

Textos para Discussão N° 78

Secretaria do Planejamento e Gestão
Fundação de Economia e Estatística Siegfried Emanuel Heuser

Estudo sobre a previsibilidade de preços no mercado *spot* de milho

Vancelei Zanin
Fábio Bandeira Guerra
Vitor Augusto Ozaki

Porto Alegre, agosto de 2010



GOVERNO DO ESTADO
RIO GRANDE DO SUL

SECRETARIA DO PLANEJAMENTO E GESTÃO

Secretário: José Alfredo Parode



DIRETORIA

Presidente: Adelar Fochezatto

Diretor Técnico: Octavio Augusto Camargo Conceição

Diretor Administrativo: Nóra Angela Gundlach Kraemer

CENTROS

Estudos Econômicos e Sociais: Sônia Rejane Unikowski Teruchkin

Pesquisa de Emprego e Desemprego: Roberto da Silva Wiltgen

Informações Estatísticas: Adalberto Alves Maia Neto

Informática: Luciano Zanuz

Editores: Valesca Casa Nova Nonnig

Recursos: Alfredo Crestani

TEXTOS PARA DISCUSSÃO

Publicação cujo objetivo é divulgar resultados de estudos direta ou indiretamente desenvolvidos pela FEE, os quais, por sua relevância, levam informações para profissionais especializados e estabelecem um espaço para sugestões. As opiniões emitidas nesta publicação são de exclusiva e de inteira responsabilidade do(s) autor(es), não exprimindo, necessariamente, o ponto de vista da Fundação de Economia e Estatística.

É permitida a reprodução deste texto e dos dados nele contidos, desde que citada a fonte. Reproduções para fins comerciais são proibidas.

Estudo sobre a previsibilidade de preços no mercado *spot* de milho

Vanclei Zanin¹
Fábio Bandeira Guerra²
Vitor Augusto Ozaki³

Resumo

O mercado do milho possui importante participação no cenário agropecuário nacional, dado, em grande parte, a sua relevância como insumo na indústria de carnes. Diante a essa situação, estratégias de comercialização de tal commodity ganham cada vez mais destaque, principalmente aquelas relacionadas à mitigação do risco de preço. Para tanto, prever preços tornou-se de suma importância. Nesse contexto, o presente trabalho busca avaliar dois diferentes “canais” de previsão de preços, um modelo econométrico de séries temporais (ARIMA) e o mercado futuro do milho. Os dados usados têm como fonte o CEPEA (Centro de Estudos Avançados em Economia Aplicada) e a BM&F (Bolsa de Mercadorias e Futuros), que disponibilizam, respectivamente, os preços à vista e futuro. Como resultado, observou-se que o modelo econométrico de previsão (dinâmico) obteve estimativas mais robustas em relação ao mercado futuro. Apesar das limitações que envolvem tal resultado, há uma evidência de que o ferramental econométrico pode ser bastante útil nas estratégias de gerenciamento de risco.

Palavras-chaves: ARIMA, Mercados Futuros, Previsão de preços, Milho.

Abstract

The corn market has an important share in the national agricultural scenario, mostly due to its relevance as an input in the meat industry. In this perspective, marketing strategies of such commodity have been gaining prominence, mainly those related to the mitigation of price risk. For such, forecasting prices has become extremely important. In this realm, the present research tries to evaluate two different price forecasting channels: a time series econometric model (ARIMA), and the corn future market. The source of data is the CEPEA (Center for Advanced Studies in Applied Economics) and the BM&F (Brazilian Mercatile and Futures Exchange), which respectively provide the spot and future prices. As a result, it was observed that the econometric model for forecasting (dynamic) obtained more robust estimates than the futures market. Despite the limitations involved in the results, there is evidence that the econometric tooling can be very useful in strategies for risk management.

KeyWords: ARIMA. Futures Market. Forecasting Prices. Corn.

Códigos JEL: C22, C53.

¹Economista, Fundação de Economia e Estatística (FEE/RS) Mestrando em Economia Aplicada Esalq/USP.

² Mestrando em Economia Aplicada Esalq/USP.

³ Professor do Departamento de Economia, Administração e Sociologia da Escola Superior de Agricultura “Luiz de Queiroz”.

1. INTRODUÇÃO

Em razão, principalmente, do acelerado crescimento econômico dos países asiáticos e do uso do milho para a produção de etanol nos Estados Unidos, a demanda mundial por essa *commodity* tem aumentado a cada ano. Não obstante essa realidade, o consumo interno também vem crescendo significativamente como consequência do fortalecimento da indústria de carnes, em especial, de aves e suínos. Tal crescimento é refletido na trajetória ascendente da produção nacional de milho, que pode ser vista no Gráfico 1, em que são apresentados os valores produzidos anualmente em milhões de toneladas, estimados pela Companhia Nacional de Abastecimento – CONAB, em 2009.

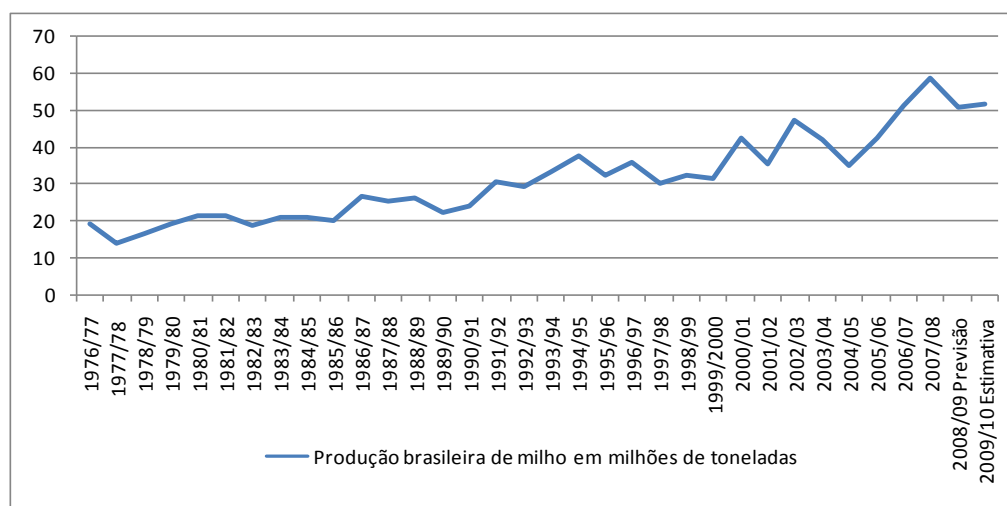


Gráfico 1 – Produção brasileira de milho, 1977 – 2009

Fonte: elaboração dos autores com dados da CONAB (2009)

Segundo dados divulgados pelo United States Department of Agriculture - USDA, em 2009, o Brasil foi o terceiro maior produtor mundial de milho, sendo superado apenas pelos EUA e China. Outro fator que ratifica a relevância do milho no agronegócio brasileiro é a sua fundamental participação nas cadeias de produção animal, onde é usado na formulação de rações, dado o seu elevado grau nutricional. Além disso, deve-se destacar a importância do milho na agricultura de subsistência, pois o seu cultivo também é realizado em pequenas propriedades.

Diante desse contexto, as estratégias de comercialização do milho, mais especificamente, as estratégias de gerenciamento de risco de preço são cada vez mais utilizadas e sofisticadas. Para tanto, muitas vezes, esforços são direcionados

para a obtenção de boas previsões em relação aos preços, pois a partir delas pode-se planejar, por exemplo, algum plano de estrutura comercial. No estudo de Bayer (2008) é colocado que a importância da precificação e previsão do preço das *commodities* decorre do fato de que, muitas empresas, para se protegerem das baixas ou aumentos dos preços, ou ainda, das variações cambiais, realizam operações financeiras, a prazo, logo a previsão do valor de uma *commodity* ou da sua variabilidade é de suma importância tanto para os produtores quanto para os compradores. Além disso, no caso do milho, que é uma *commodity* estocável, a obtenção de estimativas do preço futuro é bastante usada nos processos de “carregamento” da mercadoria, ou seja, nas tomadas de decisões relativas à estocagem da mesma, pois o milho somente será estocado caso haja expectativa de aumento nos preços.

Nesse sentido, dois importantes meios de previsão têm sido usados. O primeiro compreende os mercados futuros e o segundo os modelos econométricos de previsão. Conforme o estudo de Hull (2005), um mercado futuro pode ser definido como um mercado no qual são transacionados determinados produtos padronizados, em datas específicas, com um preço acordado para liquidação futura.

Segundo French (1986) os mercados futuros possuem duas grandes finalidades: sinalizar a expectativa futura para o preço *spot* (à vista) e oferecer ferramentas para mitigar os riscos (por meio de operações conhecidas como *hedging*). Com relação à primeira, muitos trabalhos trazem a idéia de que caso o mercado futuro em análise seja eficiente⁴, então o preço futuro negociado na bolsa poderá ser considerado um previsor não viesado do preço corrente.

Quanto aos modelos econométricos, para os autores Bressan e Lima (2002) uma das estratégias para reduzir a incerteza no processo de tomada de decisões econômicas é a utilização de modelos de previsão de séries temporais univariadas. Tais modelos são construídos a partir de processos estocásticos especiais, que buscam estimar o valor futuro da variável em questão com base somente na estrutura de correlação de seus valores passados.

¹ Mercados eficientes são definidos em Fama (1970) como aqueles em que os preços sempre refletem completamente as informações disponíveis, em um mercado caracterizado por um elevado número de indivíduos bem informados cujas decisões de compra e venda fazem com que novas informações sejam rapidamente absorvidas pelo preço de mercado.

Dada a existência desses dois canais usados para observar a expectativa dos preços, surge na literatura acadêmica certa discussão quanto à qualidade das “previsões” extraídas por cada um deles. Tomek (1997), por exemplo, argumenta que modelos econométricos de domínio público não podem ter um desempenho de previsão melhor do que os mercados futuros eficientes.

Embasado nesse debate, o presente trabalho tem como objetivo comparar a qualidade das previsões obtidas por meio dos mercados futuros e por meio de técnicas econométricas. Para tanto, foi utilizada uma série de preços *spot* diários do milho negociado em Campinas-SP no período entre 5 de janeiro de 2004 a 27 de fevereiro de 2009, a partir da qual um modelo econométrico do tipo ARIMA (Auto-Regressivos Integrados de Médias Móveis) foi construído para realizar as previsões de preços. Além disso, também foi usada uma série de preços futuros do milho, negociados na BM&F (Bolsa de Mercadorias e Futuros), com a qual foi calculado o “desvio” em relação aos preços físicos correspondentes. O critério usado para analisar a acurácia das previsões foi o “erro quadrático médio” (EQM).

O artigo está dividido, além dessa introdução, em uma primeira seção que contém uma breve fundamentação teórica sobre mercados eficientes e previsão de preços, em especial, no mercado de *commodities* agrícolas. Na segunda seção a metodologia econométrica é apresentada. Na terceira seção os dados usados na análise empírica são descritos. Em seguida são colocados os resultados obtidos pelos procedimentos econométricos, que são interpretados e brevemente discutidos. Por fim, algumas conclusões são tecidas.

2. MERCADOS EFICIENTES E PREVISÃO DE PREÇOS – TEORIA E EVIDÊNCIAS

Segundo Marques e Mello (1999), o risco na atividade agropecuária difere de outras atividades produtivas, principalmente em função da dependência dos fatores climáticos, do elevado tempo em que algumas culturas permanecem no campo sem propiciar retorno do investimento realizado, da dificuldade de negociação em face de os produtos serem perecíveis, bem como da volatilidade dos preços.

Diante a tal situação, a previsão de preços é de suma importância para o indivíduo tomador de decisão, sendo tanto maior quanto maior for seu grau de precisão. Portanto, o modelo de previsão se torna uma ferramenta de grande valia

para o agente que se defronta com decisões que envolvem riscos, notadamente na agricultura.

Com relação à capacidade de previsão, Leuthold et al. (1989) colocam que um mercado futuro eficiente é aquele que promove a mais acurada representação dos preços *spot* subseqüentes. Os autores ainda afirmam que se o modelo de previsão for mais robusto que o mercado futuro tem-se uma evidência de que o mercado não está incorporando todas as informações existentes e, portanto, é ineficiente.

O conceito de eficiência de mercado é trabalhado de maneira bastante rica em Fama (1970), onde há a caracterização de três tipos de eficiência. O primeiro é o conceito de eficiência fraca, descrito pelo autor como uma situação em que todas as informações que podem ser extraídas dos preços passados já estão incorporadas no preço à vista, por essa razão, informações precedentes não são úteis para realizar previsão de preços futuros. O segundo tipo é a eficiência semiforte, em que, além das informações dos preços passados, todas as informações públicas estão incorporadas no preço *spot*. Por fim, é colocado o conceito de eficiência forte, no qual todas as informações disponíveis estão incorporadas nos preços físicos, inclusive as possuídas pelos *insiders*.

Assim como em trabalhos já mencionados, Tomek (1997) coloca que a leitura dos preços futuros como previsores dos preços físicos esperados é diretamente relacionada com a hipótese de eficiência de mercado. Para o autor uma previsão de preços é condicionada às informações disponíveis no tempo em que a previsão é feita, e o conceito de eficiência de preços é relacionado com a questão de quão bem o mercado usa as informações disponíveis para determinar os preços. No mesmo estudo ele acrescenta que se o modelo econométrico representa razoavelmente o mercado da *commodity* e esse mercado é eficiente, então os preços correntes refletem toda informação disponível, inclusive as expectativas dos agentes, e o modelo não conseguirá melhorar a previsão de preço dada pelo mercado.

Ainda com relação ao artigo comentado acima, destacam-se três possíveis interpretações para o caso em que modelo econométrico obtenha um desempenho de previsão melhor do que o mercado: (a) os dois estão corretos, e o problema está no tamanho da amostra que é demasiadamente pequena, (b) um dos dois está

errado, se o mercado for melhor do que o modelo, então o modelo é erroneamente especificado, se o modelo é melhor do que o do mercado, então o mercado é ineficiente, entretanto é improvável que o modelo econométrico possa ter um desempenho continuamente melhor que do mercado e (c) o mercado é fracamente eficiente e não fortemente eficiente. Além disso, é ressaltado que a habilidade de previsão varia entre os diferentes mercados, onde mercados com grandes estoques contínuos, como o milho, são mais propensos em prover previsões mais acuradas. O autor finaliza dizendo que essas evidências não significam que a análise econométrica não tem valor, colocando que previsões que misturam os modelos econométricos com as informações trazidas pelos mercados futuro podem promover previsões com menor variância de erro.

Como exemplo do uso tanto dos mercados futuros quanto dos modelos econométricos na busca por boas previsões dos preços *spot*, tem-se Springgs (1981), em que é feita uma previsão para os preços mensais de milho referentes à Indiana (EUA) com o uso de dois modelos, um deles é um modelo ARIMA e o outro é um modelo que envolve a relação entre os preços mensais com o mercado futuro de milho. A amostra utilizada compreende o período entre janeiro de 1959 e agosto de 1978, usada como base para realizar a previsão de quatro meses à frente, ou seja, estimar os preços de setembro, outubro, novembro e dezembro. Como critério de desempenho foi usado o erro quadrático mínimo, e assim, o modelo ARIMA se mostrou mais eficiente para a previsão dos preços de milho para o mês de setembro, no entanto o modelo que utiliza os preços negociados nos mercados futuros se mostrou melhor para prever os preços mensais dos meses de outubro, novembro e dezembro. O autor então conclui que o modelo que usa os preços futuros, na média, foi mais eficiente para se fazer previsão dos preços mensais em Indiana para a safra do próximo ano.

No Brasil também existem diversos estudos voltados para a previsão de preços das *commodities* agropecuárias, o que reforça a relevância desse instrumental. Pinto *et al.* (2008) se propõem a analisar o comportamento dos preços *spot* das seguintes *commodities*: cacau, café, cana-de-açúcar, laranja e soja. Para atingir tal objetivo, realizaram previsão para os preços dos produtos observados por meio da metodologia ARIMA. Os autores colocam que os resultados encontrados

podem ser muito úteis na tomada de decisão, pois eles sinalizam a tendência dos preços para um horizonte de curto prazo.

Medeiros *et al.* (2006) dedicam-se ao estudo da redução do risco na comercialização do boi por meio da previsão dos preços a serem recebidos no mercado físico brasileiro. Para tanto foi usada a modelagem de séries temporais ARIMA. Os resultados obtidos foram considerados significativos e próximos aos valores de mercado, sendo favorável ao uso do modelo ARIMA como método de previsão.

Seguindo a mesma linha de estudo dos trabalhos abordados nessa seção, o presente artigo busca avaliar a qualidade da previsão no mercado brasileiro de milho. No entanto, o enfoque foi dado na diferença de desempenho das previsões advindas dos procedimentos econométricos em relação àquelas obtidas a partir do mercado futuro do milho.

3. METODOLOGIA⁵

3.1 Modelos ARIMA

Uma família de modelos muito utilizada na análise de séries temporais é a auto-regressiva integrada de média móvel, ARIMA, baseada na metodologia de Box e Jenkins (1970). Esses modelos estatísticos são ajustados aos dados, tendo como idéia subjacente que a série temporal é um processo estocástico, cuja natureza poderá ser representada a partir de um modelo matemático. Morettin e Tolo (2004) colocam a construção do modelo como um ciclo iterativo que segue a especificação, identificação, estimação e verificação da adequação do(s) modelo(s) aos dados.

A notação utilizada para designar esses modelos é ARIMA (p,d,q), em que p é a ordem dos termos da parte auto-regressiva (AR); d, o número de diferenciações necessárias para a série tornar-se estacionária, e, q a ordem dos termos da parte de médias móveis (MA). Cabe ressaltar que p, d, q são todos inteiros maiores ou iguais a zero.

O modelo ARIMA pode ser completo, ou seja, com parte AR(p) e MA(q) modelados depois que a série foi diferenciada d vezes para se tornar estacionária.

⁵ Esta seção é baseada em Morettin e Tolo (2004), Enders (2004) e Brooks (2008).

Ele pode também ser incompleto, ou seja, depois de a série ser diferenciada d vezes para se tornar estacionária, a melhor adequação para o modelo consiste somente em sua parte auto-regressiva AR (p) ou em sua parte de médias móveis MA (q).

Ainda de acordo com Morettin e Tolo (2004), um processo X para ser estacionário deve se desenvolver aleatoriamente no tempo em torno de uma média, de modo que a escolha de uma origem dos tempos não é importante. No contexto de séries temporais a estacionariedade relevante é a de segunda ordem, ou também, chamada fracamente estacionária, que é definida como:

Um processo estocástico $X=\{X(t), t \in T\}$ diz-se fracamente estacionário se e somente se:

$$E\{X(t)\} = \mu(t) = \mu, \text{ constante, para todo } t \in T.$$

$$E\{X^2(t)\} < \infty, \text{ para todo } t \in T.$$

$$\gamma(t_1, t_2) = \text{Cov}\{X(t_1), X(t_2)\} \text{ é uma função de } |t_1 - t_2|.$$

3.1.1 Modelo AR (p)

O modelo auto-regressivo é adequado para representar a série quando apenas as observações passadas da variável em questão são necessárias para explicar o comportamento da série “ x ”. Um modelo AR (p) obtido quando $d=q=0$ é dado por (1):

$$x_t = \theta_0 + \sum_i^p \phi_i x_{t-i} + a_t, \quad (1)$$

onde, $\theta_0, \phi_1 \dots \phi_p$ são parâmetros do modelo e $a_t \sim \text{RB}(0, \sigma^2)$, ou seja, o termo de erro é um ruído branco com média zero e variância constante.

O modelo AR (p), de (1) pode ser escrito na forma de operadores de diferenças, como em (2):

$$(1 - \phi_1 B - \phi_2 B^2 \dots \phi_p B^p)x_t = \theta_0 + a_t \quad (2)$$

Como AR(p) é sempre invertível, a preocupação é checar se o processo gerador da série é estacionário, para tanto, é necessário que $|B| \leq 1$. Condição equivalente a dizer que a raiz da equação característica, $\phi(B) = 0$, deve estar fora do círculo unitário.

3.1.2 Modelo MA(q)

Quando a série “x” é modelada apenas pelo processo de médias móveis, $d=p=0$, o modelo fica na forma (3):

$$x_t = \phi_0 + a_t - \theta_1 a_{t-1} - \dots - \theta_q a_{t-q} \quad (3)$$

onde, $\phi_0, \theta_1, \theta_2 \dots \theta_q$ são os parâmetros da série e $a_t \sim \text{RB}(0, \sigma^2)$.

Em termos de operador de defasagem, o modelo MA(q) pode ser escrito tal como (4):

$$(1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q) a_t = \theta(B) a_t = x_t - \phi_0 \quad (4)$$

O modelo MA(q) é uma combinação linear das q defasagens dos resíduos do modelo (a_t), que por definição são ruídos brancos (RB). Pode-se observar que o modelo MA (1) é sempre estacionário. Para que seja invertível é necessário que $|B| \leq 1$, que é equivalente a ter as raízes da equação característica, $\theta(B) = 1 - \theta_1 B - \dots - \theta_q B^q = 0$, fora do círculo unitário.

3.1.3 Modelo ARMA (p, q)

Morettin e Toloi (2004) afirmam que um modelo parcimonioso em número de parâmetros deve ser sempre buscado pelo pesquisador, e para tanto é interessante a modelagem da série com termos auto-regressivos e de médias móveis, ou seja, um modelo ARMA (p, q) que pode ser expresso como em (5):

$$x_t = \phi_0 + \sum_i^p \phi_i x_{t-i} + a_t - \sum_i^q \theta_i a_{t-i} \quad (5)$$

onde $\phi_0, \dots, \phi_p, \theta_1, \dots, \theta_q$ são parâmetros reais e $a_t \sim \text{RB}(0, \sigma^2)$.

Em termos de operadores de defasagem o modelo descrito em (5) fica da forma (6) como segue:

$$(1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p) x_t = \phi_0 + (1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q) a_t \quad (6)$$

O modelo ARMA (p,q) segue as mesmas condições de estacionariedade do processo AR (p), e, de invertibilidade do processo MA (q). Equivalentemente a dizer que as raízes características para os operadores de defasagem devem estar fora do círculo unitário.

3.1.4 Modelo ARIMA (p, d, q)

Os modelos definidos anteriormente são adequados para séries temporais estáveis ao longo do tempo, ou seja, com propriedades de um processo estacionário. Entretanto muitas séries não apresentam essa característica; para torná-las estacionárias, o procedimento padrão é sua diferenciação. Tomando-se um número finito de diferenças d a série passa a se desenvolver no tempo ao redor de uma média constante, tornando-se estacionária. Esse processo de diferenciação garante que as condições de estacionariedade e invertibilidade anteriormente colocadas sejam atendidas. O modelo ARIMA (p,d,q), descrito em (7), onde $\Delta^d x_t$ representa a diferenciação que torna a série estacionária, é equivalente ao modelo ARMA(p,q).

$$\phi(B)\Delta^d x_t = \theta(B)a_t \quad (7)$$

onde B é operador de defasagens, d é o número de diferenças para estacionarizar a série e $a_t \sim RB(0, \sigma^2)$.

Para Morettin e Tolo (2004) a fase crítica da modelagem ARIMA é sua identificação. Aqui o procedimento “chave” está na modelagem adequada das partes AR, MA e na estacionariedade da série, identificáveis por meio das funções de autocorrelação (FAC) e autocorrelação parcial (FACP).

3.2 Teste de raiz unitária

Enders (2004) coloca que a inspeção da função de autocorrelação serve como um indicador da presença de tendência na série. O lento decaimento da função de autocorrelação (FAC) é uma evidência da existência de raiz unitária, ou da não estacionariedade da série. Um teste formal muito utilizado para detectar a presença de raiz unitária na série é o teste de Dickey-Fuller (DF) de 1976, gerado para três diferentes tipos de regressão, (8), (9), (10), apresentadas a seguir:

$$\Delta x_t = \gamma x_{t-1} + \varepsilon_t \quad (8)$$

$$\Delta x_t = a_0 + \gamma x_{t-1} + \varepsilon_t \quad (9)$$

$$\Delta x_t = a_0 + \gamma x_{t-1} + a_2 t + \varepsilon_t \quad (10)$$

Onde $\Delta x_t = x_t - x_{t-1}$, ε_t é um ruído branco com média zero e variância constante.

Em (8) é representado um passeio aleatório puro, em (9) soma-se um intercepto ou uma *drift*, já em (10) além da *drift* acrescenta-se uma tendência linear ao passeio aleatório. O procedimento adequado é comparar a estatística t dos coeficientes da forma apropriada (passeio aleatório puro, com *drift*, com tendência, ou com *drift* e tendência) com o valor tabulado pela estatística τ de Dickey-Fuller, para se examinar a hipótese nula: presença de raiz unitária, contra a hipótese alternativa: série estacionária. A evidência estatística contra a hipótese nula, ou seja, em favor da estacionariedade da série é obtida se a estatística t for mais negativa do que o valor crítico tabelado por Dickey-Fuller.

Para o caso em que as variáveis dependentes são autocorrelacionadas foi proposto o teste de Dickey-Fuller aumentado (ADF), de 1979, que leva em consideração p defasagens da variável dependente até que esse problema seja solucionado. Nesse contexto a regressão fica na forma (11):

$$\Delta x_t = a_0 + \gamma x_{t-1} + \sum_2^p \beta_i \Delta x_{t-i+1} + \varepsilon_t \quad (11)$$

O teste é realizado da mesma maneira que os anteriormente citados, sendo os valores das estatísticas t negativamente maiores que os valores calculados τ de Dickey-Fuller uma evidência contra a existência de raiz unitária na série.

3.3 Previsão com modelos ARIMA

Brooks (2008) mostra que a previsão em um modelo ARIMA (p,d,q) já estacionarizado nada mais é que a esperança condicional da série modelada.

Supondo que $f_{t,s}$ denota a previsão no tempo t para s passos à frente da série x. A função de previsão será da forma (12):

$$f_{t,s} = \sum_{i=1}^p \phi_i f_{t,s-i} + \sum_{j=1}^q \theta_j u_{t+s-j} \quad (12)$$

onde $f_{t,s} = x_{t+s}, s \leq 0$; $u_{t+s} = 0, s > 0 = u_{t+s}, s \leq 0$ e ϕ_i e θ_i são os coeficientes da parte auto-regressiva e de médias móveis respectivamente.

Para parte de médias móveis a previsão tem memória somente até a defasagem q, uma vez que a esperança condicional do resíduo para períodos futuros é zero. Segue abaixo o exemplo em (13) para um processo MA(2):

$$x_t = \theta_0 + \theta_1 u_{t-1} + \theta_2 u_{t-2} + u_t \quad (13)$$

O modelo para o período t+1 é representado em (14):

$$x_{t+1} = \theta_0 + \theta_1 u_t + \theta_2 u_{t-1} + u_{t+1} \quad (14)$$

Quando se faz a previsão para esse período posterior tem-se (15):

$$f_{t,1} = E(x_{t+1/t}) = E(\theta_0 + \theta_1 u_t + \theta_2 u_{t-1} + u_{t+1} / \Omega_t) \quad (15)$$

onde θ_0 é a média incondicional e Ω_t é a informação disponível no instante t. Como a $E(u_{t+1}) = 0$ no instante t, a previsão de x um período à frente, feita no instante t é uma combinação linear dos termos de erro, conforme (16):

$$f_{t,1} = E(x_{t+1/t}) = \theta_0 + \theta_1 u_t + \theta_2 u_{t-1} \quad (16)$$

Já o modelo auto-regressivo tem memória infinita, (17) exemplifica um caso AR (2):

$$x_t = \phi_0 + \phi_1 x_{t-1} + \phi_2 x_{t-2} + u_t \quad (17)$$

Para um período à frente o modelo pode ser escrito na forma (18):

$$x_{t+1} = \phi_0 + \phi_1 x_t + \phi_2 x_{t-1} + u_{t+1} \quad (18)$$

A previsão para esse modelo apresentado em (18) é dada em (19):

$$f_{t,1} = E(x_{t+1/t}) = E(\phi_0 + \phi_1 x_t + \phi_2 x_{t-1} + u_{t+1} / \Omega_t) \quad (19)$$

Como pode ser observado a previsão é a soma do intercepto com resultado da multiplicação do coeficiente ϕ_1 pelo valor da série no período 1 e da multiplicação do coeficiente ϕ_2 com valor da série no período t-1, conforme (20):

$$f_{t,1} = E(x_{t+1/t}) = \phi_0 + \phi_1 x_t + \phi_2 x_{t-1} \quad (20)$$

A obtenção da fórmula para a previsão do ARIMA (p,d,q) já estacionarizado é a generalização das partes AR(p) e MA(q) feitas acima.

Descrito o método de previsão a partir de um modelo ARIMA, o próximo ponto a ser discutido é acurácia dessa previsão. Nesse sentido, Brooks (2008) identifica que para uma averiguação do ajuste do modelo é interessante utilizar-se de uma previsão *out-of-sample*, que consiste na divisão da amostra em duas partes, a primeira (*in sample*) para se estimar o modelo, com o qual se faz a previsão, e, então, comparam-se os resultados do modelo estimado com os valores reais que foram separados na segunda parte da amostra (*out of sample*).

Em relação à qualidade da previsão Brooks (2008) afirma que esta é afetada pelo número de períodos à frente em que é feita, quanto maior for considerado o período à frente, maior será a variância entre o valor estimado e o observado. O autor sugere como forma de se contornar o problema, fazer uma previsão recursiva, onde cada previsão é feita para um período à frente, incorporando-se o verdadeiro valor observado na previsão seguinte⁶. Para finalizar, ele sugere a utilização do erro quadrático médio (EQM) entre o valor previsto e o valor observado *out of sample* como uma forma de averiguação da adequação do modelo proposto.

4. DADOS

4.1 Descrição dos dados

Os dados utilizados para a realização desse estudo correspondem aos preços futuro e *spot* do milho para o mercado nacional. O primeiro tem como fonte a Bolsa de Mercadorias e Futuros (BM&F), onde são negociados os contratos futuros da *commodity* sob análise⁷. É importante ressaltar que para a construção da série de preço futuro foi considerado o preço referente ao contrato com vencimento mais próximo⁸. Já o preço à vista foi obtido junto ao Centro de Estudos Avançados em

⁶ Cabe ressaltar que foram feitos dois tipos de previsão (dinâmica e estática) para um horizonte temporal de dez dias.

⁷ As especificações do contrato futuro do milho podem ser encontradas no site da BM&F, <www.bmf.com.br>.

⁸ Além disso, também deve ser lembrado que em razão da mudança do contrato futuro do milho com entrega física para o contrato com liquidação financeira (criado em setembro de 2008), houve a necessidade de realizar uma análise de transferência de liquidez entre esses contratos para se estabelecer o instante em que o preço futuro observado passaria do contrato futuro com entrega física para o contrato com liquidação financeira. Dessa forma foi estabelecido que o preço futuro de 5 de janeiro de 2004 a 15 de janeiro de 2009 seria aquele referente ao contrato com entrega física, e no

Economia Aplicada (CEPEA) da ESALQ-USP, que calcula o preço físico do milho para cinco “praças”, dentre elas Campinas-SP, que foi a escolhida no presente trabalho, já que é este o local de entrega especificado no contrato futuro do milho negociado na BM&F⁹. Ainda, vale a pena destacar que tanto o preço futuro quanto o *spot* são diários, expressos em Reais (R\$) e referem-se a sacas de sessenta quilos.

O período escolhido para se trabalhar com ambos os preços foi de 5 de janeiro de 2004 a 13 de março de 2009, o que contabiliza 1281 observações para cada série. No entanto, a amostra foi segmentada em duas partes, uma parte relativa ao período entre 5 de janeiro de 2004 e 27 de fevereiro de 2009 (*in the sample*), no qual as previsões do modelo econométrico basearam-se. A outra parte da amostra corresponde ao período de 2 de março de 2009 a 13 de março de 2009 (*out of the sample*), que são os dias para os quais serão feitas as previsões do preço físico do milho, ou seja, uma horizonte de dez dias à frente, e as comparações do preço futuro negociado na bolsa com o preço físico negociado em Campinas-SP. A escolha do período para essa divisão da amostra deve-se ao fato de que março é um mês importante tanto para a colheita do milho safra verão quanto para as decisões do plantio do milho “safrinha”, período em que a previsão dos preços é bastante importante para que os participantes do mercado do milho possam traçar suas estratégias de comercialização.

4.2 Análise preliminar dos dados

Aqui é feita uma breve e superficial análise das variáveis estudadas. Para tanto, o passo inicial é avaliar as trajetórias dos preços futuro e *spot* do milho, a partir das quais é possível identificar, ainda que primariamente, o nível de relação entre tais variáveis. O gráfico a seguir mostra que as séries observadas parecem estar fortemente relacionadas, constatação que é ratificada pelo alto grau de correlação igual a 0,95. Esse resultado pode ser uma indicação de que o preço futuro do milho

período entre 16 de janeiro de 2009 e 13 de março de 2009 a referência seria o contrato com liquidação financeira.

⁹ Pereira (2006) coloca que a formação dos preços à vista das *commodities* depende da localidade onde estas são negociadas, sendo necessário incorporar os custos de frete e armazenagem. Por isso o preço físico de Campinas-SP é o mais adequado, uma vez que essa é a praça formadora de preços, logo com a menor diferença de base, ou seja, com preço mais “próximo” ao preço futuro negociado na BM&F.

(pfm) pode ser interpretado como um estimador não-tendencioso do preço à vista (psm), o que seria um ponto favorável ao uso do mercado futuro como previsor do mercado físico.

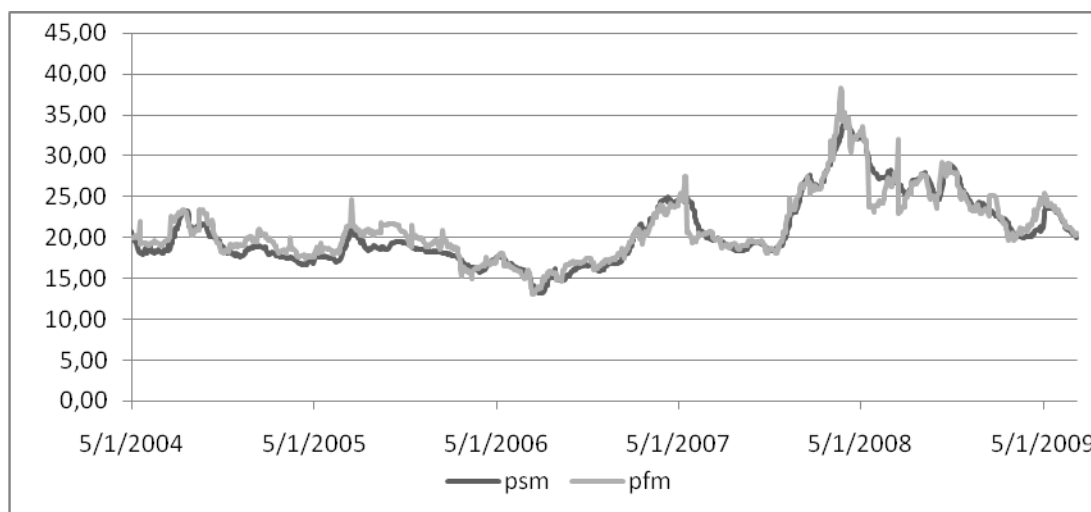


Gráfico 2 – Preços futuro e *spot* do milho, 05/01/2004 – 13/03/2009

Fonte: elaboração dos autores com dados da BM&F e do CEPEA-Esalq.

A segunda etapa dessa análise consiste na apresentação de um resumo descritivo das variáveis de interesse nesse estudo. Isso é feito na Tabela 1, apresentada logo abaixo, em que é possível notar um comportamento semelhante entre as duas séries.

Tabela 1
Análise descritiva das séries de preços

	Média	Desvio padrão	Variância	Máximo	Mínimo
PSM	20,73	4,23	17,93	34,14	13,17
PFM	21,11	4,03	16,23	38,32	13,03

Fonte: elaboração dos autores.

5. RESULTADOS

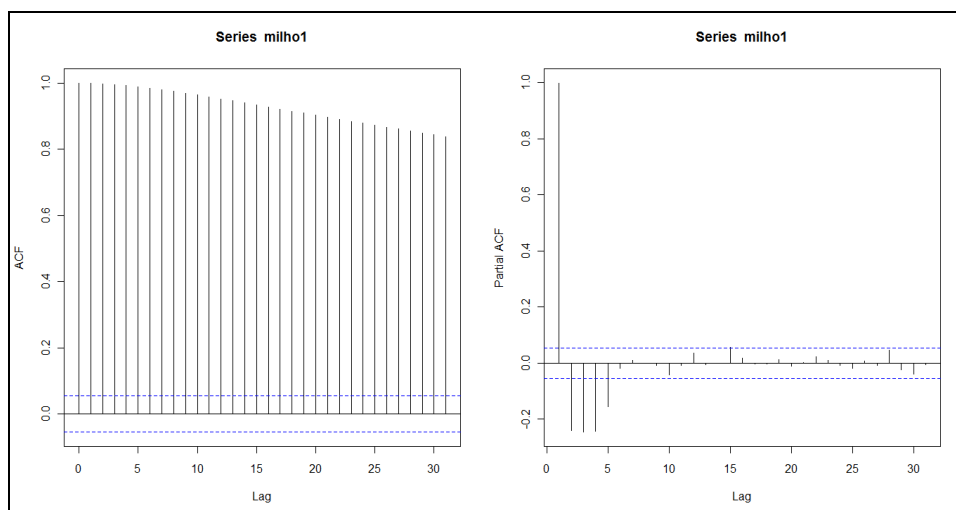
Nessa seção são apresentados os principais resultados obtidos a partir da metodologia econométrica descrita anteriormente. A ordem de apresentação dos procedimentos aplicados é dada da seguinte forma: análise de estacionariedade da série, identificação e estimação do modelo e, por fim, a previsão e avaliação da qualidade da previsão.

5.1 Análise de estacionariedade

A análise de estacionariedade é realizada em duas etapas. A primeira corresponde à interpretação das funções de autocorrelação (FAC) e autocorrelação parcial (FACP), já a segunda baseia-se no teste de raiz unitária (ADF).

A Figura 1 traz os valores da FAC e da FACP. Ao observá-las é possível notar que a série de preços *spot* do milho, provavelmente, é não estacionária, pois a FAC não declina conforme o comprimento dos *lags* aumenta, ou seja, não ocorre o “decaimento exponencial” característico de uma série estacionária.

Figura 1
FACP da série em nível



Fonte: elaboração dos autores.

Para buscar uma evidência com maior robustez quanto à estacionariedade da série, foi realizado o teste ADF. Para tanto, foi estimada uma regressão, considerada a mais adequada, em que o preço do milho é explicado por ele mesmo, defasado em quatro períodos além de uma constante (*drift*). O valor crítico tabelado por Dickey-Fuller, ao nível de 5% de significância, é igual a -2,86, inferior ao valor calculado da estatística, -2,09, portanto não se rejeita a hipótese nula de que a série em análise possui raiz unitária. Diante desse resultado, tornou-se necessário tomar a primeira

diferença da série para estacionarizá-la. Além disso, optou-se por trabalhar com os preços em logaritmo¹⁰, o que estabilizou a variância.

Com a série na primeira diferença, novamente foram observadas as funções de autocorrelação e autocorrelação parcial, que dessa vez indicaram a estacionariedade da série. Esse comportamento foi ratificado pelo teste ADF, já que o valor crítico calculado (-9,55) foi inferior ao valor tabelado (-2,86), o que rejeita a hipótese nula ao nível de 5% de significância.

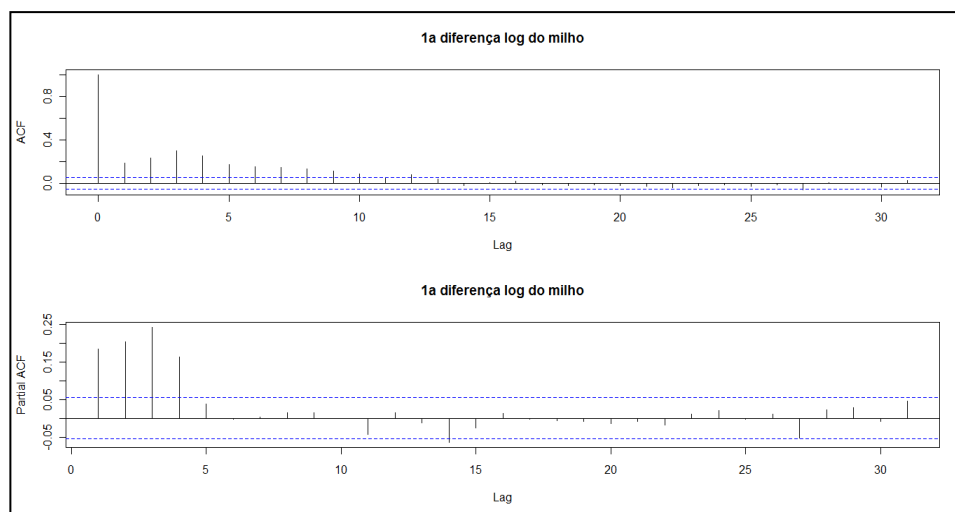
5.2 Identificação e estimação do modelo

Com a série estacionarizada, o passo a seguir é identificar o modelo que a descreve da melhor forma, através da metodologia de Box e Jenkins (1970). Para tanto são analisadas a FAC e a FACP, por meio das quais é possível identificar uma primeira sugestão do modelo. A Figura 2 fornece indícios de que o processo gerador da série observada é auto-regressivo de ordem quatro, ou seja, um ARIMA (4, 1, 0), já que a série original é integrada de ordem um.

Para verificar se o modelo ARIMA (4, 1, 0) realmente é o mais adequado foi preciso estimá-lo e, em seguida, avaliar simultaneamente três características subjacentes aos seus resultados, sendo elas: significância estatística dos coeficientes das últimas defasagens, critério de informação (Akaike - AIC) e o correlograma dos resíduos. O modelo escolhido foi aquele que apresentou a seguinte combinação: coeficientes estatisticamente significantes, resíduo com comportamento igual a um ruído branco e o menor critério de informação. Além disso, é importante destacar a preferência por modelos parcimoniosos, já enfatizada pela literatura especializada.

¹⁰ Deve-se ressaltar que a diferenciação dos preços em logaritmo corresponde aos retornos dos preços à vista.

Figura 2
FAC e FACP da série na primeira diferença



Fonte: elaboração dos autores.

Após a estimação de alguns modelos, chegou-se à conclusão de que o ARIMA (4, 1, 0), considerado uma primeira sugestão, de fato é o mais adequado para descrever o processo gerador dos preços *spot* do milho. Isso porque foi esse o modelo que combinou da melhor forma as características de seleção descritas logo acima, levando-se em consideração, principalmente, o bom comportamento dos seus resíduos (ruído branco, indicado pelo teste de adequação de Box-Pierce, em que se rejeita a hipótese nula de que todos os coeficientes de autocorrelação das trinta primeiras defasagens são iguais a zero) e a parcimônia quanto ao número de parâmetros a serem estimados. A Tabela 2, que vem na seqüência, expõe os valores do critério de Akaike calculado para os modelos testados que merecem destaque. Ao observá-la é possível notar que o modelo escolhido possui o menor critério de informação em relação aos demais, o que corrobora a sua escolha.

Tabela 2
Critério de Akaike – AIC

	Modelo			
	ARIMA (4,1,0)	ARIMA (3,1,0)	ARIMA (4,1,1)	ARIMA (4,1,2)
AIC	-8462,92	-8431,05	-8462,33	-8460,40

Fonte: elaboração dos autores.

Os valores estimados para o modelo ARIMA (4, 1, 0) são apresentados na Tabela 3. Deve-se ressaltar que os valores entre parênteses correspondem ao desvio padrão das estimativas de cada coeficiente, por meio do qual é possível verificar a significância estatística dos mesmos (todos significantes ao nível de 5%).

Tabela 3
Modelo ARIMA (4, 1, 0)

Parâmetros	Valor estimado
	0,0575
φ_1	(0,0277)
	0,1415
φ_2	(0,0270)
	0,2290
φ_3	(0,0270)
	0,1640
φ_4	(0,0277)
Variância do erro estimada	0,0000974

Fonte: elaboração dos autores.

5.3 Previsão

Encontrada a modelagem adequada para descrever a trajetória dos preços à vista do milho, a etapa a seguir é a realização da previsão para um horizonte de dez dias, correspondente ao período: 02/03/2009 a 13/03/2009. Para tanto, dois métodos de previsão foram utilizados. O primeiro (PPM1), mais simples, é feito de maneira estática, ou seja, não há entrada de novas informações no procedimento de previsão. Já o segundo (PPM2) contempla a possibilidade de retroalimentação, que faz com que cada nova previsão incorpore o valor observado no período anterior. Além disso, também é colocada a “previsão” feita pelo mercado futuro (PFM). Essas três formas de previsão foram comparadas com os valores reais do preço *spot* do milho (PSM) no período citado acima. Os resultados dessas previsões podem ser vistos na Tabela 4, que traz, além das três previsões, os seus respectivos desvios em relação ao preço físico efetivo.

Tabela 4
Previsão do preço *spot* do milho

Data	PSM	PFM	Desvio	PPM1	Desvio	PPM2	Desvio
2/3/2009	20,60	20,98	-0,38	20,96	-0,36	20,95	-0,35
3/3/2009	20,65	20,85	-0,20	20,97	-0,32	20,59	0,06
4/3/2009	20,50	20,85	-0,35	20,97	-0,47	20,60	-0,10
5/3/2009	20,35	20,60	-0,25	20,93	-0,58	20,37	-0,02
6/3/2009	20,43	20,55	-0,12	20,92	-0,49	20,26	0,17
9/3/2009	20,42	20,45	-0,03	20,92	-0,50	20,38	0,04
10/3/2009	20,50	20,30	0,20	20,91	-0,41	20,37	0,13
11/3/2009	20,17	20,55	-0,38	20,90	-0,73	20,49	-0,32
12/3/2009	20,17	20,40	-0,23	20,89	-0,72	20,17	0,00
13/3/2009	20,03	20,31	-0,28	20,89	-0,86	20,14	-0,11

Fonte: elaboração dos autores.

Para comparar a qualidade de previsão dos três “métodos” usados, é utilizado o erro quadrático médio (EQM) descrito na seção 2. Deve-se lembrar que quanto menor for o EQM melhor é a previsão obtida. A seguir a Tabela 5 mostra que o modelo econométrico PPM2 promoveu estimativas dos valores futuros mais acuradas. Isso indica que o modelo PPM2 não só se mostrou melhor que o modelo estático, o que era de se esperar dado que o PPM2 incorpora os valores reais em cada nova estimação, mas também que o próprio mercado futuro. Esse resultado deve ser interpretado com cautela, visto que o horizonte de previsão é pequeno e que o mercado de milho pode ter eficiência fraca, e não forte ou semi-forte. Feita essa ressalva, o resultado pode ser identificado como uma evidência a favor da estimação de modelos econométricos para se buscar prever valores futuros e vai ao encontro de resultados já encontrados em outros estudos como o de Springgs (1981). Contudo, deve-se ter em conta que foram em previsões de curto prazo (no caso um passo a frente) que o modelo se mostrou como melhor previsor que o mercado futuro, não ocorrendo o mesmo para um espaço de previsão maior sendo assim, o mercado continua tendo importância como indicador de comportamento de preços para um lapso temporal maior.

Tabela 5
Qualidade da previsão

	Mercado futuro	Modelo 1	Modelo 2
EQM	0,07004	0,3249281	0,02984

Fonte: elaboração dos autores.

6. CONCLUSÃO

O presente estudo teve como objetivo avaliar a capacidade de os instrumentos econométricos de séries temporais preverem os preços *spot* do milho. Além disso, buscou-se comparar a qualidade de previsão baseada na econometria e no mercado futuro, na tentativa de apontar um previsor mais fiel do preço à vista.

Os resultados apresentados na seção anterior mostraram que o modelo econométrico dinâmico (PPM2) obteve estimativas para os valores previstos mais robustas. Isso pode ser uma evidência da utilidade do ferramental econométrico na previsão de preços, usada em estratégias de gerenciamento de risco.

Entretanto esse resultado deve ser visto com “cuidado” quando se trata de colocá-lo como uma evidência da ineficiência no mercado, uma vez que a análise feita possui algumas limitações. Uma dessas limitações é a forma simplista com que os preços futuros foram usados, outra é o pequeno horizonte temporal de previsão, uma terceira limitação é o fato de que foram observados os preços apenas para a “praça” de Campinas-SP e, portanto o resultado aqui encontrado não deve e nem pode ser estendido diretamente para outras localidades, indicando a importância de outros estudos levarem em conta o diferencial de base entre as diferentes praças formadoras de preços regionais no País. Além dessas limitações, também se deve lembrar que a previsão dinâmica, apesar de mais robusta, pode não ser efetivamente utilizada pelos agentes, e nesse caso o mercado futuro se torna um “melhor previsor” em relação ao modelo estático de previsão (PPM1).

Por fim, algumas outras sugestões para futuras extensões desse estudo são feitas. Primeira, realizar a mesma análise para outras *commodities* brasileiras. Segunda, abordar o conceito de eficiência de mercado por meio de uma análise mais

robusta, como, por exemplo, a co-integração, em que é possível incluir outros fatores na sua determinação, tais como o custo de carregamento. Terceira, comparar modelos econométricos com diferentes horizontes de previsão. Última, apesar da boa aderência do modelo ARIMA, testar a qualidade de outras modelagens como, por exemplo, ARCH (Auto-Regressivos com Heteroscedasticidade Condicional), GARCH (Auto-Regressivos com Heteroscedasticidade Condicional Generalizado) e ARFIMA (Auto-Regressivos Fracionários Integrados de Médias Móveis).

BIBLIOGRAFIA

BAYER, F. M. "Previsão do preço e da volatilidade de *commodities*, por meio de modelagens ARFIMA-GARCH". Dissertação de mestrado apresentada ao Curso de Mestrado do Programa de Pós-Graduação em Engenharia de Produção, Área de concentração em Gerência da Produção, da Universidade Federal de Santa Maria (UFSM, RS), Santa Maria, 2008.

BRESSAN, A. A.; LIMA, J. E.. "Modelos de previsão de preços aplicados aos contratos futuros de boi gordo da BM&F". **Nova Economia**, p.117-140, Belo Horizonte, 2002.

BROOKS, C. **Introductory Econometrics for Finance**. Cambridge, 2º Edição, 2008.

CONAB. Companhia Nacional de Abastecimento. Disponível em: www.conab.gov.br. Acesso em: 28 de nov. de 2009.

ENDERS, W. **Applied Econometric Times Series**. John Wiley & Sons Inc, 2º Edição. 2003.

FAMA, E. "Efficient Capital Markets: A Review of Theory and Empirical Work", **Journal of Finance**, Vol. 25, p. 238 – 417, May, 1970.

FRENCH, K. R. "*Detecting Spot Price Forecasts In Futures Prices*", **The University of Chicago Press**, Vol. 59, Nº. 2, S39-S54, April, 1986.

HULL, J. C. Fundamentos dos mercados futuros e de opções. São Paulo: Bolsa de Mercadorias e Futuros / Cultura Editores Associados, 2005. 597p.

LEUTHOLD, R. M., *et al.* "An examination of the necessary and sufficient conditions for market efficiency: the case of hogss." **Applied economics**. Vol. 21, p. 193-204, Jan, 1989.

MARQUES, P. V.; MELLO, P.C. Mercados Futuros de Commodities Agropecuárias: Exemplos e Aplicações aos Mercados Agropecuários. São Paulo: Bolsa de Mercadorias e Futuros, 1999, 212 p.

MEDEIROS, A. L. *etal.* “Modelagem ARIMA na previsão do preço da arroba do boi gordo” **Sociedade Brasileira de Economia, Administração e Sociologia Rural**. Fortaleza, 2006.

MORETTIN, P. A.; TOLOI, C. M. C. **Análise de Séries Temporais**. Edgar Blucher, 2004.

PAVÃO, A. R. “Impactos econômicos da introdução do milho Bt11 no Brasil: uma abordagem de equilíbrio geral inter-regional”. Dissertação de mestrado apresentada ao Departamento de Economia, Administração e Sociologia, da Escola Superior de Agricultura Luiz de Queiroz (ESALQ-USP). Piracicaba, 2008.

PEREIRA, L. M. “Modelo de formação de preços de commodities agrícolas aplicado ao mercado de açúcar e álcool”. Tese apresentada ao Departamento de Economia da Escola de Economia, Administração e Contabilidade de São Paulo, Universidade de São Paulo. São Paulo, 2009.

PINTO, P. A. L. A. *etal.* “Aplicação do modelo ARIMA à previsão do preço das *commodities* agrícolas brasileiras”, **Sociedade Brasileira de Economia, Administração e Sociologia Rural**. Rio Branco, 2008.

SPRIGGS J. “Forecasts of Indiana Monthly Farm Prices Using Univariate Box-Jenkins Analysis and Corn Futures Prices”. **North Central Journal of Agricultural Economics**, Vol. 3, Nº1, pp.81-87. Jan, 1981.

TOMEK, W. G. “Commodity futures prices as forecasts”. **Review of Agricultural Economics**, 1997.

USDA. United States Department of Agriculture. Disponível em: www.usda.gov
Acesso em: 28 de nov. de 2009.