

MODELAGEM E PREVISÃO DA PRODUÇÃO BRASILEIRA DE BIODIESEL UTILIZANDO SÉRIES TEMPORAIS

ERIKA FIALHO MORAIS XAVIER¹, NYEDJA FIALHO MORAIS BARBOSA², SÍLVIO FERNANDO ALVES XAVIER JÚNIOR^{3*}, JADER DA SILVA JALE¹, TIAGO ALESSANDRO ESPÍNOLA FERREIRA¹

¹ Programa de Pós Graduação em Biometria e Estatística Aplicada (DEIFO), Universidade Federal Rural de Pernambuco.

² CCBSA, Universidade Estadual da Paraíba.

³ Depto de Estatística, Centro de Ciências e Tecnologia, Universidade Estadual da Paraíba, Campina Grande/PB.

* Autor para correspondência: profsilviofernando@gmail.com

Recebido em 25 de novembro de 2017. Aceito em 24 de dezembro de 2018. Publicado em 28 de dezembro de 2018.

RESUMO - Os biocombustíveis são, atualmente, as melhores alternativas para a redução do uso de combustíveis fósseis, que representam a maior parte da matriz energética brasileira. Oriundo principalmente a partir de biomassa vegetal, o biodiesel é um combustível limpo e renovável que pode substituir quase todos os derivados de petróleo sem modificações nos motores, diminuindo a dependência dos combustíveis fósseis e causando menores prejuízos ao meio ambiente, se comparado aos prejuízos ocasionados por combustíveis derivados do petróleo. Com base na preocupação ambiental e através de análises de séries temporais, foram propostos possíveis modelos de predição para a produção de brasileira de biodiesel, utilizando dados da produção nacional em m³ no período entre janeiro de 2005 e abril de 2018, disponibilizados livremente através do site da ANP - Agência Nacional do Petróleo, Gás Natural e Biocombustíveis. Os modelos de previsão foram estimados considerando a sazonalidade apresentada pela série temporal, através da metodologia de Box e Jenkins e do algoritmo de alisamento exponencial de Holt-Winters. Os modelos propostos responderam satisfatoriamente às expectativas, mostrando erros de predição relativamente pequenos, bem como um bom ajuste aos dados. O trabalho apresentou resultados que contribuem positivamente para a sociedade, na busca de recursos cada vez mais ecologicamente sustentáveis.

PALAVRAS-CHAVE: BIOCMBUSTÍVEIS, SARIMA, SARIMAX, HOLT-WINTERS

MODELING AND FORECASTING OF BRAZILIAN BIODIESEL PRODUCTION USING TEMPORARY SERIES

ABSTRACT - Biofuels are currently the best alternative for reducing the use of fossil fuels, which account for the major part of the Brazilian energy matrix. Mainly based on plant biomass biodiesel is a clean and renewable fuel that can replace almost all petroleum products without modification in the engines, reducing dependence on fossil fuels and causing less damage to the environment, compared to the damages caused by petroleum-based fuels. Based on environmental concern and through time series analysis, possible predictive models for biodiesel production in Brazil were proposed, using data from national production in m³ in the period between January 2005 and April 2018, freely available through the website of ANP – the Brazilian National Agency of Petroleum, Natural Gas and Biofuels. The prediction models were estimated considering the seasonality presented by time series, making use of Box and Jenkins methodology and the Holt-Winters exponential smoothing algorithm. The proposed models responded satisfactorily to expectations; showing relatively small prediction errors, as well as a good fit to the data. The work presented results that contribute positively to society, in the search for increasingly environmentally sustainable resources.

KEYWORDS: BIOFUELS, SARIMA, SARIMAX, HOLT-WINTERS

MODELADO Y PREVISIÓN DE LA PRODUCCIÓN BRASILEÑA DE BIODIESEL UTILIZANDO SERIES TEMPORALES

RESUMEN - Los biocombustibles son actualmente las mejores alternativas para la reducción del uso de combustibles fósiles, que representan una gran parte de la matriz energética brasileña. Derivado sobre todo de la biomasa vegetal, el biodiésel es un combustible limpio y renovable que puede sustituir casi todos los derivados del petróleo sin modificaciones en los motores, reduciendo la dependencia de los combustibles fósiles y causando menores daños al medio ambiente, si se compara con los daños ocasionados por combustibles derivados del petróleo. Con base en la preocupación ambiental y a través de análisis de series temporales, se propusieron posibles modelos de predicción para la producción brasileña de biodiésel, utilizando datos de la producción nacional en m³ en el período entre enero de 2005 y abril de 2018, disponibles libremente a través del sitio de la ANP - Agencia Nacional del Petróleo, Gas Natural y Biocombustibles. Se han estimado los modelos de previsión considerando la estacionalidad presentada por la serie temporal, a través de la metodología de Box y Jenkins y del algoritmo de suavizamiento exponencial de Holt-Winters. Los modelos propuestos respondieron satisfactoriamente a las expectativas, mostrando errores de predicción relativamente pequeños, así como un buen ajuste a los datos. El trabajo presentó resultados que contribuyen positivamente a la sociedad en la búsqueda de recursos cada vez más ecológicamente sostenibles.

PALABRAS CLAVE: BIOCMBUSTIBLES, SARIMA, SARIMAX, HOLT-WINTERS.

INTRODUÇÃO

Devido aos compromissos assumidos com a questão ambiental desde o Protocolo de Kyoto (http://unfccc.int/kyoto_protocol/items/2830.php), uma das maiores preocupações do mundo atual tem sido o suprimento de energia limpa para as próximas décadas, razão pela qual fizeram renascer a atenção quanto às fontes alternativas de energia (Souza *et al.* 2012). As principais energias renováveis são obtidas através de fontes renováveis, tais como energia solar, eólica, hidráulica e dos biocombustíveis. Em particular, esta última fonte, os biocombustíveis são a melhor alternativa para a redução do uso de combustíveis fósseis, os quais representam a maior parte da matriz energética brasileira (Carmo *et al.* 2009).

O biodiesel é um combustível renovável produzido principalmente a partir de fontes de biomassa vegetal e animal, associada ao metanol ou etanol (Freitas e Penteadó 2006). Sua principal vantagem é que, comparado a outros combustíveis provenientes de petróleo, causa menor poluição ao meio ambiente, pois trata-se de um combustível limpo e renovável que, produzido em larga escala, pode gerar custos mais baixos do que os combustíveis oriundos do petróleo.

Os resíduos e co-produtos gerados na produção do biodiesel podem ser aproveitados, fazendo com que toda a cadeia de produção seja sustentável e economicamente viável. Uma outra vantagem é que sua composição não possui enxofre, portanto sua queima é isenta de SO₂ e SO₃, óxidos causadores da chuva ácida (Mota e Pestana 2011). O biodiesel pode substituir quase todos os derivados de petróleo sem modificações nos motores, o que reduz a dependência dos combustíveis fósseis (Parente 2003; Domingos *et al.* 2012).

A proposta de utilização do biocombustível no Brasil existe desde 1920 (Osaki e Batalha 2011), entretanto, foi só a partir da crise do petróleo na década de 1970 que o governo federal criou o Programa Nacional do Álcool – PRÓÁLCOOL, com o decreto número 76.593 de 14 de novembro de 1975, incentivando a iniciação da produção do álcool oriundo da cana-de-açúcar como um biocombustível.

Em 13 de janeiro de 2005 o governo federal criou o Programa Nacional de Produção e uso do Biodiesel – PNPB, através da lei número 11.097, que estabeleceu percentuais obrigatórios de biodiesel no óleo diesel comercializado em território nacional, tendo como principais finalidades: a implantação de um programa sustentável para a produção de biocombustíveis no país, a garantia de preços competitivos e o aumento na produção do biodiesel. Diante desta informação, surge o questionamento: qual o impacto gerado na produção nacional do biodiesel a partir da implantação do PNPB? Neste sentido, o acúmulo de dados da produção nacional de biodiesel pode ser

explicado sequencialmente em função do tempo através da modelagem de séries temporais, técnica que permite o entendimento do padrão temporal de comportamento e descrição do fenômeno (Oliveira *et al.* 2015).

A aplicabilidade dos modelos de séries temporais pode ser vista nas mais diversas áreas de estudo (Tablada *et al.* 2016): na saúde (Bicalho *et al.* 2014), na meteorologia (Liska *et al.* 2013), nas ciências econômicas, entre outras, em virtude da dependência temporal apresentada pela maioria das variáveis (Barbosa *et al.* 2015). Uma das classes de modelos de séries temporais mais difundidos é a abordagem proposta por Box e Jenkins (Box e Jenkins 2015), que consiste em decompor a série temporal em componentes autorregressivos de médias móveis (Sharma *et al.* 2009), com o objetivo de analisar o comportamento, as tendências e as correlações dos dados observados (Oliveira *et al.* 2015) e, a partir desta análise, obter estimativas viáveis e confiáveis para o fenômeno em estudo. Serra (2013) realizou estudos de econometria em séries temporais analisando as interações de volatilidade entre os mercados de biocombustíveis e combustíveis fósseis.

Este trabalho tem por objetivo analisar a produção brasileira de biodiesel ao longo do tempo, através de técnicas estatísticas de análise de séries temporais estimando um possível modelo de previsão para a produção de biodiesel no Brasil no período entre 2005 e 2018.

MATERIAL E MÉTODOS

Os dados utilizados são referentes às unidades mensais em m^3 e compreendidos entre janeiro de 2005 e abril de 2018, totalizando 160 pontos. Os dados da produção nacional foram utilizados e estão disponíveis no site da ANP - Agência Nacional do Petróleo, Gás Natural e Biocombustíveis (<http://www.anp.gov.br/>).

Procedimento Box & Jenkins

O procedimento de análise proposto por Box e Jenkins (Box e Jenkins 2015) fundamenta-se na construção de funções baseadas em um ciclo iterativo que utiliza os próprios dados da série temporal para encontrar uma estrutura matemática de previsão. Este procedimento é feito inicialmente pela especificação do modelo através das funções de autocorrelação e autocorrelação parcial, seguido da estimação dos parâmetros do modelo, bem como do diagnóstico do modelo ajustado.

Um dos modelos mais populares entre a classe dos modelos Box-Jenkins é o modelo autorregressivo de ordem p e com médias móveis de ordem q , $ARMA(p, q)$, definido pela Equação (2.1.1)

$$Z_t = \mu + \sum_{i=1}^p \phi_i Z_{t-i} + a_t + \sum_{i=1}^q -\theta_i a_{t-i} \quad (2.0.1)$$

ou simplesmente

$$\Phi(B)Z_t = \Theta(B)a_t \quad (2.1.2)$$

onde a_t é o ruído branco e B é o operador de retardo temporal $BZ_t = Z_{t-1}$; $\Phi(B)$ é o polinômio autorregressivo em B dado por $\Phi(B) = 1 - \Phi_1 B - \Phi_2 B^2 - \dots - \Phi_p B^p$, e $\Theta(B)$ é o polinômio de médias móveis em B dado por $\Theta(B) = 1 - \Theta_1 B - \Theta_2 B^2 - \dots - \Theta_q B^q$.

Diz-se que uma série temporal é estacionária quando segue uma distribuição cujo todos os seus momentos sejam constantes ao longo do tempo. Contudo, é possível se definir o conceito de estacionariedade de ordem f onde do primeiro ao f -ésimo momento são constantes. Assim, de forma fraca, mais utilizada na literatura, considera-se no mínimo ordem $f = 2$; média e variância constantes no tempo pra se ter um processo estacionário. O modelo $ARMA(p, q)$ deve ser utilizado em séries estacionárias. Entretanto, para séries que apresentam alguma

forma de não estacionariedade, pode-se realizar o processo de integração de ordem d através da Equação (2.1.3), tomando diferenças sucessivas na série original até que seja obtida uma série estacionária.

$$\nabla^d Z_t = (1 - B)^d Z_t \quad (2.1.3)$$

Neste caso, o processo passa a ser definido por um modelo $ARIMA(p,d,q)$.

Séries temporais que apresentam flutuações sazonais podem ser estimadas utilizando um modelo $SARIMA(p,d,q) \times (P,D,Q)$, onde P indica a ordem autorregressiva sazonal, Q a ordem sazonal de médias móveis e D a diferenciação sazonal. A forma matemática para a modelagem $SARIMA(p,d,q) \times (P,D,Q)$ é definida pela Equação (2.1.4)

$$\Phi(B)\phi(B^s)[(1 - B)^d(1 - B^s)^D - \mu]Z_t = \theta(B)\Theta(B^s)a_t \quad (2.1.4)$$

onde s representa o período de sazonalidade, $(1 - B^s)Z_t = Z_t - Z_{t-s}$ define a diferenciação sazonal, $\Phi(B^s) =$ representa o polinômio autorregressivo com o período de sazonalidade e $\Theta(B^s) = -\sum_{i=1}^q \Theta_i B^{s \times i}$.

Pode-se ainda incluir nos modelos uma estrutura de regressão, obtendo-se assim um processo $SARIMAX$ com modelo definido através da Equação (2.1.5).

$$\Phi(B)\Phi(B_s)[(I - B)_q(I - B_s)_D]\Sigma^t = \pi + X\beta + \theta(B) + \Theta(B_s)\alpha^t \quad (2.1.5)$$

onde X é uma matriz $n \times k$ de regressores não aleatórios e β representa um vetor de parâmetros $k \times 1$.

Alisamento exponencial Holt-Winters

O alisamento exponencial de Holt (Holt 1957) é um algoritmo formado por componentes não observáveis de nível e tendência. O nível é uma estimativa suavizada do valor dos dados no final de cada período e a tendência é uma estimativa suavizada da variação da média no final de cada período. O alisamento exponencial de Holt-Winters (Winters 1960) é uma variação do algoritmo de Holt, que tem por objetivo principal a incorporação de movimentos sazonais.

Este algoritmo tem como base as fórmulas de recorrência, definidas para séries sazonais aditivas através das Equações (2.2.1).

$$\begin{aligned} N_t &= \alpha(Z_t - F_{t-s}) + (1 - \alpha)(N_{t-1} + T_{t-1}) \\ T_t &= \beta(N_t - N_{t-1}) + (1 - \beta)T_{t-1} \\ F_t &= \gamma(Z_t - N_t) + (1 - \gamma)F_{t-s} \end{aligned} \quad (2.2.1)$$

Onde α, β e γ são constantes de suavização, com $0 < \alpha < 1$, $0 < \beta < 1$ e $0 < \gamma < 1$. N_t , T_t e F_t representam os fatores de nível, tendência e sazonalidade, respectivamente. A previsão h passos a frente é definida através das Equações (2.2.2)

$$\begin{aligned} \hat{Z}_t(h) &= N_t + hT_t + F_{t+h-s}, h = 1, 2, \dots, s \\ \hat{Z}_t(h) &= N_t + hT_t + F_{t+h-2s}, h = s + 1, s + 2, \dots, 2s \end{aligned} \quad (2.2.2)$$

A forma de correção dos erros é definida através das Equações (2.2.3)

$$\begin{aligned} N_t &= N_{t-1} + T_{t-1} + \alpha e_t \\ T_t &= T_{t-1} + \alpha \beta e_t \\ F_t &= F_{t-s} + \gamma(1 - \alpha)e_t \end{aligned} \quad (2.2.3)$$

Deste modo, os aumentos no nível, na tendência e no fator de sazonalidade são controlados por a , $a\beta$ e $\gamma(1 - a)$, respectivamente.

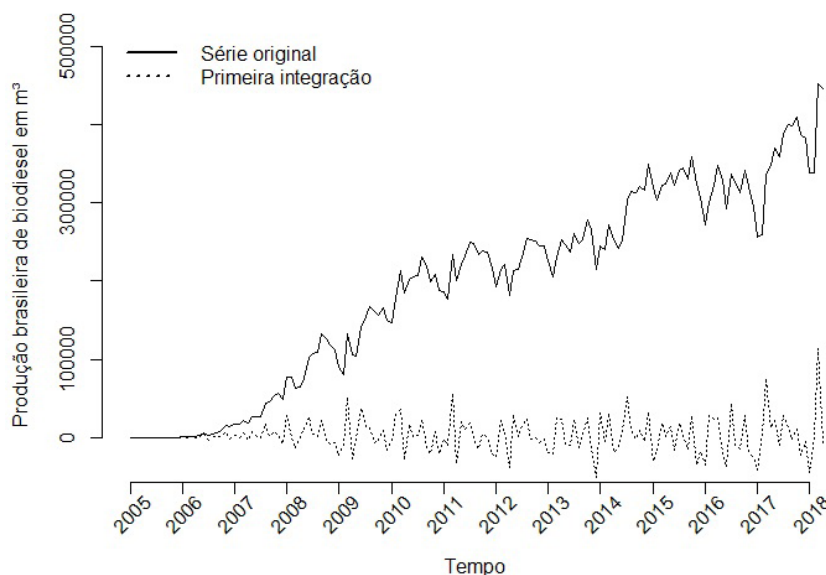
Maiores informações a respeito da modelagem de Box-Jenkins, bem como do alisamento exponencial de Holt-Winters, podem ser obtidas em Box e Jenkins (1976), Box *et al.* (1994), Morettin e Toloí (2004), Cryer e Chan (2008). As análises apresentadas foram feitas utilizando os pacotes *tseries* (Trapletti e Hornik 2016) e *forecast* (Hyndman e Khandakar 2008; Hyndman 2016) do Software R Core Team (2015).

RESULTADOS E DISCUSSÃO

A introdução do biodiesel na matriz energética brasileira ganhou maior destaque após janeiro de 2005, quando foi sancionada a Lei número 11.097, em conjunto com a criação do Programa Nacional de Produção e uso do Biodiesel – PNPB. O período entre os anos de 2005 e 2013 foi considerado como um período de adaptação às novas regras de produção do biodiesel, por esta razão, estipulou-se a adição de 2% de biodiesel no óleo diesel comercializado em território nacional entre os anos de 2005 a 2007, em caráter autorizativo. Este percentual passou a ser obrigatório entre os anos de 2008 a 2012. A partir de 2013, o percentual obrigatório da adição de biodiesel passou a ser de 5% em todo território nacional. A lei 13263 de 23 de março de 2016 (que altera a lei 13033 de 24 de setembro de 2014) estipulou o percentual obrigatório de 8%, 9% e 10%, para 12, 24 e 36 meses após a data de promulgação da lei, respectivamente. Com estas medidas, o aumento na produção nacional de biodiesel foi significativo e notório. Pode-se observar através da Figura 1 a crescente produção de biodiesel no Brasil, especialmente a partir de 2008, quando o percentual de 2% passou a ser obrigatório em território nacional. Neste ano, a produção brasileira de biodiesel ultrapassou os 50.000 m³ e apresentou uma tendência crescente nos anos seguintes. Além da relação com a legislação vigente, o aumento da produção de biocombustíveis pode também ser associado ao aumento da geração de eletricidade renovável (Ozturk 2016).

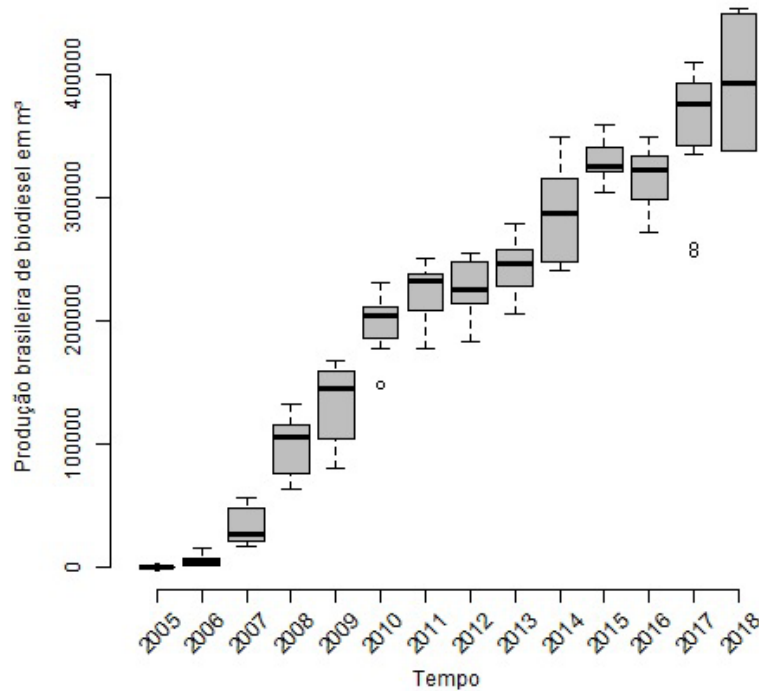
Há indícios da não estacionariedade da série para a produção brasileira de biodiesel no decorrer do tempo em estudo (Figura 1), o que pôde ser confirmado com o teste de Dickey-Fuller, que apresentou um p-valor = 0.2517. Os resultados do teste de raiz unitária em Ozturk (2016) indicam a mistura da ordem de integração, implicando em um estudo de um modelo não estacionário. No entanto, os autores não verificaram a classe de modelos SARIMA em busca de um modelo que se ajustasse à produção e previsão de biocombustíveis. Com uma transformação nos dados originais, realizou-se uma diferença na série original obtendo-se uma série estacionária.

Figura 1. Série temporal original e com uma integração para a produção brasileira de biodiesel em m³.



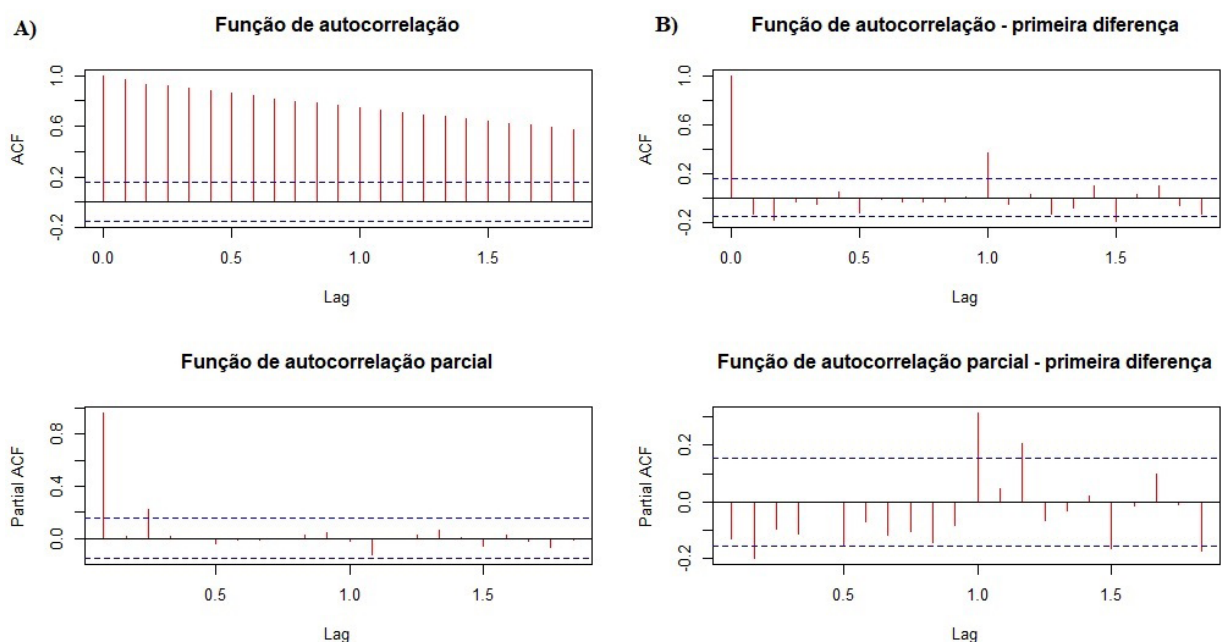
Os boxplots para a dispersão mensal da produção brasileira de biodiesel por ano evidenciam a variabilidade anual da produção (Figura 2). Observam-se maiores variabilidades nos anos 2008, 2009, 2014 e 2018, bem como as menores variabilidades nos anos 2005 e 2006. A série apresentou *outliers* nos anos de 2010 e 2017.

Figura 2. Gráficos Boxplot para dispersão mensal da produção nacional de biodiesel ao longo dos anos.



A Figura 3 apresenta as funções de autocorrelação e autocorrelação parcial da série. A Função de autocorrelação apresentou um decaimento lento, indicando a necessidade de diferenciação para estacionariedade da série; obtida através da primeira diferença. A análise gráfica para as funções de autocorrelação e autocorrelação parcial sugere que um bom ajuste pode ser obtido por um modelo Autorregressivo com Médias Móveis (ARMA), em virtude do decaimento da função de autocorrelação e do truncamento da função de autocorrelação parcial (Figura 3b).

Figura 3. A) Função de autocorrelação e autocorrelação parcial para a série original da produção nacional de biodiesel. B) Função de autocorrelação e autocorrelação parcial para a série da produção nacional de biodiesel com a primeira diferença.



A Figura 4 apresenta a decomposição da série em componentes não observáveis de sazonalidade, da tendência e aleatoriedade dos resíduos. Observa-se a existência de sazonalidade e tendência nos dados. Assim, são apresentados indícios que sugerem que o ajuste de modelos da classe SARIMA pode ser uma boa indicação para a modelagem da produção brasileira de biodiesel. A escolha do modelo para ajuste da série temporal foi feita com base nos critérios AIC (Akaike Information Criterion) e BIC (Bayesian Information Criterion). Os modelos que apresentaram menores critérios AIC e BIC foram o $SARIMAX(1,1,1) \times (1,2,1)$ e o $SARIMA(1,1,1) \times (1,2,1)$, respectivamente. A Tabela 1 apresenta os coeficientes para o modelo estimado.

Figura 4. Decomposição da série temporal de produção do biodiesel em componentes de sazonalidade, da tendência e aleatoriedade dos resíduos.

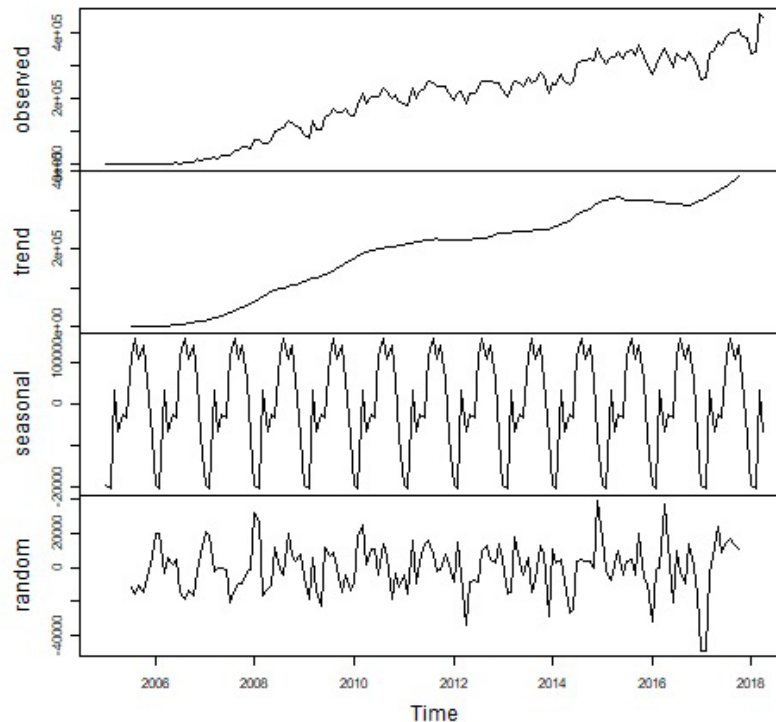


Tabela 1. Coeficientes do modelo SARIMA estimado.

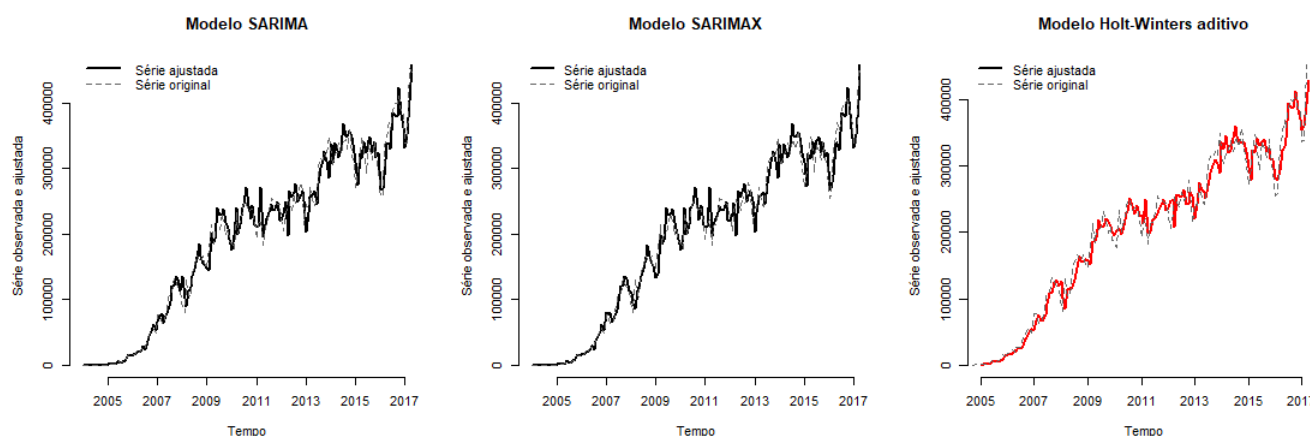
$SARIMA(1,1,1) \times (1,2,1)$		AIC = 3102.44		BIC = 3116.97	
	AR(1)	MA(1)	SAR(1)	SMA(1)	
Coef.	0.1492	-0.4661	-0.4833	-0.9940	
S. E.	0.2078	0.1807	0.0737	0.1087	

O sensível aumento na produção brasileira de biodiesel devido à criação do Programa Nacional de Produção e uso do Biodiesel – PNPB foi incorporado ao modelo de previsão através da criação de variáveis *dummy*. A primeira variável regressora considerou o aumento na produção a partir de janeiro de 2008, quando a adição de 2% de biodiesel no óleo diesel comercializado passou a ser obrigatória. A segunda variável considerou o aumento na produção a partir de janeiro de 2013, quando o percentual obrigatório passou a ser de 5%, a terceira variável foi quando o percentual obrigatório passou a ser de 10% a partir de setembro de 2017. A escolha dos modelos utilizando as variáveis regressoras foi feita novamente utilizando como critério de seleção os menores valores do AIC e BIC. A Tabela 2 apresenta os coeficientes dos dois modelos com menores índices AIC e BIC: o primeiro modelo considera o aumento a partir de janeiro de 2013 e o segundo modelo considera o efeito das duas variáveis regressoras em conjunto.

Tabela 2. Coeficientes do modelo SARIMAX estimado.

<i>SARIMAX</i>		AIC = 3101.25		BIC = 3118.68	
	AR(1)	MA(1)	SAR(1)	SMA(1)	DUM(1)
Coef.	0.1615	-0.4708	-0.4731	-0.9941	26119.44
S. E.	0.2248	0.1980	0.0750-	0.1083	14590.30

Foram considerados ainda modelos de previsão gerados a partir do alisamento exponencial do Holt-Winters. As constantes de suavização para o modelo aditivo foram $\alpha = 0.604$, $\beta = 0.011$ e $\gamma = 0.731$. Os ajustes da série temporal de produção nacional do biodiesel a partir de cada um dos modelos estimados podem ser vistos através da Figura 5.

Figura 5. Ajuste da série temporal de produção nacional de biodiesel através dos modelos SARIMA, SARIMAX, e do algoritmo de alisamento exponencial Holt-Winters aditivo.

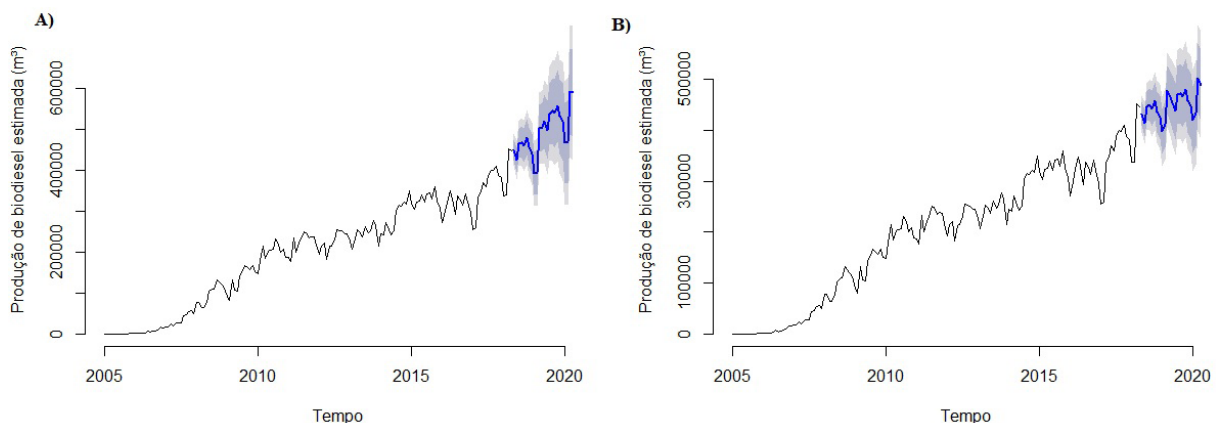
O erro relativo de previsão para as estimativas fornecidas pelos modelos foi calculado e utilizado como critério para avaliação dos ajustes. O modelo SARIMA apresentou um erro médio relativo de 3.2%, o modelo SARIMAX apresentou erro médio relativo de 3.18%. O menor erro médio relativo de previsão foi -0.88%, obtido através do algoritmo de alisamento exponencial de Holt-Winters aditivo. A Tabela 3 apresenta a produção brasileira de biodiesel observada para os meses janeiro, fevereiro, março e abril de 2018, bem como a produção estimada a partir dos modelos SARIMA, SARIMAX e Holt-Winters aditivo e os respectivos erros relativos de previsão. Os erros relativos de previsão para os modelos SARIMA, SARIMAX foram bastante próximos, enquanto o erro relativo de previsão obtido pelo algoritmo de Holt-Winters aditivo foi consideravelmente menor.

Tabela 3. Valores estimados, valores observados e erro relativo de previsão para a produção brasileira de biodiesel em m³ a partir dos modelos utilizados.

Modelo SARIMA			
Período	Prod. observada	Prod. estimada	Erro relativo de previsão
Jan / 2018	337818.12	332455.7	1.58%
Fev / 2018	338266.87	350022.0	-3.47%
Mar / 2018	452444.15	405472.3	10.38%
Abr / 2018	446177.27	426920.1	4.31%
Modelo SARIMAX			
Período	Prod. observada	Prod. estimada	Erro relativo de previsão
Jan / 2018	337818.12	332637.6	1.53%
Fev / 2018	338266.87	349908.4	-3.44%
Mar / 2018	452444.15	405702.4	10.33%
Abr / 2018	446177.27	427017.5	4.29%
Modelo Holt-Winters aditivo			
Período	Prod. observada	Prod. estimada	Erro relativo de previsão
Jan / 2018	337818.12	358505.0	-6.12%
Fev / 2018	338266.87	374494.7	-10.71%
Mar / 2018	452444.15	419332.8	7.32%
Abr / 2018	446177.27	419499.5	5.98%

A Figura 6 apresenta a previsão para a produção de biodiesel ajustada pelo modelo SARIMA e pelo algoritmo Holt-Winters aditivo. Pode-se observar através da análise gráfica que ambos os modelos estimam tendência de aumento na produção de biodiesel. Entretanto, o ajuste pelo algoritmo de Holt-Winters aditivo (Figura 6b) apresenta estimativas de crescimento um pouco maiores que o ajuste proposto pelo modelo SARIMA (Figura 6a). Em média, a série estimada através do modelo ajustado pelo algoritmo de Holt-Winters aditivo apresentou valores mais próximos da série observada. A tendência de aumento na produção de biocombustíveis também analisada em outros estudos, entre eles pode-se destacar Ozturk (2016), que utilizou a metodologia de regressão por mínimos quadrados confirmando que todos os indicadores sustentáveis têm uma associação significativa com a produção de biocombustíveis, pois o consumo total de energia primária aumenta a produção de biocombustíveis. Panichelli e Gnansounou (2015) consideraram diferentes opções de modelagem para examinar o impacto potencial da produção de biocombustíveis na mudança do uso da terra e nas emissões de gases de efeito estufa. Os resultados mostram implicações favoráveis que abordam a produção de biocombustíveis para mitigar as emissões de gases de efeito estufa e as mudanças no uso da terra.

Figura 6. A) Previsão para a produção nacional de biodiesel em m³ a partir do ajuste do modelo SARIMA, bem como B) do ajuste obtido pelo algoritmo de alisamento exponencial de Holt-Winters aditivo, com intervalos de confiança de 80% e 95% para as previsões.



As Figuras 7, 8 e 9 apresentam os gráficos de diagnóstico para verificação da qualidade do ajuste dos modelos estimados SARIMA, SARIMAX e Holt-Winters aditivo, respectivamente. O primeiro gráfico mostra os resíduos padronizados, em seguida temos o gráfico das autocorrelações residuais, que tem por finalidade conferir se os resíduos são não-correlacionados, e o terceiro é o gráfico dos p-valores para o teste Portmanteau de Ljung-Box, p-valores > 0.05 confirmam a evidência da não rejeição da hipótese de que o modelo está bem especificado.

Os gráficos dos resíduos padronizados evidenciam que os resíduos não apresentam valores altos no modelo SARIMA (Figura 7), nos modelos SARIMAX (Figura 8) e Holt-Winters aditivo (Figura 9). Nos gráficos das autocorrelações residuais, observa-se que os resíduos parecem ser correlacionados, pois as autocorrelações ultrapassam o limite de -0.2 e 0.2 . Na Figura 9 os p-valores para o teste de Ljung-Box, o modelo Holt-Winters aditivo apresentaram-se mais distantes do limite de confiança que os outros modelos.

Figura 7. Diagnóstico Modelo SARIMA – Gráfico dos resíduos padronizados, da função de autocorrelação dos resíduos e dos p-valores para o teste Portmanteau de Ljung-Box.

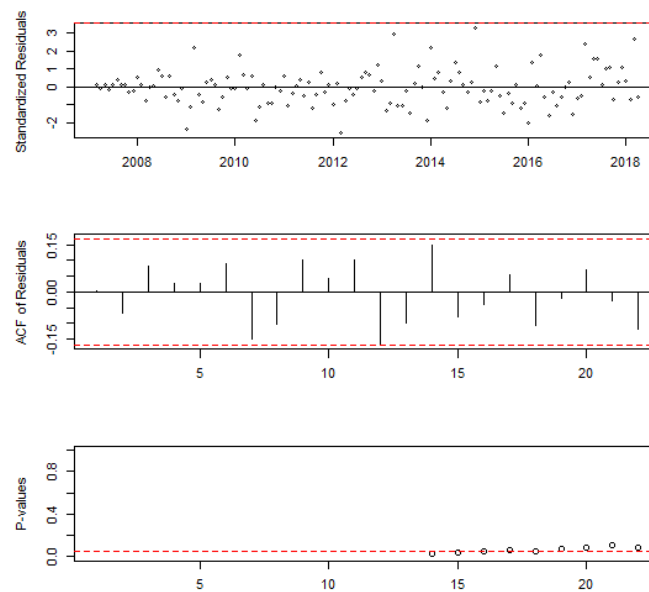


Figura 8. Diagnóstico Modelo SARIMAX – Gráfico dos resíduos padronizados, da função de autocorrelação dos resíduos e dos p-valores para o teste Portmanteau de Ljung-Box.

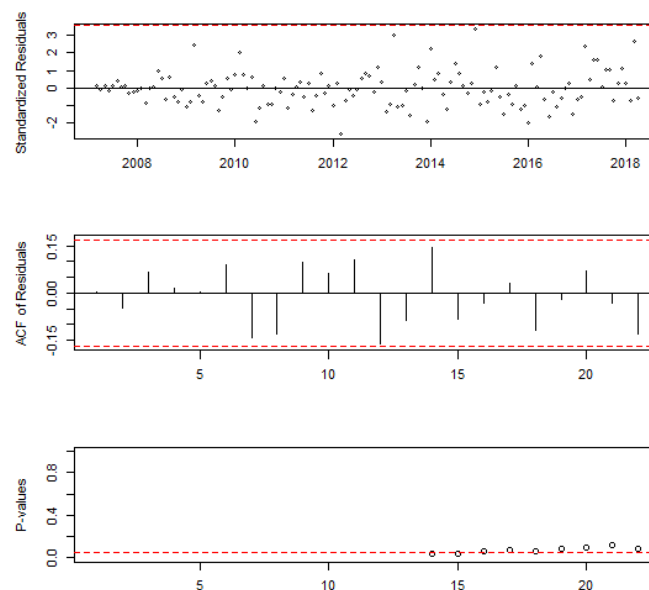
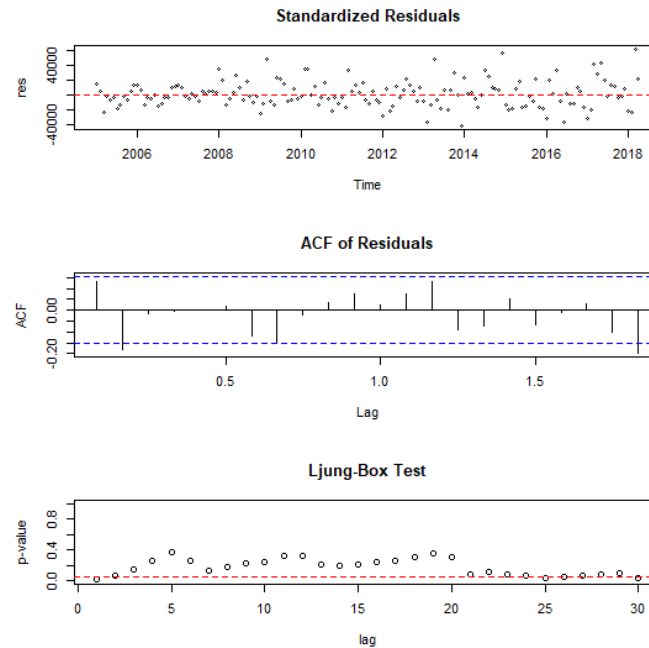


Figura 9. Diagnóstico Modelo Holt-Winters aditivo – Gráfico dos resíduos padronizados, da função de autocorrelação dos resíduos e dos p-valores para o teste Portmanteau de Ljung-Box.



CONCLUSÃO

A partir das análises da série de dados da produção de biodiesel no Brasil, entre janeiro de 2005 e abril de 2018, foi possível estimar um modelo $SARIMA(1,1,1) \times (1,2,1)$, com $AIC = 3102.44$ e $BIC = 3116.97$. Duas variáveis *dummy* foram utilizadas como variáveis regressoras para os períodos de janeiro de 2008, quando o percentual de 2% biodiesel obrigatório nos biocombustíveis passou a ser implementado em território nacional, e janeiro de 2013, quando o percentual obrigatório passou a ser de 5%. Estas duas variáveis regressoras forneceram os modelos SARIMAX, com $AIC = 3101.25$ e $BIC = 3118.68$. Foi proposto ainda um ajuste para a produção nacional do biodiesel por meio do algoritmo de alisamento exponencial de Holt-Winters aditivo.

Os modelos foram ajustados aos dados e forneceram estimativas para a produção nacional de biodiesel para os meses de janeiro, fevereiro, março e abril de 2018. Os valores estimados foram comparados aos valores observados, gerando estimativas para os erros relativos de previsão dos modelos. O modelo Holt-Winters aditivo forneceu, em média, os menores erros relativos de previsão, seguido do modelo SARIMAX. Deste modo, pode-se concluir que os melhores ajustes foram obtidos por meio dos modelos Holt-Winters aditivo. Não obstante, os modelos SARIMA e SARIMAX apresentaram valores para AIC, BIC e erros relativos de previsão. O modelo Holt-Winters aditivo mostrou-se mais adequado para o fornecimento de estimativas da produção nacional do biodiesel, de acordo com os critérios utilizados para verificação dos ajustes: as análises do erro relativo de previsão, juntamente com os resíduos padronizados, o gráfico das autocorrelações residuais, bem como o gráfico dos p-valores para o teste de Ljung-Box.

Assim, o trabalho apresenta resultados que indicam a tendência de aumento na produção do biodiesel, o que contribuirá positivamente, em conjunto com políticas públicas na busca de recursos cada vez mais ecologicamente sustentáveis, bem como a minimização de impactos ambientais causados pelos combustíveis usuais.

REFERÊNCIAS

- Babosa EC, Sáfadi T, Nascimento M, Nascimento ACC, Silva CHO e Manuli RC. 2015. Metodologia Box & Jenkins para a previsão de temperatura média mensal da cidade de Bauru (SP). **Rev. Bras. Biom.** 33(2): 104 – 117.
- Bicalho CC, Sáfadi T e Charret IC. 2014. The influence of climatic factors on dengue epidemics in the cities Cuiabá (Mato Grosso state) and Lavras (Minas Gerais state), Brazil, using statistical methods. **Rev. Bras. Biom.** 32(2): 308 – 322.
- Box GEP, Jenkins GM, Reinsel GC e Ljung GM. 2015. **Time Series Analysis: Forecasting and Control**. 15 ed. Wiley: San Francisco, 712p.
- Box GEP, Jenkins GM e Reinsel GC. 1994. **Time Series Analysis: Forecasting and Control**. 2 ed. Prentice-Hall: New York, 500p.
- Carmo BBT, Pontes HLJ, Albertin MR, Barros Neto JF, Dutra NGS. 2009. Avaliação da Demanda por Biodiesel em Função de um Modelo de Previsão de Demanda por Diesel. **Revista Produção On Line**, 9(3): 511 – 535.
- Cryer JD e Chan KS. 2008. **Time Series Analysis: With Applications in R**. Springer-Verlag: New York, 303p.
- Domingos CA, Pereira DD, Cardoso LS, Teodoro RA, Castro VA. 2012. **Biodiesel – Proposta de Um Combustível Alternativo**. Revista Brasileira de Gestão e Engenharia, 5: 134 – 178.
- Freitas C e Penteadó M. 2006. **Biodiesel – Energia do Futuro**. Letra Boreal, São Paulo, 353p.
- Holt CC. 1957. Forecasting Trends and Seasonal by Exponentially Weighted Moving Averages. **Carnegie Institute of Technology**, 2(52).
- Hyndman RJ. 2016. **Forecasting functions for time series and linear models**. R package version 7.1.
- Hyndaman RJ e Khandakar Y. 2008. Automatic time series forecasting: the forecast package for R. **Journal of Statistical Software**, 26(3):1-22.
- Liska GR, Bortolini J, Sáfadi T, Beijo LA. 2013. Estimativas da velocidade máxima de vento em Piracicaba – SP via séries temporais e teoria de valores extremos. **Rev. Bras. Biom.**, 31(2): 295 – 309.
- Morettin PA e Toloi CMC. 2004. **Análise de Séries Temporais**. ABE – Projeto Fisher e Edgard Blucher, 564p.
- Mota CJA e Pestana CFM. 2011. Co-produtos da produção de biodiesel. **Revista Virtual de Química**, 3(5): 416 – 425.
- Oliveira MRG, Cantalice JRB, Ferreira TAE, Filho MC, Cruz DV, Falcão APST. 2015. Estudo estatístico do coeficiente de escoamento da bacia hidrográfica do Riacho Jacu no Sertão do Pajeu – PE. **Rev. Bras. Biom**, 33(3): 277 – 290.
- Osaki M e Batalha MO. 2011. Produção de biodiesel e óleo vegetal no Brasil: Realidade e desafio. **Organizações Rurais & Agroindustriais**, 13(2): 227 – 242.
- Ozturk, I. 2016. Utilizing biofuels for sustainable development in the panel of 17 developed and developing countries. **Gcb Bioenergy**, 8(4): 826 – 836.

- Panichelli, L e Gnansounou, E. 2015. Impact of agricultural-based biofuel production on greenhouse gas emissions from land-use change: key modelling choices. **Renewable and Sustainable Energy Reviews**, 42: 344 – 360.
- Parente EJS. 2003. **BIODIESEL: Uma Aventura Tecnológica num País Engraçado**. Tecbio, Fortaleza, 66p.
- R Core Team. 2015. R: **A language and environment for statistical computing**. R Foundation for Statistical Computing, Vienna, Austria. <https://www.R-project.org/>.
- Serra T. 2013. Time-series econometric analyses of biofuel-related price volatility. **Agricultural Economics**, 44(1): 53 - 62.
- Sharma P, Chandra A e Kaushik SC. 2009. Forecasts using Box–Jenkins models for the ambient air quality data of Delhi City. **Environ Monit Assess**, 157: 105 – 112.
- Souza O, Schulz MA, Fischer GAA, Wagner TM, Sellin N. 2012. Energia Alternativa de Biomassa: Bioetanol a Partir da Casca e da Polpa de Banana. **Revista Brasileira de Engenharia Agrícola e Ambiental**, 16(8): 915 – 921.
- Tablada CJ, Ramírez FAP, Vieira GIA, Vasconcelos JM. 2016. Modelagem e previsão de exportações do Brasil fazendo uso de séries temporais. **Rev. Bras. Biom.** 34(1): 33 – 48.
- Trapletti A e Hornik K. 2016. tseries: **Time Series Analysis and Computational Finance**. R package version 0.10-35.
- Winters, PR. 1960. Forecasting sales by exponentially weighted moving averages. **Management Science**, 6: 324 – 342.