

**Análise da previsão do preço do etanol hidratado no estado de São Paulo : uma aplicação do modelo arima****Analysis of the forecasting of the hydrated ethanol price in the state of São Paulo: an application of the arima model**

DOI:10.34117/bjdv5n10-045

Recebimento dos originais: 10/09/2019

Aceitação para publicação: 04/10/2019

**Carlos Alberto Gonçalves da Silva**

Doutor em Engenharia de Produção – COPPE/UFRJ

Pós-Doutor em Economia Aplicada pela Universidade Federal Fluminense -UFF

Professor Visitante do PPGCE - *Universidade do Estado do Rio de Janeiro (UERJ)*, Rio de Janeiro.

Professor Sênior Colaborador do Departamento de Engenharia de Produção. CEFET-RJ.

E-mail: carlos.silva@uerj.br e gon7silva@gmail.com

**RESUMO**

O presente trabalho se propõe a analisar o comportamento dos preços médios do etanol hidratado no Estado de São Paulo. Para tanto, procurou-se realizar uma previsão para os preços deste produto a partir da metodologia ARIMA. O período analisado pela modelagem compreende novembro de 2002 a dezembro de 2017. O modelo ARMA (1,1) estimado foi eficiente e apresentou projeção satisfatória para previsão dos preços do etanol hidratado. Os resultados obtidos fornecem uma ferramenta de análise para o mercado desta *commodity*, na medida em que demonstram a tendência dos preços para um horizonte de curto prazo, servindo de auxílio à tomada de decisão de agentes que comercializam este bem.

**Palavras-chave:** Modelagem ARIMA, *commodity* etanol, Previsão.

**ABSTRACT**

The present work proposes to analyze the behavior of the average prices of hydrous ethanol in the State of São Paulo. In order to do so, a forecast was made for the prices of this product based on the ARIMA methodology. The period analyzed by modeling comprises November 2002 to December 2017. The ARMA (1,1) estimated it was efficient and presented satisfactory projection for hydrated ethanol prices forecast. The results obtained provide an analysis tool for the market of this commodity, as they demonstrate the trend of prices for a short term horizon, serving as an aid to the decision-making of agents that market this commodity.

**Keywords:** ARIMA modeling, ethanol commodity, Forecast.

**1.INTRODUÇÃO**

A utilização do etanol como combustível nos veículos automotores vem se modificando ao longo do tempo no Brasil. O Plano Nacional do Álcool (PNA) viabilizou o renascimento do etanol

como combustível automotivo e deu no Brasil a liderança nas políticas globais de transição dos combustíveis fósseis para as fontes renováveis de energia.

O elevado preço do petróleo, no final dos anos 2000, contribuiu para formulação de programas de substituição de combustível fósseis. Neste cenário de incerteza sobre o preço do óleo natural no mercado internacional, o etanol adquire grande importância.

No Brasil, maior produtor mundial de cana-de-açúcar, o governo federal observou uma oportunidade de expandir o mercado e tornar o país também o maior e exportador do produto. Passadas quatro décadas e meia, o álcool é o principal combustível renovável utilizado globalmente, graças ao pioneirismo brasileiro, o país surge como líder na produção de etanol.

A difusão dos veículos movidos com motores *flex*, capazes de operar com qualquer mescla de gasolina e etanol, minimizou os efeitos de conjunturas petrolíferas desfavoráveis na trajetória de difusão do uso do etanol combustível. Depois de um período de estagnação na década de 1990, a parcela do etanol no consumo de combustíveis da frota brasileira de automóveis voltou a crescer continuamente, concomitantemente com a trajetória de aumento no preço do petróleo deslanchada no início da década passada.

A introdução dos veículos *flex* no mercado brasileiro em 2003 revolucionou o mercado do etanol hidratado e marca a retomada da expansão do seu consumo<sup>1</sup>. A opção pelo consumo de etanol deixou de ser realizada no momento da aquisição do automóvel para ser uma decisão no ato de abastecimento, sempre que o preço do etanol seja competitivo. Paralelamente, a escalada do preço do barril de petróleo, repassada para o preço da gasolina, deu forte impulso competitivo ao etanol<sup>2</sup>. Desde então, a venda de veículos *flex* passou a ser dominante no mercado de veículos do ciclo Otto e o consumo de etanol hidratado cresceu rapidamente, até o final da década passada onde teve seu ápice em 2013 com 2,9 milhões de unidades vendidas de autoveículos<sup>3</sup> flex (Figura 1).

---

<sup>1</sup> Veículos com motores flexíveis capazes de operar com qualquer mescla de gasolina com etanol hidratado.

<sup>2</sup> Preços do etanol hidratado inferiores a 70% ao preço da gasolina tornam seu consumo economicamente competitivo para os proprietários dos veículos *flex*.

<sup>3</sup> Autoveículos inclui automóveis, comerciais leves, caminhões e ônibus.

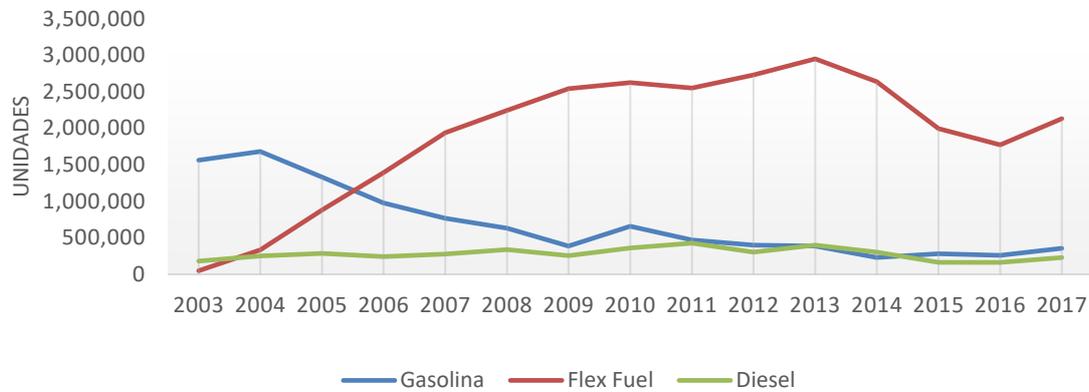


Figura 1 - Produção de autoveículos\* entre 2000 e 2017

Fonte: Elaborado pelo autor a partir de dados do Anuário Estatístico ANFAVEA 2018.

(\*) Autoveículos montados e desmontados (CKD).

A crise econômica global e a descoberta de vastos reservatórios de petróleo no pré-sal brasileiro induziram profunda mudança na política de precificação da gasolina e na política tributária adotada para os combustíveis. Essas mudanças provocaram drástica reversão na trajetória de expansão do consumo do etanol hidratado a partir de 2009, mergulhando o mercado desse combustível em fortes incertezas. Essas incertezas necessitam ser removidas para que seja preservado o papel necessário do etanol na política brasileira de mudanças climáticas e na nova conjuntura mundial.

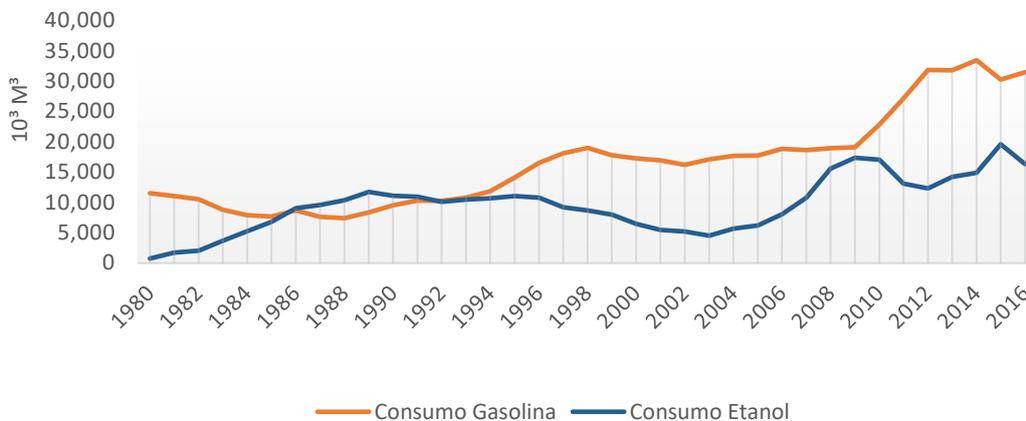


Figura 2 - Consumo de gasolina e etanol entre 1980 e 2016

Fonte: Elaborado pelo autor a partir de dados do Balanço Energético Nacional (BEN).

Desde então, o consumo e produção de etanol hidratado permanecem estagnados, colocando em risco a continuidade da trajetória ambientalmente benigna do sistema de transporte individual da economia brasileira (Figura 2).

Nesse contexto, este trabalho tem por objetivo avaliar a potencialidade da utilização de modelos de séries temporais na previsão de preços do etanol hidratado no Estado de São Paulo, ou seja, poderão ser empregados no planejamento e comercialização da *commodity*, gerando informações que darão suporte ao produtor em suas decisões.

Além desta introdução, o trabalho é composto por mais quatro partes. Na segunda parte, expor-se-á revisão de literatura. Na terceira explicar-se-á o modelo ARIMA e suas propriedades teóricas. Na quarta apresentar-se-á a análise e discussão dos resultados obtidos. Na seção seguinte serão expostas as conclusões sobre o trabalho. Finalmente, apresentar-se-ão as referências bibliográficas utilizadas.

## 2. REVISÃO DE LITERATURA

Verificam-se inúmeros trabalhos relacionados à previsão dos preços de *commodities*, devido a sua importância no auxílio à tomada de decisões de produtores e investidores.

A modelagem ARIMA e a técnica de Box e Jenkins (1976) para previsão de séries temporais são metodologias amplamente utilizadas nesse tipo de trabalho.

O estudo de Souza *et al.*(2007) analisa a confiabilidade da metodologia Box & Jenkins, aplicando o modelo ARIMA na previsão do preço da madeira serrada pago as exportações do Estado do Paraná. Os autores procuram examinar a acuracidade desses métodos econométricos visando dar

suporte e auxiliar na tomada de decisões por parte das empresas florestais. De acordo com os resultados, a metodologia foi capaz de formular um modelo estocástico com nível explicativo satisfatório para a série em estudo.

Lamounier e Leite (2007) procuram detectar a existência dos componentes estocásticos e/ou determinísticos de tendência, ciclo, e sazonalidade nos preços do mercado spot do café no Brasil. Os autores explicam que existem diversos métodos para a análise do componente sazonal em uma série de tempo, como: a análise visual do gráfico da série, os modelos de regressão linear com variáveis independentes binárias, e a utilização de métodos mais sofisticados como dos modelos de análise espectral, dos modelos de Box e Jenkins sazonais (SARIMA) e de alisamento sazonal.

A eficácia dos modelos de séries temporais na previsão dos preços de *commodities* é destaque nos estudos de Arêdes e Pereira (2008). Os autores procuraram avaliar a potencialidade da utilização de modelos de séries temporais na previsão de preços do trigo no Estado do Paraná. Os modelos de séries temporais estimados foram o ARIMA, o SARIMA, o ARCH, o GARCH e o TARARCH. Os resultados evidenciaram que todos os modelos são eficazes na previsão do preço do trigo, dado que os preços previstos são próximos aos observados.

O trabalho de Marques *et al.*(2010) procurou apresentar e analisar projeções de produção, consumo, exportação e preço de exportação do café. As projeções das variáveis descritas foram geradas a partir da utilização dos modelos de séries temporais.

De acordo com os autores, os modelos ARIMA, formulados por Box e Jenkins (1976), seguem a proposição de que uma série temporal não-estacionária pode ser modelada a partir de diferenciações e da inclusão de um componente autoregressivo e de um componente média móvel.

Felipe (2012) analisou a série de preços da soja do Norte do Paraná de janeiro de 2000 a outubro de 2011. O autor procurou determinar o modelo que melhor explica a dinâmica temporal dos preços da soja, utilizando a metodologia Box e Jenkins (1976). Observou-se que dentre o conjunto de modelos testados para previsão do preço da soja, o melhor modelo foi o AR (5), ou seja, a série preços da soja do Norte do Paraná pode ser explicada por um processo autoregressivo de ordem cinco.

### **3. METODOLOGIA E DADOS**

#### **3.1 TESTES DE RAIZ UNITÁRIA E SAZONALIDADE**

Para testar a estacionariedade das séries, serão utilizados neste trabalho, os testes ADF (Dickey – Fuller Aumentado) (1979 e 1981) e KPSS (Kwiatkowski, Phillips, Schmidt e Shin (1992), bem como o teste HEGY (2004) no sentido de verificar a existência de sazonalidade.

O teste de Dickey-Fuller Aumentado (ADF) consiste na estimação da seguinte equação por Mínimos Quadrados Ordinários (MQO):

$$\Delta Y_t = \alpha + \beta t + \gamma Y_{t-1} + \sum_{i=1}^p \delta_i \Delta Y_{t-i} + \varepsilon_t \quad (1)$$

onde  $\Delta Y_t$  é o operador de diferenças ( $Y_t - Y_{t-1}$ ),  $\alpha$  é uma constante,  $\beta t$  é a componente de  $\gamma$  tendência do modelo,  $\gamma$  é o coeficiente que permite testar a estacionariedade (se  $\gamma = 0$ ,  $Y$  tem uma raiz unitária),  $p$  é o número defasados a incluir no modelo e  $\varepsilon_t$  é o termo de erro aleatório ou perturbação estocástica.

O teste KPSS [Kwiatkowski, Phillips, Schmidt e Shin (1992)], surgiu como uma forma de complementar a análise dos testes de raiz unitária tradicional, como teste ADF. Ao contrário do teste ADF, o teste KPSS, considera como hipótese nula que a série é estacionária, ou estacionária em torno de uma tendência determinística, contra a hipótese alternativa que um caminho aleatório está presente.

Para avaliar a existência ou não de sazonalidade na série dos preços do milho, serão utilizados os testes de Hylleberg, Engle, Granger e Yoo, conhecido como teste de HEGY (1990), uma vez que a sazonalidade é muito comum em séries temporais mensais. Este teste verifica a existência de raiz unitária sazonal e raiz unitária regular (não-sazonal / tendência estocástica). De acordo com Lütkepohl e Krätzig (2004), o teste para dados mensais é dado pelo seguinte modelo estimado por MQO (mínimos quadrados ordinários):

$$\Delta_{12} y_t = \pi_1 y_{1,t-1} + \pi_2 y_{2,t-1} + \pi_3 y_{3,t-1} + \pi_4 y_{3,t-2} + \pi_5 y_{4,t-1} + \pi_6 y_{4,t-2} + \pi_7 y_{5,t-1} + \pi_8 y_{5,t-2} + \pi_9 y_{6,t-1} + \pi_{10} y_{6,t-2} + \pi_{11} y_{7,t-1} + \pi_{12} y_{7,t-2} + u_t + \varepsilon_t \quad (2)$$

Para analisar sua significância estatística empregam-se as estatísticas  $t$ , bem como o teste F (para significância conjunta dos parâmetros) e seus valores críticos são apresentados por Franses (1990) e Franses e Hobijn (1997).

Segundo Lütkepohl e Krätzig (2004), o processo  $Y_t$  terá raiz unitária não sazonal na frequência zero (tendência estocástica), caso  $\pi_1 = 0$ ; e terá raiz unitária sazonal, se ao menos um da ordem  $\pi_i$  ( $i=2,3,\dots,12$ ) for igual a zero.

### 3.2 MODELOS UNIVARIADOS DE SÉRIE DE TEMPO

Modelos univariados são aqueles em que os valores correntes de uma série de tempo são relacionados apenas com seus próprios valores passados ou com os valores correntes e passados dos resíduos da série temporal.

A combinação dos modelos auto-regressivos (AR) e de médias móveis (MA) resulta no modelo denominado ARMA. Assim, tratando-se de uma série temporal  $Y_t$  e considerando-se suas primeiras diferenças  $Y_t = Y_t - Y_{t-1}$ , com  $p$  defasagens para a variável e  $q$  para os erros aleatórios, pode-se expressar o modelo ARMA(p,q) da seguinte maneira:

$$Y_t = \mu + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \dots + \phi_p Y_{t-p} + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \dots - \theta_q \varepsilon_{t-q} \quad (3)$$

O modelo ARIMA (p, d, q), onde  $p$  refere-se ao número de defasagens da série,  $d$  é a ordem de integração e  $q$  é o número de defasagens dos erros aleatórios. Portanto, utiliza-se os procedimentos sugeridos por Box e Jenkins (1976), cujas etapas são as seguintes:

- a) Identificação - Determina-se os valores apropriados de  $p$ ,  $d$  e  $q$  e utiliza-se um correlograma apresentando as funções de autocorrelação parcial (FACP) e de autocorrelação simples (FAC), podendo determinar a ordem apropriada dos componentes AR e MA.
- b) Estimação – esta etapa vem após a identificação e consiste em estimar os parâmetros autoregressivos e de médias móveis. Verifica-se a significância estatística dos parâmetros sugeridos para o modelo.
- c) Verificação – Consiste em verificar se o modelo estimado ajusta-se adequadamente à série ou não e se os resíduos calculados apresentam ou não problemas de autocorrelação, bem como existência de heterocedasticidade. Os critérios de informação de Akaike e de Schwarz ajudam neste diagnóstico, bem como a utilização da estatística  $Q$  de Ljung-Box. O teste  $Q$  é usado para testar se um conjunto de autocorrelações de resíduos é ou não estatisticamente diferente de zero. A estatística  $Q$  conterà aproximadamente uma distribuição  $X^2$  com  $K-p-q$  graus de liberdade.
- d) Previsão - O modelo ARIMA estimado pode ser utilizado nas previsões, tornando-se possível calcular previsões de  $Y_t$  para  $n$  períodos seguintes, desde que sejam conhecidos  $Y_{t-1}$  e  $Y_t$ .

### 3.3 DADOS

O período da análise deste trabalho estende-se de novembro de 2002 a dezembro de 2017, com periodicidade mensal. Os preços médios do etanol hidratado no Estado de São Paulo foram obtidos do *site* do Centro de Estudos Avançados em Economia Aplicada (CEPEA-ESALQ/USP), sendo os preços atualizados monetariamente para dezembro de 2017 pelo Índice Geral de Preços – Disponibilidade Interna (IGP-DI) da Fundação Getúlio Vargas. Os *software* utilizado foi Eviews 8.0.

## 4. RESULTADOS EMPÍRICOS E ANÁLISES

### 4.1 ANÁLISES GRÁFICAS E TESTES PRELIMINARES

As Figuras 3 e 4 mostram o comportamento das séries de cotações dos preços e retornos diários dos preços da *commodity* etanol hidratado no período considerado.

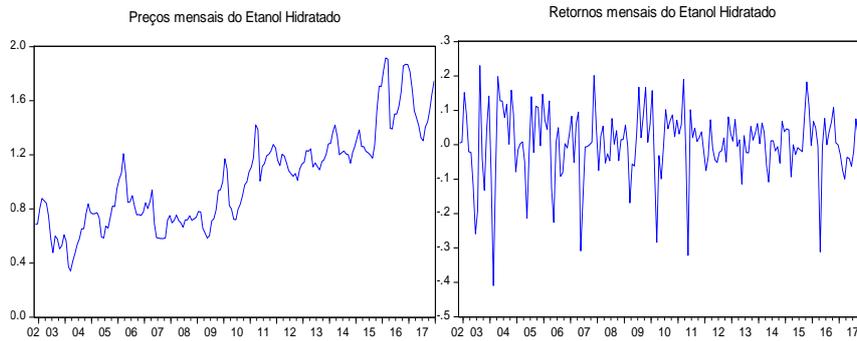


Figura 3- Preços médios mensais do etanol hidratado (R\$/litro)      Figura 4 – Retornos médios mensais do etanol hidratado

Na inspeção visual da Figura 3, observa-se um comportamento com tendência crescente apresentando oscilações atingindo o seu pico de preço em fevereiro de 2016. Dessa forma foi preciso testar a normalidade e estacionariedade da série mensais dos preços do etanol hidratado para aplicação dos modelos ARIMA. Visualmente não permite afirmar que existe a presença da sazonalidade, fato este que será testado a partir do teste HEGY.

Algumas estatísticas descritivas básicas são apresentadas na Tabela 1. Observa-se que os preços médios do etanol hidratado apresentam uma distribuição normal ao valor de curtose (3,2880). A estatística de Jarque-Bera indica p-valor de 0,3925, maior do que 5%, ou seja, não rejeita a normalidade da distribuição da série.

Tabela 1 – Sumário estatístico dos preços médios do etanol hidratado

Estatísticas	Média	Mediana	Máximo	Mínimo	Desvio padrão
Valores	1,4728	1,4606	2,3738	0,7411	0,2678

Estatísticas	Assimetria	Curtose	Jarque-Bera	p-valor	Observações
Valores	0,2023	3,2880	1,8703	0,3925	182

\

O Q-Q Plot representa um dos métodos gráficos mais utilizados na verificação da normalidade de séries temporais. O procedimento empregado consiste na comparação gráfica dos quantis teóricos da distribuição normal com os quantis dos dados amostrais.

A Figura 5 indica que a distribuição normal parece bem aderente aos dados. Assim sendo, todos os testes não rejeitaram a hipótese de normalidade da série analisada.

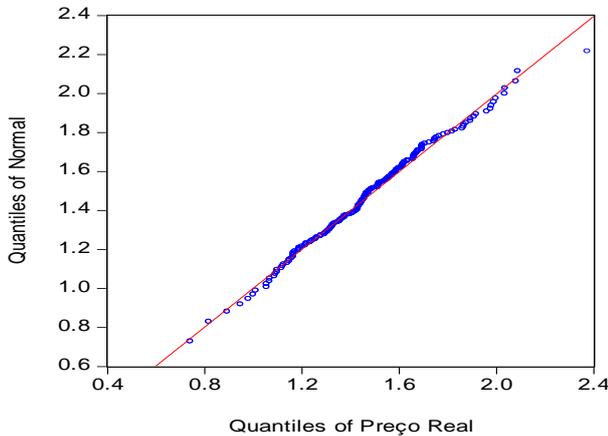


Figura 5 - Plot Q-Q Retornos médios mensais dos preços do etanol hidratado

## 4.2 TESTES DE RAIZ UNITÁRIA E SAZONALIDADE

Inicialmente, analisou as funções de autocorrelações (FAC) e de autocorrelações parciais (FACP). O comportamento dessas funções indica qual o modelo a ser utilizado. Ao realizar o estudo das funções em nível, de acordo com o correlograma, observa-se na tabela 2, um decaimento lento dos lag's o que indica de não estacionariedade, bem como a existência de sazonalidade, os quais serão verificados através dos testes de estacionariedade de ADF e KPSS e o teste de sazonalidade de HEGY.

Tabela 2 - Função de autocorrelação e autocorrelação parcial do preço médio mensal do etanol hidratado em nível

Sample: 2002M11 2017M12  
Included observations: 182

Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob
1	0.855	0.855	135.35	0.000	
2	0.657	-0.279	215.53	0.000	
3	0.508	0.116	263.85	0.000	
4	0.345	-0.246	286.24	0.000	
5	0.193	0.018	293.30	0.000	
6	0.096	0.019	295.05	0.000	
7	0.065	0.143	295.87	0.000	
8	0.081	0.081	297.13	0.000	
9	0.106	0.003	299.30	0.000	
10	0.128	-0.011	302.49	0.000	
11	0.161	0.054	307.54	0.000	
12	0.167	-0.077	313.02	0.000	
13	0.136	-0.017	316.71	0.000	
14	0.082	-0.084	318.05	0.000	
15	0.001	-0.098	318.05	0.000	
16	-0.077	-0.009	319.25	0.000	
17	-0.125	0.033	322.41	0.000	
18	-0.155	-0.022	327.33	0.000	
19	-0.159	0.034	332.52	0.000	
20	-0.131	-0.005	336.07	0.000	
21	-0.107	-0.070	338.47	0.000	

Para investigar formalmente a não estacionariedade da série, foram utilizados os testes Dickey-Fuller Aumentado (ADF) e KPSS [Kwiatkowski, Phillips, Schmidt e Shin]. De acordo com a tabela 3, as séries temporais de preços do etanol hidratado em nível, foram estacionárias com constante e tendência, considerando o nível de significância de erro de 5%.

Tabela 3 – Testes de raiz unitária Dikey Fuller Aumentado (ADF) e Kwiatkowski, Phillips, Schmidt e Shin (KPSS)

Variável	ADF	Valor Crítico (5%)	KPSS	Valor Crítico (5%)	Decisão
Milho	-5,4995	-3,4348	0,0776	0,1460	Rejeita $H_0$

Fonte: Elaboração própria com os dados da pesquisa.

O teste escolhido para verificar a sazonalidade das variáveis foi o teste de Hylleberg, Engle, Granger e Yoo (HEGY), por se tratar de séries temporais de preços das *commodities*, sujeitos à existência de ciclos e tendências e à variação sazonal da produção. Os valores contidos na tabela 4 mostram que, para a variável, em nível, se rejeita a hipótese de raiz unitária regular ( $\pi_1 = 0; H_0$ : possui raiz unitária regular) com significância de 5%, logo a série é estacionária em nível. A hipótese de raiz unitária sazonal ( $\pi_2 = 0; H_0$ : possui raiz unitária sazonal) foi rejeitada ao nível de significância

de 5%. Portanto, a série não possui raiz unitária sazonal. Vale ressaltar que os resultados são semelhantes com os testes ADF e KPSS.

Tabela 4 - Teste HEGY para raiz unitária dos preços do etanol hidratado com tendência e sazonalidade

$\pi$	Teste HEGY	Valor crítico	Decisão
$\pi_1$	-2,710	-2,520 *	Rejeita $H_0$
$\pi_2$	-3,760	-2,540 **	Rejeita $H_0$

Fonte: Resultados da pesquisa.

Nota: Valores críticos de Franses e Hobijn (1997).

$t(\pi_1) = -3,40$  a 1%;  $-2,82$  a 5%;  $-2,52$  a 10%.  $t(\pi_2) = -2,54$  a 1%;  $-1,94$  a 5%;  $-1,60$  a 10%.

(\*) indica rejeição da hipótese nula a 10% de significância (\*\*) indica rejeição da hipótese nula

a

1% de significância.

Com relação à sazonalidade, esta não foi identificada pelo teste HEGY. Assim, foram estimados modelos ARMA, conforme apresentados na tabela 7. No processo de modelagem e escolha dos modelos ARMA, devemos testar modelos diferentes e verificar a escolha do mais adequado com base nos menores valores dos critérios de informação Akaike (AIC), Schwarz (SIC), coeficiente de desigualdade de Theil-U, raiz do erro quadrático médio (RMSE) e erro absoluto médio (MAE), bem como a estatística de Durbin-Watson (DW) que também detecta a presença de autocorrelação nos resíduos de uma análise de regressão, ou seja, quando DW é aproximadamente igual a 2,0, ele indica que não existe autocorrelação nos resíduos. O coeficiente de Theil-U, avalia o desempenho da previsão em relação à previsão ingênua ou trivial. Previsão ingênua ou trivial significa que a estimativa do valor futuro é igual ao valor atual. O coeficiente de Theil analisa a qualidade de uma previsão através dos seguintes valores:

- $U > 1$ , significa que o erro do modelo é maior do que da previsão ingênua;
- $U < 1$ , significa que o erro do modelo é menor que da previsão ingênua.

O coeficiente de Theil-U menor do que 1 já indica uma previsão melhor que a previsão ingênua; quanto mais próximo o mesmo for de zero, melhor será o resultado da previsão. A estatística Theil-

U aplicada aos dados estimados e observados de novembro de 2002 a dezembro de 2017 é 0,0426 (tabela 5), indicando também uma boa qualidade do modelo em relação aos valores preditos.

Desta forma, após várias simulações e com os critérios mencionados anteriormente, a melhor estimativa foi para o ARMA (1,1), devido aos menores valores observados para os critérios de AIC e SIC, coeficiente de Durbin-Watson (DW), RMSE, MAE e Theil-U. (Tabela 5)

Tabela 5–Modelos estimados e critérios utilizados para seleção do melhor modelo

Modelos	AIC	SIC	DW	RMSE	MAE	Theil-U
ARMA (1,1) *	- 1,2542	- 1,2012	2,10	0,1271	0,0933	0,0426
ARMA (1,2)	- 1,1346	- 1,0815	1,39	0,1349	0,0987	0,0452
ARMA (2,1)	- 1,1853	- 1,1321	1,51	0,1315	0,0958	0,0441
ARMA (2,2)	- 0,3532	- 0,3000	0,93	0,1994	0,1510	0,0670
ARMA (1,3)	- 1,1237	- 1,0707	1,41	0,1357	0,0970	0,0454
ARMA (3,1)	- 0,8887	- 0,8353	1,20	0,1526	0,1192	0,0513

Fonte: Elaboração própria com base nos resultados da pesquisa.

(\* ) melhor modelo de previsão

A equação estimada do modelo foi:

$$Y_t = 1,4733 + 0,7319Y_{t-1} + 0,5331\varepsilon_{t-1}$$

*p – valor*      (0,0000)   (0,0000)   (0,0000)

Na figura 6, podem-se verificar os preços observados, os preços previstos e os valores residuais no período analisado. Portanto, as estimativas servem de embasamento para maior segurança em investimentos financeiros, dando suporte para produtores em suas tomadas de decisão.

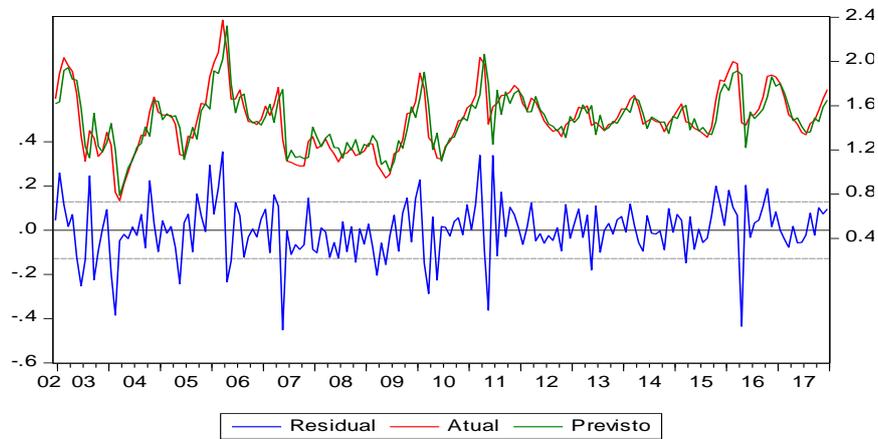


Figura 6 - Preços do etanol hidratado observados, previstos e valores residuais no período de novembro de 2002 a dezembro de 2017

Após a escolha do modelo é necessário realizar a análise de resíduos (ajustados). Se o modelo for adequado aos dados espera aleatoriamente em torno de zero com variância aproximadamente constante. Também espera-se que os resíduos sejam independentes e possuam distribuição normal. Na tabela 7 é apresentado o Teste Ljung-Box amostral é aproximadamente Qui-quadrado dos resíduos, cujas hipóteses podem ser interpretadas da seguinte maneira:

$$H_0: \chi^2 = 0 \text{ (Resíduos independentes).}$$

$$H_1: \chi^2 \neq 0 \text{ (Resíduos não são independentes).}$$

Pode-se verificar que todas as observações possuem um p-valor alto, indicação de que os resíduos são independentes.

Tabela 7 - Função de autocorrelação parcial e autocorrelação dos resíduos e teste Ljung-Box.

Sample: 2002M11 2017M12  
Included observations: 181

Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob	
		1	0.127	0.127	2.9712	0.085
		2	0.105	0.091	5.0231	0.081
		3	0.092	0.070	6.6060	0.086
		4	-0.063	-0.093	7.3505	0.118
		5	-0.003	-0.001	7.3524	0.196
		6	0.003	0.012	7.3543	0.289
		7	-0.034	-0.022	7.5701	0.372
		8	-0.033	-0.035	7.7804	0.455
		9	-0.104	-0.097	9.8831	0.360
		10	-0.062	-0.029	10.639	0.386
		11	-0.077	-0.049	11.795	0.379
		12	0.018	0.055	11.860	0.457
		13	0.048	0.047	12.318	0.502
		14	0.119	0.111	15.132	0.369
		15	0.054	0.006	15.713	0.401
		16	-0.049	-0.089	16.200	0.439
		17	0.096	0.093	18.073	0.384
		18	-0.094	-0.113	19.868	0.340
		19	-0.065	-0.058	20.740	0.351
		20	-0.052	-0.068	21.287	0.380
		21	-0.065	-0.005	22.166	0.390

O procedimento de previsões realizado neste trabalho evidencia o modelo que apresenta a melhor capacidade preditiva, uma vez que realiza previsões para cada modelo, com reajuste a cada passo, e verifica o modelo em que as previsões se aproximam mais dos valores reais da série temporal. Observa-se que as previsões obtidas pelo processo ARMA (1,1) são mais próximas dos valores observados da série dos preços do etanol hidratado, evidenciando o bom desempenho de previsão do modelo proposto.

## 5. CONCLUSÕES

A metodologia de previsão de Box-Jenkins é muito ampla, flexível e altamente subjetiva, mas, também é uma importante ferramenta de gestão que pode auxiliar o processo de tomada de decisão e planejamento futuro. Entretanto, o ponto fundamental para um mercado competitivo como o do etanol é o de que uma antecipação de preços gera uma série de benefícios a todos os interessados neste mercado.

O modelo ARMA (1,1) estimado no presente estudo foi eficiente e apresentou projeção satisfatória para previsão dos preços do etanol hidratado no Estado de São Paulo. Os resultados obtidos fornecem uma ferramenta de análise para o mercado, na medida em que demonstram a tendência dos preços para um horizonte de curto prazo, servindo de auxílio à tomada de decisão de agentes que comercializam este bem.

Apesar da modelagem ARMA apresentar um adequado poder de previsão de curto prazo, recomenda-se, em outros trabalhos de pesquisa, proceder à estimação utilizando outros modelos, tais como: ARMAX e volatilidade por meio da modelagem da família ARCH.

### REFERÊNCIAS

AKAIKE, H. A- *New look at the statistical model identification*. IEEE Transactions on Automatic Control, Waschiton, v.19, p.716-723, 1973.

ARÊDES, A.F.; PEREIRA, M. W.G. - Potencialidade da utilização de modelos de séries temporais na previsão do preço do trigo no Estado do Paraná. **Revista de Economia Agrícola**, São Paulo, v. 55, n. 1, p. 63-76, jan./jun., 2008.

BOX, G. P.; JENKINS, G. M. *Time series analysis, forecasting and control*. Holden-Day, San Francisco 1976.

DICKEY, D. A. & FULLER, W. A. - Likelihood ratio statistics for autoregressive Time series with unit root . **Econometrica**, 49(4): 1057-1072, Julho, 1981.

FELIPE, I. J. S. – Aplicação de modelos Arima em séries de preços de soja no Norte do Paraná. **Tekhne e Logos**, Botucatu, SP, v.3, n.3, Novembro. 2012.

FRANSES, P. H. - Testing for seasonal unit root in monthly data. Econometric Institute Report 9032A. Rotterdam: Erasmus University, 1990.

FRANSES, P. H.; HOBIJN, B. - Critical values for unit root tests in seasonal time series. **Journal of Applied Statistics**. v. 24, p. 25-46, 1997.

HYLLEBERG, S.; ENGLE, R.; GRANGER, C.; YOO, B. - Seasonal Integration and Cointegration. **Journal of Econometrics**, 44, 215-38, 1990.

JARQUE, C.; BERA, A. (1987). Test for normality of observations and regression residuals. **International Statistical Review**. 55(2), 163-172.

LAMOUNIER, W.M.- Tendência, ciclos e sazonalidade nos preços do café . **Gestão e Produção**, São Carlos, v. 14, n. 1, p. 13-23, jan.-abr. 2007.

LÜTKEPOHL, H.; KRÄTZIG, M. Applied time series econometrics. New York: Cambridge University Press, 2004. 323p.

MARQUES, D. V.; SOUZA, G.S.; SOUZA, M. O.; MARRA, R.- Análise do mercado de café via modelos de previsão. Simpósio de Pesquisa Operacional, 42. 2010, Bento Gonçalves. (2010).

PHILLIPS, P.C.B.; PERRON, P. (1988). Testing for a unit root in time series regression. **Biometrika**, 75(3), 335-346.

SOUZA, V. S.; NUNES, B.H.C.; ALMEIDA, A. N.; EISFELD, C. L. - Utilização da metodologia de Box & Jenkins na previsão do preço futuro pago as exportações paranaenses de madeira serrada. V Encontro de Economia Paranaense, 2007.

ZIVOT, E.; W.K. ANDREWS. - 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, 1992.