

Modelos Box-Jenkins de produção de biodiesel no Brasil

Verônica Dalmolin Cattelan (UFSM) veronica_vdc@hotmail.com

Renata Rocha de Oliveira (UFSM) reoliveira.2808@gmail.com

Adriano Mendonça Souza (UFSM) amsouza.sm@gmail.com

Julio Cezar Mairesse Siluk (UFSM) jsiluk@ufsm.br

Janis Elisa Ruppenthal (UFSM) profjanis@gmail.com

Resumo:

A fabricação de biodiesel está distribuída por diversos países pelo mundo, com legislação própria e estratégia comercial definida em cada um deles. Essa proliferação pelos continentes facilita a entrada do produto na lista de commodities mundiais, favorecendo o desenvolvimento de seu mercado em uma escala global. Nesse contexto, a metodologia Box-Jenkins foi utilizada para analisar dados históricos de produção de biodiesel e obter previsões da quantidade produzida no Brasil, que com a rápida evolução do seu programa de biodiesel, está inserido no cenário mundial, alcançando o posto de quarto maior produtor de biodiesel do mundo. A obtenção dos modelos mais adequados foi baseada na análise de gráficos e em testes estatísticos próprios da metodologia, os quais subsidiaram a decisão de adotar o modelo SARIMA(2,1,0)(1,1,0)₁₂ para previsão da produção nacional desse biocombustível.

Palavras chave: Modelos Box-Jenkins, Modelos SARIMA, Séries Temporais, Produção de Biodiesel

Box-Jenkins model of biodiesel production in Brazil

Abstract

The production of biodiesel is spread over many countries around the world with its own laws and business strategy defined in each of them. This proliferation by continents facilitates entry of the product in the list of world commodities, favoring the development of your market on a global scale. In this context, the methodology Box-Jenkins was used this article to analyze historical data for biodiesel production and to obtain estimates of the amount produced in Brazil, with the rapid development of its biodiesel program, is well positioned in the global scenario, reaching the rank Room largest biodiesel producer in the world. The choice of the optimum model was based on the analysis of charts and statistical tests themselves, which lead to the decision to adopt the SARIMA(1,1,0)(1,1,0)₁₂ model for forecasting the national production of this biofuel.

Key-words: Box-Jenkins Models, SARIMA Models, Time Series, Biodiesel Production

1. Introdução

Os dois principais biocombustíveis líquidos usados no Brasil são o etanol extraído de cana-de-açúcar e, em escala crescente, o biodiesel, que é produzido a partir de óleos vegetais ou de gorduras animais e adicionado ao diesel de petróleo em proporções variáveis. Estes, são derivados de biomassa renovável que podem substituir, parcial ou totalmente, combustíveis derivados de petróleo e gás natural em motores a combustão ou em outro tipo de geração de energia (ANP, 2014).

Os biocombustíveis poluem em menor quantidade por emitirem menos compostos do que os combustíveis fósseis no processo de combustão dos motores e também porque seu processo de produção tende a ser mais limpo (LEE, 2011; LEONARDI et. al, 2011; KOHLHEPP, 2011; ANP, 2014). Pioneiro mundial no uso de biocombustíveis, o Brasil alcançou uma posição almejada por muitos países que buscam fontes renováveis de energia como alternativas estratégicas ao petróleo. Nesse contexto, a ANP (2014) relata que atualmente cerca de 45% da energia e 18% dos combustíveis consumidos no Brasil já são renováveis e no resto do mundo, 86% da energia vêm de fontes energéticas não-renováveis.

Na comparação com o diesel de petróleo, o biodiesel também tem significativas vantagens ambientais. Estudos do National Biodiesel Board (associação que representa a indústria de biodiesel nos Estados Unidos) demonstraram que a queima de biodiesel pode emitir em média 48% menos monóxido de carbono; 47% menos material particulado (que penetra nos pulmões); 67% menos hidrocarbonetos. Como esses percentuais variam de acordo com a quantidade de B100 adicionado ao diesel de petróleo, no B3 essas reduções ocorrem de modo proporcional (ANP, 2014).

A Lei nº 11.097, publicada em 13 de janeiro de 2005, introduziu o biodiesel na matriz energética brasileira e ampliou a competência administrativa da Agência Nacional do Petróleo, Gás Natural e Biocombustíveis (ANP) e assumiu as atribuições de especificar e fiscalizar a qualidade dos biocombustíveis e garantir o abastecimento do mercado, em defesa do interesse dos consumidores. A Agência também executa as diretrizes do Conselho Nacional de Política Energética para os biocombustíveis e tem as funções de estabelecer as normas regulatórias, autorizar e fiscalizar as atividades relacionadas à produção, transporte, transferência, armazenagem, estocagem, importação, exportação, distribuição, revenda e comercialização e avaliação de conformidade e certificação de biocombustíveis (ANP, 2014).

O presente artigo tem por objetivo principal, analisar dados históricos de produção de biodiesel e obter previsões da quantidade produzida desse biocombustível no Brasil por meio da metodologia de Box-Jenkins. O estudo a partir dessa abordagem se justifica, no sentido de apresentar a escolha dos modelos adequados para esta variável e poderá subsidiar decisões no cenário brasileiro de produção de biodiesel, verifica-se também o comportamento da série em relação a sua tendência e sazonalidade.

A fim de atender ao objetivo deste trabalho relacionado à previsão de produção de biodiesel no Brasil, no período de Junho/2013 a Maio/2014, inicialmente, reuniram-se os dados no período de Março/2005 a Maio/2013, correspondentes à produção de Biodiesel no Brasil a partir do Banco de dados da ANP (2014). Desta forma, empregou-se a metodologia de Box-Jenkins para o valor de produção, em volume, na série em estudo.

2. Metodologia

O conjunto de dados considerado neste artigo corresponde a 110 observações mensais de produção de biodiesel em m³ (metros cúbicos) no Brasil, no período de março de 2005 a maio de 2014. Estes dados foram obtidos a partir do site oficial da ANP.

Inicialmente, analisou-se o gráfico do histograma, de forma a verificar a necessidade de aplicação de alguma forma de transformação dos dados de modo a variância e tornar o efeito sazonal aditivo.

Posteriormente, verificou-se o gráfico da série original juntamente com a série de primeira diferença para identificar em qual poderia ser realizada a modelagem, através dos testes de estacionariedade de Dickey- Fuller e KPSS. Para verificar a presença de tendência e sazonalidade, realizou-se a análise da função de autocorrelação (fac), função de

autocorrelação parcial (facp). Através da mesma, juntamente com o gráfico da série, sugeriu-se alguns modelos concorrentes.

Após a estimação dos modelos concorrentes, foi feita a análise de resíduos para identificarmos se existe o comportamento ruído branco, que como pressuposto de um modelo adequado, os resíduos também devem se comportar de forma estacionária.

Para a escolha de um modelo, após encontrar vários modelos que representem as séries estudadas, estes modelos serão denominados de modelos concorrentes e utilizam-se critérios de penalização para auxiliar na escolha do melhor modelo, sendo: o Critério de Informação Akaike (AIC) e o Bayesian Information Criterion (BIC), os quais utilizam o valor maximizado da Função de Verossimilhança para o modelo estimado (L), o número de parâmetros (n) e o tamanho da amostra (T) (SOUZA; SOUZA; MENEZES, 2012).

Através da Metodologia que será abordada, Box-Jenkins, foram descritos no referencial os modelos ARIMA, SARIMA E ARFIMA, ao fato mostraram ser os modelos mais adequados para concorrer e descrever a série. A partir do escolhido realizaram-se as previsões para os meses de junho de 2013 até maio de 2014. A análise da série foi feita utilizando o software EVIEWS 8.0.

2.1 Modelagem Box-Jenkins

Uma das técnicas quantitativas mais difundida na década de 70 é a metodologia de Box-Jenkins, que partem da ideia que cada valor da série (temporal) pode ser explicado por valores prévios, a partir do uso da estrutura de correlação temporal que geralmente há entre os valores da série Box; Jenkis, 1970). Estes modelos têm sido largamente utilizados para modelagem e previsão em aplicações médicas, ambientais, financeiras e de engenharia (WERNER e RIBEIRO, 2003; MORETTIN e TOLOI, 2006; PEREIRA et. al, 1986; FAVA, 2000).

Conforme Morettin e Toloí (2006) considera-se série temporal qualquer conjunto de observações ordenadas no tempo. Os autores citados ainda definem que as séries temporais são compostas por 4 elementos: a tendência (evidencia o sentido do deslocamento da série ao longo do tempo); o ciclo (periodicidade do movimento da série ao longo dos anos); a sazonalidade (picos repetitivos em intervalos iguais de tempo dentro da série, geralmente inferior a um ano); e o erro (corresponde a variabilidade intrínseca dos dados, não podendo ser modelado).

Através da modelagem Box-Jenkins, serão estudados os modelos de séries temporais denominados ARIMA (autorregressivos integrados e de médias móveis) e suas variações. Para Clemente (2012), o Teorema de Wald consiste em ajustar modelos autorregressivos integrados de médias móveis a um conjunto de dados, sendo este teorema, a base da metodologia de Box-Jenkins. Sua modelagem baseia-se em um ciclo iterativo, em que a escolha da estrutura do modelo é endógena, ou seja, a partir dos dados.

De acordo com Morettin e Toloí (2006), a metodologia Box-Jenkins é composta de um ciclo iterativo, de forma que a escolha da estrutura do modelo é baseada nos próprios dados, tais como:

- a) considera-se para a análise uma classe geral de modelos, chamada especificação;
- b) a fase de identificação é realizada com base na análise de autocorrelações, autocorrelações parciais e outros critérios;
- c) após identificado o modelo, segue para a fase de estimação, que corresponde a estimação dos parâmetros de cada modelo;

d) e por fim, último estágio, verificação ou diagnóstico do modelo ajustado, através de uma análise de resíduos, para certificar-se que o modelo é adequado para a finalidade da técnica, a previsão.

Caso o modelo não seja adequado, repete-se o ciclo, voltando a fase inicial de identificação (MORETTIN e TOLOI, 2006).

2.1.1 Modelos Não-Estacionários

Para Werner e Ribeiro (2003), acontece a não-estacionariedade quando uma série temporal apresenta média e variância dependentes do tempo. Essa série apresenta a inclinação dos dados e eles não permanecem em torno de uma linha horizontal do tempo, e/ou; a variação dos dados não permanece constante sobre o tempo, ou seja, as flutuações aumentam ou diminuem com o passar do tempo, indicando que a variância não é constante (WERNER e RIBEIRO, 2003; GUJARATI, 2000).

A detecção da não-estacionariedade de uma série pode ser feita a partir da análise gráfica das funções de autocorrelação (fac) e autocorrelação parcial (facp), em busca de padrões, ou realizando testes estatísticos de raiz unitária, em se tratando desse trabalho utilizou-se os testes de Dickey-Fuller e KPSS.

2.1.2 Modelos Autoregressivos Integrados de Médias Móveis (ARIMA)

De acordo com Morettin e Tolo (2006), a partir do desenvolvimento matemático de uma função simples, existe a dedução de como se chegaram aos modelos generalizados. Nos modelos ARIMA, os autores trazem que se $W_t = \Delta^d Z_t$ for estacionária, pode-se representar W_t por um modelo ARMA(p,q), ou seja,

$$\phi(B)W_t = \theta(B)a_t \quad (1)$$

Se W_t for uma diferença de Z_t , então Z_t é uma integral de W_t , daí diz-se que Z_t segue um modelo autoregressivo, integrado, de médias móveis, ou modelo ARIMA,

$$\phi(B)\Delta^d Z_t = \theta(B)a_t \quad (2)$$

de ordem (p,d,q) e escreve-se ARIMA(p,d,q), se p e q são as ordens de $\phi(B)$ e $\theta(B)$, respectivamente (MORETTIN e TOLOI, 2006).

Na equação 1 todas as raízes de $\phi(B)$ estão fora do círculo unitário. É equivalente a escrever o modelo

$$\varphi(B)Z_t = \theta(B)a_t \quad (3)$$

onde $\varphi(B)$ é um operador autoregressivo não-estacionário, de ordem p+d, com d raízes iguais a um (sobre o círculo unitário) e as restantes p fora do círculo unitário, ou seja,

$$\varphi(B) = \phi(B)\Delta^d = \phi(B)(1 - B)^d \quad (4)$$

Observa-se que é indiferente escrever $\varphi(B)Z_t$ ou $\varphi(B)\tilde{Z}_t$, pois $\Delta^d Z_t = \Delta^d \tilde{Z}_t$, para $d > 1$.

Desta forma, Morettin e Tolo (2006) deduzem que, o modelo 2, supõe que a d-ésima diferença da série Z_t pode ser representada por um modelo ARIMA, estacionário e invertível. Em grande parte dos casos usuais, $d = 1$ ou $d = 2$, que correspondem a dois casos interessantes e comuns de não-estacionariedade homogênea:

a) séries não-estacionárias quanto ao nível: oscilam ao redor de um nível médio durante algum tempo e depois pulam para outro nível temporário. Para que se tornem estacionárias é suficiente tomar apenas uma diferença;

b) séries não-estacionárias quanto à inclinação: oscilam em uma direção por um tempo e depois mudam para outra direção, também temporária. Para estacionarizá-las é necessário uma segunda diferença.

2.1.3 Modelos Sazonais

Para Werner e Ribeiro (2003) os modelos ARIMA exploram a autocorrelação entre os valores de série em instantes sucessivos, mas quando observados em períodos inferiores a um ano, a série pode apresentar autocorrelação para uma estação de sazonalidade s . Os modelos contemplam as séries que apresentam autocorrelação sazonal são chamados de SARIMA (WERNER e RIBEIRO, 2003).

Abreu et al. (2012) definem os modelos ARIMA como aqueles que exploram a correlação de valores observados em instantes de tempo consecutivos, enquanto os modelos ARIMA sazonais, representados por SARIMA, apresentam correlação entre os instantes de tempo distantes entre si por m ou múltiplos de m , sendo que m é o número de observações contidas em um ano ($m = 12$ para dados mensais e $m = 4$ para dados trimestrais)

De acordo com Morettin e Toloí (2006) quando $\{Z_t\}$ exibe um comportamento sazonal determinístico com período 12, um modelo pode ser utilizado:

$$Z_t = \mu_t + N_t \quad (5)$$

onde μ_t é uma função determinística periódica, satisfazendo $\mu_t - \mu_{t-12} = 0$, ou

$$(1 - B^{12})\mu_t = 0 \quad (6)$$

e N_t é um processo estacionário que pode ser modelado por um ARMA(p,q) (MORETTIN e TOLOI, 2006).

Assim, N_t satisfaz a equação

$$\phi(B)N_t = \theta(B)a_t. \quad (7)$$

Dessa forma, para um modelo sazonal determinístico, aplicando a diferença sazonal $(1 - B^{12})$ à expressão (5), obtém-se

$$(1 - B^{12})Z_t = (1 - B^{12})\mu_t + (1 - B^{12})N_t \quad (8)$$

e de acordo com (6), têm-se

$$(1 - B^{12})Z_t = (1 - B^{12})N_t. \quad (9)$$

Substituindo (7) em (9), obtém-se

$$\phi(B)(1 - B^{12})Z_t = \theta(B)(1 - B^{12})a_t \quad (10)$$

ou

$$\phi(B)W_t = \theta(B)(1 - B^{12})a_t, \quad (11)$$

onde $W_t = (1 - B^{12})Z_t$ (MORETTIN e TOLOI, 2006).

2.1.4 Modelos de Memória Longa - ARFIMA

Essa família de processos, para fins de modelagem da séries com comportamento de decaimento lento, ou memória longa, foi escolhida pelo efeito do parâmetro d em observações distantes, que decai hiperbolicamente conforme a distância aumenta, quando os efeitos dos parâmetros ϕ e θ têm decaimento exponencial. Dessa forma, d deve ser escolhido com o objetivo de explicar a estrutura de correlação de ordens altas da série, enquanto os parâmetros ϕ e θ explicam a estrutura de correlação de ordens baixas, por isso esses modelos ARFIMA são denominados de modelos de memória longa (SOUZA et al, 2010; SOUZA et al, 2011).

Qualquer número real $d > -1$, define-se o operador de diferença fracionária

$$(1 - B)^d = \sum_{k=0}^{\infty} \binom{d}{k} (-B)^k \quad (12)$$

$$= 1 - dB + \frac{1}{2!}d(d-1)B^2 - \frac{1}{3!}d(d-1)(d-2)B^3 + \dots$$

diz-se que $\{Z_t\}$ é um processo autoregressivo fracionário integrado de médias móveis, ou ARFIMA(p,d,q) com $d \in (-\frac{1}{2}, \frac{1}{2})$, se $\{Z_t\}$ for estacionário e satisfizer a equação

$$\phi(B)(1 - B)^d Z_t = \theta(B)a_t, \quad (13)$$

onde $\{a_t\}$ é ruído branco e $\phi(B)$ e $\theta(B)$ são polinômios em B de graus p e q, respectivamente (MORETTIM e TOLOI, 2006; SOUZA et al, 2010; SOUZA et al, 2011).

3. Análise dos resultados

A análise da Figura 1 auxilia a identificação dos possíveis modelos a serem utilizados. A Figura mostra a tendência crescente de volume de produção (m^3), ao longo dos 10 anos, indicando não-estacionariedade da série.

Após, realiza-se a análise das funções de autocorrelação (FAC) e autocorrelação parcial (FACP), Figura 2, da série original, possibilitando a visualização dos modelos possíveis juntamente com a Figura 1.

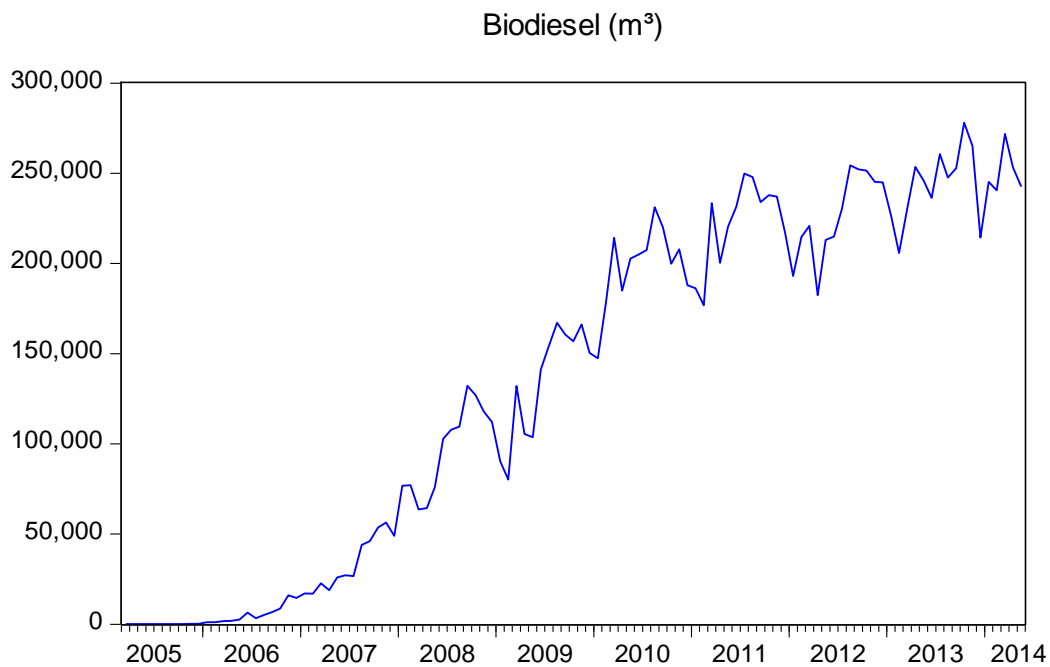


Figura 1 – Série de produção de Biodiesel

A Figura 2, contendo a FAC da série, revela que as autocorrelações apresentam decaimento lento, o que poderia indicar uma característica fracionária na série, chamados modelos ARFIMA. Pode-se visualizar que na FACP teríamos um modelo AR(1) por haver apenas um lag discrepante e bem definido, indicando que este seria o modelo um modelo concorrente adequado.

Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob	
		1	0.969	0.969	107.06	0.000
		2	0.944	0.083	209.59	0.000
		3	0.921	0.038	308.21	0.000
		4	0.899	-0.00...	402.93	0.000
		5	0.875	-0.03...	493.47	0.000
		6	0.857	0.083	581.13	0.000
		7	0.834	-0.06...	665.03	0.000
		8	0.812	-0.00...	745.35	0.000
		9	0.793	0.026	822.61	0.000
		1...	0.773	-0.01...	896.82	0.000
		1...	0.751	-0.03...	967.62	0.000
		1...	0.731	-0.01...	1035.3	0.000
		1...	0.701	-0.16...	1098.1	0.000
		1...	0.674	0.004	1156.8	0.000
		1...	0.642	-0.10...	1210.7	0.000

Figura 2 – FAC e FACP da Série Original de Biodiesel

Após a análise da FAC e FACP da série original, realizou-se dois testes de estacionariedade principais nessa metodologia, Dickey-Füller Aumentado (ADF) e KPSS.

O Teste ADF tem como hipótese nula a estacionariedade da série, ou seja, se $p < 0,05$ de probabilidade, a série é estacionária. Nesse teste a hipótese alternativa é de não-estacionariedade da série quando $p < 0,05$ de probabilidade. Conforme a Figura 3, o Teste ADF indica a não-estacionariedade da série de biodiesel, pois o valor de p foi maior que 5% de significância.

	t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic	-0.905721	0.7831
Test critical values:		
1% level	-3.491928	
5% level	-2.888411	
10% level	-2.581176	

*MacKinnon (1996) one-sided p-values.

Tabela 3 – Teste de Estacionariedade Dickey-Füller da Série Original, indicando não-estacionariedade para $p > 0,05$ de probabilidade.

O Teste KPSS se utiliza de hipóteses inversas ao Teste ADF, cujas hipóteses se apresentam: H_0 : Não há evidência empírica que ateste a estacionariedade na série, e H_1 : quando há evidência empírica de que a série é estacionária, desta forma para $p < 0,05$ de probabilidade. Na Figura 4, abaixo, pode-se compreender que o resultado do teste KPSS veio corroborar com os resultados obtidos no teste ADF, apresentando não-estacionariedade na série original do Biodiesel.

	LM-Stat.
Kwiatkowski-Phillips-Schmidt-Shin test statistic	1.16808...
Asymptotic critical values*:	
1% level	0.73900...
5% level	0.46300...
10% level	0.34700...

*Kwiatkowski-Phillips-Schmidt-Shin (1992, Table 1)

Tabela 4 – Teste de Estacionariedade KPSS da Série Original

Após a realização dos testes de estacionariedade e, da constatação que não será possível utilizar a série original para a modelagem, realizou-se a primeira diferença. Como pode-se ver, ambos os testes ADF e KPSS, indicaram a estacionariedade da série diferenciada. A Figura 5 mostra o resultado do teste ADF, cujo valor de p foi praticamente zero, mostrando a série diferenciada ser totalmente estacionária.

	t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic	-10.15713	0.0000
Test critical values:		
1% level	-3.491928	
5% level	-2.888411	
10% level	-2.581176	

*MacKinnon (1996) one-sided p-values.

Tabela 5 – Teste ADF da série com primeira diferença, resultado positivo para estacionariedade, com $p < 0,05$ de probabilidade

O teste KPSS da série diferenciada corroborou com o teste acima, o ADF. A Figura 6 traz o KPSS indicando a estacionariedade da série diferenciada, auxiliando na detecção da estacionariedade para a hipótese alternativa.

	LM-Stat.
Kwiatkowski-Phillips-Schmidt-Shin test statistic	0.09498...
Asymptotic critical values*:	
1% level	0.73900...
5% level	0.46300...
10% level	0.34700...

*Kwiatkowski-Phillips-Schmidt-Shin (1992, Table 1)

Tabela 6 – Teste KPSS da série original, indicando estacionariedade da série, resultando em um $p > 0,05$ de probabilidade

Os modelos concorrentes utilizados para representar a produção de biodiesel em cm^3 , estão no Quadro 1.

Modelo	p-value dos coeficientes	AIC	BIC
SARIMA(1,1,0)(1,1,0) ₁₂	$\phi_1: -0,215435$ $\Phi_1: -0,368503$	22,27926	22,33673
SARIMA(2,1,0)(1,1,0) ₁₂	$\phi_1: -0,294931$ $\phi_2: -0,354000$ $\Phi_1: +0,499022$	22,20982	22,29664
ARFIMA(1,d,0)	$d = 0,486387$ $\phi_1 = 0,634622$	22,635736	-
ARFIMA(1,d,1)	$d = 0,443528$ $\phi_1 = 0,988856$ $\theta_1 = -0,825915$	22,387545	-

Quadro 1 – Critérios de comparação para verificar o melhor modelo

O modelo selecionado foi um SARIMA (2,1,0)(1,1,0)₁₂, por apresentar melhores estatísticas de avaliação. Dentro da Análise da série realizou-se também testes para identificação de modelos ARFIMA, nisso se conseguiu chegar a dois modelos concorrentes visíveis, mas

devido aos critérios de escolha de modelos, realmente optou-se pelo autoregressivo sazonal. Pode-se visualizar o comportamento do modelo escolhido através da fac e facp da análise de resíduos, como segue:































Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob	
		1	-0.07...	-0.07...	0.5223	
		2	-0.26...	-0.27...	6.7659	
		3	0.175	0.139	9.5194	
		4	0.015	-0.03...	9.5411	0.002
		5	-0.17...	-0.10...	12.353	0.002
		6	-0.13...	-0.19...	13.945	0.003
		7	-0.04...	-0.15...	14.133	0.007
		8	0.043	-0.01...	14.314	0.014
		9	-0.03...	-0.05...	14.436	0.025
		10	0.044	0.050	14.628	0.041
		11	0.115	0.051	15.947	0.043
		12	-0.10...	-0.12...	16.973	0.049
		13	0.028	0.010	17.055	0.073
		14	0.256	0.210	23.904	0.013
		15	-0.21...	-0.14...	28.824	0.004

Figura 3 – FAC e FACP dos resíduos do modelo escolhido SARIMA (2,1,0)(1,1,0)₁₂

Com o intuito de realizar previsões para a produção brasileira de biodiesel no período de 06/2013 a 05/2014, fez-se a previsão in-sample, para 12 meses.

3. Conclusão

Atualmente, a União Europeia, em especial a Alemanha, os Estados Unidos, e o Brasil são os maiores mercados mundiais de Biodiesel. Hoje o Brasil conta com indústria de biodiesel consolidada, com mais de 50 usinas aptas a produzir e comercializar biodiesel, com uma capacidade instalada superior a 6 milhões de metros cúbicos.

Esse combustível renovável permite a economia de divisas com a importação de petróleo e óleo diesel e também reduz a poluição ambiental, além de gerar alternativas de empregos em áreas geográficas menos atraentes para outras atividades econômicas e, assim, promover a inclusão social. A disponibilização de energia elétrica para comunidades isoladas, hoje de elevado custo em função dos preços do diesel, também deve ser incluída como forma de inclusão, que permite outras, como a inclusão digital, o acesso a bens, serviços, informação, à cidadania e assim por diante.

Há que se considerar ainda uma vantagem estratégica que a maioria dos países importadores de petróleo vem inserindo em suas prioridades a redução da dependência das importações de petróleo. Deve-se enfatizar também que a introdução do biodiesel aumenta a participação de fontes limpas e renováveis em nossa matriz energética, somando-se principalmente à hidroeletricidade e ao álcool e colocando o Brasil numa posição ainda mais privilegiada nesse aspecto, no cenário internacional. A médio prazo, o biodiesel pode tornar-se importante fonte de divisas para o país, somando-se ao álcool como fonte de energia renovável oferecida pelo Brasil à comunidade mundial.

Estando entre os maiores produtores e consumidores de biodiesel do mundo, realizar previsão de produção desse biocombustível no Brasil torna-se uma atividade importante, pois pode revelar as tendências em um momento em que a contínua elevação do percentual de adição de biodiesel ao diesel demonstra o sucesso do Programa Nacional de Produção e Uso do Biodiesel na produção e no uso em larga escala de biocombustíveis.

A aplicação da metodologia de Box-Jenkins requer do pesquisador atenção e prática para escolher o modelo que melhor define o comportamento temporal dos dados. Com a aplicação dessa metodologia, apoiada em testes estatísticos para subsidiar a modelagem, foi possível definir modelos apropriados para previsão da produção. Desse modo, a principal contribuição deste artigo pode ser resumida em aspectos como a apresentação passo a passo do emprego da metodologia; o uso sequencial de gráficos e testes estatísticos para subsidiar a construção e a escolha dos modelos mais adequados; e a ênfase na modelagem orientada à previsão da produção de biodiesel no Brasil. Para atender ao objetivo do estudo, foi necessário utilizar um modelo contendo termos diferenciais e sazonais, SARIMA(2,1,0)(1,1,0)₁₂ para previsão da produção nacional desse biocombustível.

Referências

- ABREU, T., ARAÚJO, K. A. M., LOPES, M. L. M., LOTUFO, A. D. P. *Metodologia híbrida utilizando os modelos arima e redes neurais artificiais para previsão de cargas elétricas*. Anais do XIX Congresso Brasileiro de Automática, 2012.
- ANP. *Agência Nacional do Petróleo, Gás Natural e Biocombustíveis*. 2014.
- BOX, G.; JENKINS, G. M. *Time series analysis: Forecasting and control*. San Francisco: Holden-Day, 1970.
- CLEMENTE, A., CLEMENTE, L. T. *Aplicação da metodologia Box-Jenkins para previsão do ICMS do estado do Paraná de agosto de 2011 a julho de 2012*. Revista Economia e Tecnologia. Ano 07, Vol.27. Outubro/Dezembro 2011.
- FAVA, V. L. *Manual de econometria*. São Paulo: Editora Atlas, 2000.
- GUJARATI, D. N. *Econometria básica*. São Paulo: Makron books, 2000.
- KOHLHEPP, G. *Análise da situação da produção de etanol e biodiesel no Brasil*. Estudos Avançados, vol.24(68), p.223-253, 2011.
- LEE, D. H. *Algal biodiesel economy and competition among bio-fuels*. Bioresource Technology, vol.102(1), p.43-49, 2011.
- LEONARDI, A.; SCARTON, L.M.; PADULA, A.D.; CORONEL, D. A. *Biodiesel no Brasil: Uma estrutura de mercado determinada pela política de leilões*. Revista Brasileira de Gestão e Desenvolvimento Regional, vol.7(2), p.166-181, 2011.
- MORETTIN, P. A., TOLOI, C. M. C. *Análise de Séries Temporais*. 2ª Ed. – São Paulo: Blusher, 2006.
- PEREIRA, B. B.; PAIS, M. B. Z.; SALES, P. R. H. *Análise Espectral de séries temporais: Uma introdução para Engenharia, Economia e Estatística*. Rio de Janeiro: Arte Final Leasing Editorial; Eletrobrás S. A., 1986.
- SOUZA, F. M.; ALMEIDA, S.G.; SOUZA, A. M. *Previsão do preço da gasolina para a região sul do Brasil*. IJIE – Iberoamerican Journal of Industrial Engineering. Julho/2011.
- SOUZA, M. A.; SOUZA, M. F.; MENEZES, R. *Procedure to evaluate multivariate statistical process control using ARIMA-ARCH models*. Japan Industrial Management Association, v. 63, 2012.
- SOUZA, F. M.; ALMEIDA, S.G.; FELICIANE, A. V., SOUZA, A. M. *Uso de modelo de memória longa: previsão de equipamentos para a agroindústria*. IJIE – Iberoamerican Journal of Industrial Engineering. Julho/2011.
- WERNER, L., RIBEIRO, J. L. D. *Previsão de demanda: uma aplicação dos Modelos Box-Jenkins na área de assistência técnica de computadores pessoais*. Revista Gestão e Produção. Vol. 10, nº1, p.47-67, Abril/2003.