiz unitária na série mostrada pelas estatísticas ADF e PP. Embora esta especificação não seja adequada, seria a escolhida em caso de não identificar uma quebra nos preços, procedimento usualmente seguido em muitos trabalhos aplicados que omitem a importância de testar a presença de quebra estrutural nos dados.

Os modelos univariados foram selecionados segundo o procedimento de Box & Jenkins (1976), este procedimento é padrão na literatura e por isso mostramos apenas seus desempenhos de previsão.

<sup>12</sup> O período de identificação e teste corresponde ao período de 1997 até 2007 e o usado para o desempenho de previsão de 2008 e 2009. Para a estimativa dos valores previstos foi usada a opção *dynamic* do Eviews que computa previsão para mais de um passo à frente.

TABELA 4. MEDIDAS DE DESEMPENHO DA PREVISÃO FORA DA AMOSTRA PARA 2008 E 2009.

Modelo	RMSE	MAE	MAPE	THIEL's U
1	0,106922	0,091101	2,090988	0,012117
2	0,098389	0,084907	1,948231	0,011157
3	0,089397**	0,078234**	1,795403**	0,01145**
4	0,080279*	0,073423*	1,686080*	0,009111*
5	0,102765	0,087070	1,997746	0,011650
AR(1)	0,094735***	0,087747***	2,014574***	0,010735***
AR(2)	0,117658	0,110008	2,525892	0,013295
ARMA(1,1)	0,132208	0,124530	2,858869	0,014913
ARIMA(1,1,0)	0,097788	0,090331	2,074493	0,011075
ARIMA(1,1,1)	0,100670	0,093622	2,149837	0,011397

Notas: \*Melhor modelo, \*\* segundo melhor modelo, \*\*\* terceiro melhor modelo

Na Tabela 4 pode-se observar que o modelo 4 tem o melhor desempenho de previsão dentre os modelos estudados e dentre os modelos univariados o modelo com melhor desempenho foi o AR(1). A estatística MAPE mostra o erro percentual médio absoluto de previsão da ordem de 1,69% para o modelo 3 e de 2% para a especificação univariada o que representa valores da ordem de 11% superiores aos fornecidos pelo modelo 4. Os ganhos segundo a estatística THIEL's U são mais evidentes, o modelo 4 tem uma estatística de 0,0091 e o AR(1) de 0,0107 o que representa valores superiores em 17,58% ao do modelo 4.

A Tabela 5 mostra os resultados de previsão considerando três anos fora da amostra. Os modelos foram estimados com informação até 2006 deixando três anos para a comparação do desempenho de previsão<sup>13</sup>. Pode observar-se que o modelo 3 é a melhor especificação dentre os modelos de defasagem distribuída e dentre os modelos univariados o modelo ARIMA(1,1,1) foi o melhor. A estatística MAPE mostra o erro percentual médio absoluto de previsão da ordem de 1,72% para o modelo 3 e 3,54% para a especificação univariada. Isto representa erros percentuais médios de previsão superiores em 105% dados pela especificação univariada. A estatística THIEL's U mostra valores superiores para o modelo univariado em 45,34%.

13 Foram usados os mesmo modelos identificados na etapa anterior com informação até 2007.

TABELA 5. MEDIDAS DE DESEMPENHO DA PREVISÃO FORA DA AMOSTRA DE 2007 ATÉ 2009

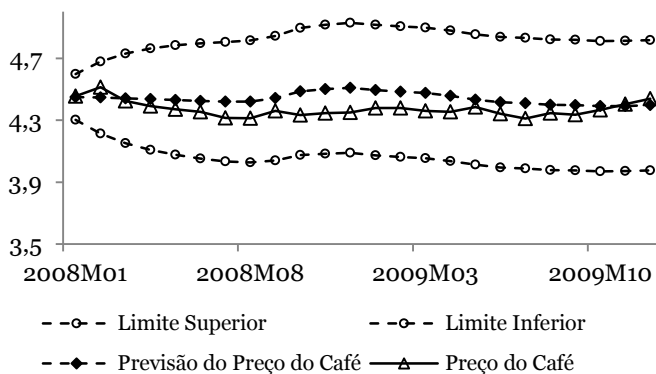
Modelo	RMSE	MAE	MAPE	THIEL's U
1	0,113442	0,098858	2,257913	0,012756
2	0,095900***	0,081579***	1,861832*** 0,010808***	
3	0,088001*	0,075520*	1,723474*	0,009925*
4	0,088254**	0,079382**	1,811575**	0,009945**
5	0,103725	0,088424	2,018132	0,011679
AR(1)	0,182595	0,175699	4,010149	0,020348
AR(2)	0,170464	0,164496	3,752210	0,019020
ARMA(1,1)	0,190095	0,183502	4,187698	0,021165
ARIMA(1,1,0)	0,262868	0,255066	5,820113	0,029037
ARIMA(1,1,1)	0,162597	0,155342	3,543873	0,018160

Notas: \*Melhor modelo, \*\* segundo melhor modelo, \*\*\* terceiro melhor modelo

Em resumo, as duas melhores especificações são os modelos 3 e 4 que dependem de defasagens da taxa de câmbio e preço do café. Na Figura 2 são apresentados os valores reais da série de preço do café, os valores previstos fora da amostra pelos modelos 3, 4 e a especificação univariada AR(1) para valores fora da amostra no período 2008 e 2009.

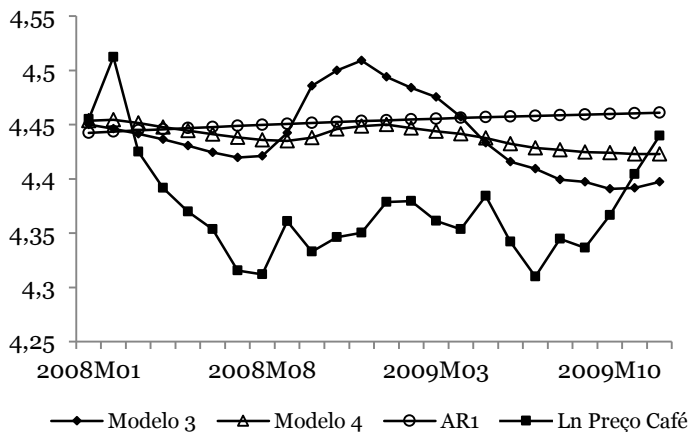
Na Figura 3 se mostra os valores observados da série de preços do café e os valores previstos pelo modelo 3 e seus intervalos de previsão. O intervalo de confiança dos valores previstos é calculado através do desvio padrão que é dado pela relação  $t_{\alpha/2}(n-k) \times SE$ , em que  $SE$  é o desvio padrão dos erros de previsão. Desta Figura pode observar-se que os valores previstos pelo modelo estão dentro do intervalo de confiança (ao nível de 95%) o que representa valores aceitáveis de previsão.

FIGURA 2. PREVISÃO DOS PREÇOS DE CAFÉ RECEBIDOS PELOS PRODUTORES - FORA DA AMOSTRA.



\* Preço do café arábica recebido pelos Produtores. Período Jan 2008- Dez 2009.

FIGURA 3. VALORES OBSERVADOS DO PREÇOS DE CAFÉ E PREÇOS PREVIS-  
TOS PELO MODELO



Período Janeiro 2008 a Dezembro 2009, com 3 com  $\pm 1,96 \times \text{sef}$ .

### 3.7 Previsão do modelo identificado

Finalmente, comparamos os resultados de previsão da melhor especificação com os dados efetivos realizados para o ano de 2010<sup>14</sup>. Para isto, estimamos o modelo 3 com toda a informação disponível até dezembro de 2009. O modelo estimado dado pela especificação 3 é mostrado a seguir:

$$\widehat{LnP_t} = 0.42 + 0.0024 \times DT_t + 0.084 \times \Delta txcam_t - 0.0024 \times \Delta txcam_{t-1} \\ (0.1322) \quad (0.0011) \quad (0.0582) \quad (0.0583) \\ 1.195 \times LnP_{t-1} - 0.290 \times LnP_{t-1} \\ (0.0769) \quad (0.0755)$$

Os valores em parênteses representam o erro padrão dos parâmetros estimados. Na Tabela 6 mostram-se os resultados de previsão para os casos estático e dinâmico. Os resultados de previsão dinâmica fornecem os valores previstos para os períodos  $t + 1, t + 2, \dots, t + 12$  utilizando informação do preço até o período  $t$ . Na previsão estática os valores dos preços a cada período são obtidos usando a informação efetiva do período anterior, ou seja, fornecem valores adequados um passo à frente. Por este motivo os erros cometidos pela previsão estática são menores, em particular mostramos a estatística MAPE para dois períodos. O MAPE\* é calculado para previsão seis passos à frente e o MAPE\*\* para previsão 12 passos à frente. Pode observar-se que o MAPE\* para o modelo estático é de 1,65% menor que 3,16% do dinâmico. A Figura 4, mostra os resultados de previsão dinâmica e estática para o ano 2010.

TABELA 6. PREVISÃO DO PREÇO DE CAFÉ DADO PELO MODELO 3 EM 2010. UNIDADES: R\$/SC 60KG

2010	Dados Efetivos	Estático	Dinâmico
Janeiro	83.99	85.62	85.62
Fevereiro	82.40	84.06	86.02
Março	81.93	81.61	85.45
Abril	82.02	81.67	84.81
Mai	83.15	82.46	84.79
Junho	87.16	83.43	84.57
Julho	85.99	87.62	84.04
	MAPE*	1,65%	3,16%
Agosto	88.60	85.29	83.70

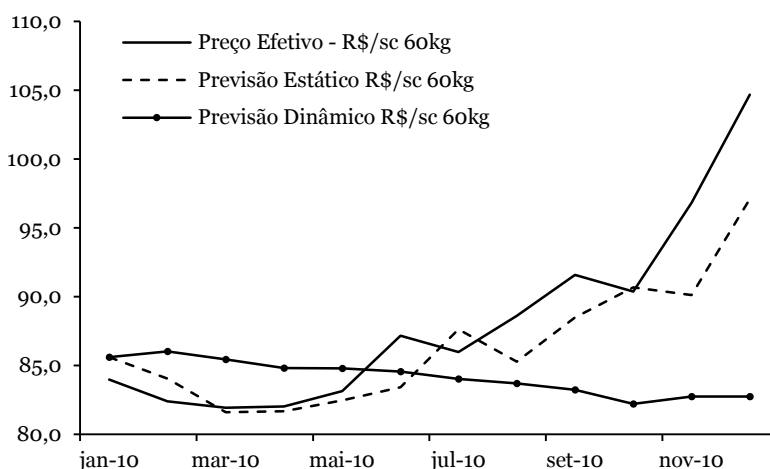
<sup>14</sup> Este trabalho foi desenvolvido em 2010 quando não se encontravam disponíveis as informações para o ano 2010. Na data de hoje pode ser utilizada esta informação para comparar os resultados de previsão do melhor modelo com os dados reais.

TABELA 6 (continuação). PREVISÃO DO PREÇO DE CAFÉ DADO PELO MODELO 3 EM 2010. UNIDADES: R\$/SC 60KG

Setembro	91.58	88.49	83.23
Outubro	90.37	90.67	82.21
Novembro	96.80	90.12	82.75
Dezembro	104.66	97.10	82.75
	<i>MAPE**</i>	2,78%	6,70%

Nota : MAPE \* é calculado para o seis primeiros meses e MAPE \*\* para o ano todo. Apenas em Janeiro de 2010 os valores de previsão para o estático e dinâmico coincidem.

FIGURA 4. COMPARAÇÃO DA PREVISÃO DO MODELO COM DADOS REAIS PARA O ANO 2010.



## Considerações Finais

Esta pesquisa objetiva encontrar um modelo econométrico de séries temporais, adequado à previsão dos preços de café a fim de auxiliar os produtores e comerciantes na tomada de decisões sobre os investimentos de café. Para atender a esse objetivo, foram estimados modelos dinâmicos de defasagem distribuída a fim de identificar o mais adequado à previsão de preço do café no Brasil. Os modelos dinâmicos de defasagem distribuída têm sido muito úteis para este fim do ponto de vista da facilidade na sua implementação e tem atendido os objetivos propostos. Adicionalmente foram estimados modelos univariados para comparar com os modelos inicialmente apresentados. A modelagem de séries temporais foi aplicada para a série de preços do café arábica recebido pelos produtores (café tipo 6 bebida dura).

Os resultados da modelagem identificaram uma quebra estrutural nos preços reais e, em particular, uma dependência com a taxa de câmbio. O modelo escolhido foi aquele que forneceu uma porcentagem de erro médio absoluto de previsão, MAPE, da ordem de 1,6%. Comparando os resultados de previsão deste modelo com dados reais (valores em nível) para o ano de 2010 observamos valores do MAPE de 1,6% para os primeiros seis meses e 2,78% para todo o período como observado na Tabela 6.

A proposta desta pesquisa é factível e as ferramentas econométricas usadas para prever o preço do café arábica demonstram ser preços. Outras pesquisas podem ser realizadas nesse campo considerando modelos econométricos mais sofisticados e explorando outras variáveis que expliquem o preço do café. Pode considerar-se endogeneidade nos preços do café através do uso de modelos VAR e a introdução de restrições de co-integração e quebra estrutural. Por outro lado, seria interessante comparar o desempenho de previsão dos modelos VAR com os propostos aqui e ver se o custo na modelagem compensa os benefícios em termos de ganhos na previsão.

## Referências

- ABIC. (2009) *Associação da Indústria de Café*, Estatísticas - média mensal dos preços recebidos pelos produtores.
- ALMEIDA Pinto, P. A. L. et al. (2008), *Aplicação do modelo arima à previsão do preço das commodities agrícolas brasileiras*, XLVI Congresso da Sociedade Brasileira de Economia, Administração e Sociologia Rural, Rio Branco - Acre.
- BACHA, C. J. C. (1998), A cafeicultura brasileira nas décadas de 80 e 90 e suas perspectivas, *Preços agrícolas*, Economia Cafeeira, ESALQ/USP - São Paulo.
- BAUM, 2004. "Stata: The language of choice for time series analysis?" Boston College Working Papers in Economics 598, Boston College Department of Economics. [Downloadable!] Published as: Christopher F Baum, 2005. "Stata: The language of choice for time-series analysis?" *Stata Journal*, StataCorp LP, vol. 5(1), pages 46-63, March.
- BOX, G. E. P. & JENKINS, G. M. (1976), *Time series analysis forecasting and control*, San Francisco: Holden Day. Edição revisada.
- BRESSAN, A. A. (2004), *Tomada de decisão em futuros agropecuários com modelos de previsão de séries temporais*, RAE eletrônica, Vol. 03, n. 1, art. 09.
- CLEMENTE, J., MONTANES, A., & REYES, M. (1998), *Testing for a unit root in variables with a double change in the mean*, *Economics Letters*, 59, 175--182.
- CNA - Confederação da Agricultura e Pecuária do Brasil (2008), Indicadores rurais, Agricultura gera retração dentro da porteira, Boletim mensal elaborado pela Superintendência da CNA em parceria com Centro de Inteligência em Mercados da Universidade Federal de Lavras. Lavras - MG, Ano XI n. 91.
- ENDERS, W. (2004), *Applied econometric time series*, Hoboken.
- GRANGER, C.W.J., & NEWBOLD, P. (1974), Spurious regressions in econometrics,

*Journal of Econometrics*, 2, p. 111-120.

IPEA. (2010), *Ipeadata* - base de dados. Inflação - IGP-M - mensal.

LORD, M. J. (1991), Price Formation in Commodity Markets, *Journal of Applied Econometrics*, Vol. 6, n. 3, p. 239-254.

MAPA, (2010) - Ministério da Agricultura e Pecuária, Estatísticas.

MARGARIDO, M. A. & BARROS, G. S. de C. (2000), *Transmissão de preços agrícolas internacionais para preços agrícolas domésticos no Brasil*, Instituto de Economia Agrícola, São Paulo, vol. 2, n. 47, p. 53-81.

MATIELLO, J. B. coord. (2005), *Cultura de café no Brasil: novo manual de recomendações*, Ed. 2005. Rio de Janeiro-RJ e Varginha-MG: Ministério da Agricultura, da Pecuária e do Abastecimento -- PROCAFÉ.

OTERO, J. G. & MILAS, C. (1998), Modelling the Behaviour of the spot prices of various types of coffee. The Warwick Economics Research Paper Series (TWERPS) 524, University of Warwick, Department of Economics.

PERRON, P. (1989), The great crash, the oil price shock, and the unit root hypothesis, *Econometrica*, 57, 1361--1401.

\_\_\_\_\_. (1997), Further Evidence on Breaking Trend Functions in Macroeconomic Variables, *Journal of Econometrics*, Vol. 80, p. 355-85.

PERRON, P. & VOGELSANG, T. J. (1992), Nonstationary and level shifts with an application to purchasing power parity, *Journal of Business and Economic Statistics*, 10, p. 301--320.

RIBEIRO, K. C. S.; SOUSA, A. F. & ROGERS, P. (2006), Preços do café no Brasil: variáveis preditivas no mercado à vista e futuro, *Revista de gestão USP*, São Paulo, vol. 13, n. 1, p. 11-30.

SHRESTHA, M. B. & CHOWDHURY, K. (2005), A sequential procedure for testing unit roots in the presence of structural break in time series data: an application to quarterly data of nepal, 1970-2003, *International Journal of Applied Econometrics and Quantitative Studies*, Vol.2-2, p. 1-16.

VOGELVANG, E. (1980), A quarterly econometric model for the price formation of coffee on the world market, *Free University Press*, Amsterdam.

\_\_\_\_\_. (1988), A Quarterly Econometric Model of the World Coffee, Working, H. Amsterdam (Holanda), *Economy*, Free University Press, 230 p.

\_\_\_\_\_. (1990) Serie research memoranda: Testing for co-integration with spot prices of some related agricultural commodities. Faculteit der Economische Wetenschappen en Econometrie Amsterdam. January.

\_\_\_\_\_. (1992). Hypothesis testing concerning relationships between spot prices of various types of coffee, *Journal of Applied Econometrics*, 7, p. 191-201.

ZIVOT, E. & ANDREWS, D. (1992), Further evidence on the great crash, the oil-price shock and the unit root hypothesis, *Journal of Business and Economic Statistics* 10, p. 251--270.

## Apêndice

TABELA A. ANÁLISE DE CORRELAÇÃO ENTRE AS VARIÁVEIS CONSIDERADAS

	Ln pibale	Ln pibbel	Ln pibbra	Ln pibita	Ln pibjap	Ln pibusa	Ln precoara	Ln txcam	Ln txjur
Lncredrur	1								
Lnpi bale	0,67	1							
Lnpi bbel	0,55	0,94	1						
Lnpi bbra	0,85	0,90	0,83	1					
Lnpi bita	0,50	0,95	0,93	0,80	1				
Lnpi bjap	0,75	0,89	0,78	0,88	0,87	1			
Lnpi busa	0,72	0,96	0,96	0,93	0,94	0,89	1		
Lnprecoara	-0,11	-0,65	-0,69	-0,50	-0,68	-0,37	-0,62	1	
Lntxcam	-0,13	0,46	0,65	0,28	0,61	0,25	0,52	-0,71	1
Lntxjur	0,06	-0,29	-0,33	-0,17	-0,30	-0,17	-0,22	0,24	-0,39

\* Lncredur é o crédito rural, Lnpi bale é o Pib da Alemanha, Lnpi bbel é Pib da Bélgica, Lnpi bbra é o Pib do Brasil, Lnpi bita é o Pib da Itália, Lnpi bjap é o Pib do Japão, Lnpi busa é o Pib dos Estados Unidos, Lnprecoara é o preço café arábica, Lntxcam é a taxa de câmbio e Lntxjur é a taxa de juro.

TABELA B. ANÁLISE DO CRITÉRIO DE INFORMAÇÃO E ESTATÍSTICAS DE ESTIMAÇÃO

Modelo	AIC	SC	F	R2-ajustado	DW
1	-2.298444	-2.144039*	369.2266	0.944833	1.982449
2	-2.303506	-2.171158	441.8319	0.944710	2.010558
3	-2.288864	-2.156516	435.0494	0.943895	1.982757
4	-2.282845*	-2.172555	535.9267	0.943139	1.969825
5	-2.301943	-2.147538	370.5926	0.945026	2.006490

\*Indica os menores valores de cada critério. A estatística F mostra que as variáveis explicativas em conjunto são significativas.

A estatística DW é próximo de 2 mostrando evidências de não autocorrelação de primeira ordem nos modelos selecionados. R2-ajustado indica um bom ajuste dos dados aos modelos candidatos.

TABELA C. MEDIDAS DE DESEMPENHO DA PREVISÃO

	RMSE	
Raíz do Erro Médio Quadrático		$\sqrt{\sum_{t=T+1}^{T+h} (\hat{y}_t - y_t)^2 / h}$
	MAE	
Erro Médio Absoluto		$\sum_{t=T+1}^{T+h}  \hat{y}_t - y_t  / h$
	MAPE	
Erro Médio Absoluto Percentual		$100 \times \sum_{t=T+1}^{T+h} \left  \frac{\hat{y}_t - y_t}{y_t} \right  / h$
	THEIL's U	
Coefficiente U de Theil		$\frac{\sqrt{\sum_{t=T+1}^{T+h} (\hat{y}_t - y_t)^2 / h}}{\sqrt{\sum_{t=T+1}^{T+h} \hat{y}_t^2 / h} + \sqrt{\sum_{t=T+1}^{T+h} y_t^2 / h}}$

Recebido em: 17 de agosto de 2012

Aceito em: 9 de outubro de 2012