

Upgrade de previsão: comportamento da produção de biodiesel brasileiro por meio de modelagem Box & Jenkins

O processo produtivo de biodiesel está distribuído por diversos países pelo mundo, com legislação própria e estratégia comercial definida em cada um deles. Essa disseminação pelos continentes facilita a entrada do produto na lista de commodities mundiais, favorecendo o desenvolvimento de seu mercado em uma escala global. Dessa forma, realizou-se a conceitualização da metodologia Box & Jenkins, a fim de 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 contexto internacional, se tornando um dos maiores produtores de biodiesel do mundo. Considerou-se nessa análise, inclusive, a variação produtiva ocorrida em virtude da pandemia de Covid-19 iniciada em março de 2020. A escolha de modelos de previsão concorrentes, mais adequados para a previsão, baseou-se em análise gráfica e testes estatísticos próprios da metodologia, os quais subsidiaram a decisão de adotar o modelo SARIMA (1,1,2) (1,1,0)12 para previsão da produção nacional dessa modalidade de biocombustível em m³. Esse estudo vem sofrendo atualizações, desde 2014, quando de sua primeira versão e publicação em congresso, para que se fosse possível atingir muita análise da modelagem perpassando os anos, conseguindo captar as influências externas mais recentes.

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

Forecast upgrade: behavior of brazilian biodiesel production through Box & Jenkins modeling

The biodiesel production process is spread across several countries around the world, with their own legislation and commercial strategy defined in each of them. This spread across continents facilitates the entry of the product into the list of world commodities, favoring the development of its market on a global scale. Thus, the Box & Jenkins methodology was conceptualized to analyze historical biodiesel production data and obtain forecasts of the amount produced in Brazil, which, with the rapid evolution of its biodiesel program, is inserted in the international context, becoming one of the largest biodiesel producers in the world. This analysis also considered the production variation that occurred due to the Covid-19 pandemic that started in March 2020. The choice of competing forecast models, more suitable for prediction, was based on graphic analysis and own statistical tests methodology, which supported the decision to adopt the SARIMA model (1,1,2) (1,1,0)12 to forecast the national production of this type of biofuel in m³. This study has been updated since 2014, when it was first published and published in a congress, so that it would be possible to achieve a high degree of modeling analysis over the years, managing to capture the most recent external influences.

Keywords: Box-Jenkins Models; SARIMA Models; Time Series; Biodiesel Production.

Topic: **Pesquisa Operacional**

Received: **01/12/2022**

Approved: **25/11/2022**

Reviewed anonymously in the process of blind peer.

Renata Rocha de Oliveira 
Universidade Federal de Santa Maria, Brasil
<http://lattes.cnpq.br/8846571240091795>
<https://orcid.org/0000-0002-2057-0217>
renata.oliveira@ufsm.br

Verônica Dalmolin Cattelan 
Universidade Federal de Santa Maria, Brasil
<http://lattes.cnpq.br/8984050070043680>
<https://orcid.org/0000-0002-9508-6497>
veronica_vdc@hotmail.com

Luis Felipe Dias Lopes 
Universidade Federal de Santa Maria, Brasil
<http://lattes.cnpq.br/1074372911061770>
<https://orcid.org/0000-0002-2438-0226>
lflopes67@yahoo.com.br



DOI: 10.6008/CBPC2179-684X.2022.004.0012

Referencing this:

OLIVEIRA, R. R.; CATTELAN, V. D.; LOPES, L. F. D.. Upgrade de previsão: comportamento da produção de biodiesel brasileiro por meio de modelagem Box & Jenkins. **Revista Brasileira de Administração Científica**, v.13, n.4, p.164-176, 2022. DOI: <http://doi.org/10.6008/CBPC2179-684X.2022.004.0012>

INTRODUÇÃO

O surgimento do biodiesel ocorreu no início do século XX, durante a Exposição Mundial de Paris, em 1900, quando Rudolf Christian Karl Diesel (1858 – 1913), inventor dos motores que levam seu nome, testou com sucesso um combustível à base de óleo de amendoim, que apresentou um desempenho equivalente ao do combustível fóssil à base de petróleo. A partir de 1970, com a elevação do preço do petróleo, o biodiesel se tornou uma ideia viável.

O início dos estudos para desenvolvimento de uma política para utilização do biodiesel no Brasil tivera início no ano de 2003, acompanhados da criação de uma Comissão Executiva Interministerial do Biodiesel (CEIB) e do Grupo Gestor (GG) pelo governo federal (ANP, 2014) (ANP, 2017) (ANP, 2022). Em 13 de janeiro de 2005, houve a introdução do biodiesel na matriz energética brasileira por meio da lei nº 11.097, após um período experimental em 2004, todavia, de 2005 a 2007, a comercialização passou a ser voluntária.

O biodiesel caracteriza-se por ser um combustível derivado de biomassa renovável que se obtém por meio do processo de transesterificação, sendo esse um processo amplamente, para produzir biocombustíveis, utilizando como matéria prima os óleos vegetais e a gordura de origem animal (TAPANES et. al, 2013) (ANP, 2014) (ANP, 2017) (ANP, 2022). Os principais biocombustíveis líquidos produzidos e utilizados no Brasil são o etanol, extraído de cana-de-açúcar e, em escala crescente, o biodiesel, produzido a partir de óleos vegetais ou de gorduras animais adicionados ao diesel de petróleo em proporções variáveis. Ambos são derivados de fontes renováveis podendo substituir, parcial ou totalmente, os combustíveis derivados de petróleo e gás natural em motores à combustão ou em outro tipo de geração de energia (ANP, 2014) (ANP, 2017) (ANP, 2022).

Os biocombustíveis são considerados menos poluentes, pois emitem uma quantidade menor de gases do efeito estufa (GEE) que os combustíveis fósseis no processo de combustão dos motores e porque seu processo de produção tende a ser mais limpo (LEE, 2011) (LEONARDI et. al, 2011) (KOHLHEPP, 2011) (ANP, 2014) (ANP, 2017) (ANP, 2022). Pioneiro mundial no uso de biocombustíveis, o Brasil alcançou uma posição almejada por muitos países que buscam fontes renováveis de energia como alternativa estratégica ao petróleo. Nesse contexto, a ANP (2014) relata que cerca de 45% da energia e 18% dos combustíveis consumidos no Brasil provém de fontes renováveis enquanto no restante do mundo, 86% da energia vem de fontes não-renováveis.

Na comparação com o diesel de petróleo, o biodiesel apresenta significativas vantagens ambientais. Estudos do National Biodiesel Board demonstraram que a queima de biodiesel pode emitir em média 48% menos monóxido de carbono; 47% menos de material particulado; 67% menos de hidrocarbonetos. Como esses percentuais variam de acordo com a quantidade de B100 (biodiesel 100% puro, sem mistura de diesel mineral) adicionado ao diesel de petróleo, no B3 (biodiesel composto por 97% de diesel mineral e 3% de biodiesel) essas reduções ocorrem de modo proporcional (ANP, 2014) (ANP, 2017) (ANP, 2022).

A partir da lei que introduziu o biodiesel na matriz energética brasileira, houve a ampliação da competência administrativa da Agência Nacional do Petróleo, Gás Natural e Biocombustíveis (ANP),

assumindo as atribuições de especificar e fiscalizar a qualidade dos biocombustíveis e garantir o abastecimento do mercado. 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) (ANP, 2017) (ANP, 2022).

Tendo em vista a crescente produção de biodiesel no país desde 2005, a partir de março de 2020 até os dias atuais, o ano de 2021 acabou por ser um período de transformação. Pela análise dos dados extraídos da ANP, pode-se constatar que a produção reduziu a metade, considerando os fatores pandêmicos como: processo de vacinação mundial em andamento, alta taxa de mortalidade e contaminação, isolamento social, período pós 2020 de readaptação das rotinas sociais no mundo. A pandemia do Coronavírus afetou o mundo e teve seu início em novembro de 2019, na província de Wuhan, na China (WHO, 2022).

O artigo tem por objetivo reanalisar dados históricos de produção de biodiesel visando à obtenção de previsões da quantidade produzida desse biocombustível no Brasil por meio da metodologia de Box & Jenkins. Justifica-se o estudo a partir dessa abordagem por apresentar a escolha dados interessantes ao estudo de modelos adequados para variável produção de biodiesel (em metros cúbicos) podendo subsidiar decisões no cenário brasileiro, verificando-se também o comportamento da série em relação a sua tendência e sazonalidade. Esse trabalho vem sendo atualizado, desde 2014, quando de sua primeira versão e publicação em congressos, a fim de que se pudesse atingir muita análise da modelagem através do tempo, considerando as variabilidades de acontecimentos que afetaram o mundo.

Com a finalidade de atender ao objetivo desse trabalho relacionado a determinar um modelo adequado para previsão de produção de biodiesel no Brasil, será realizada uma previsão, a partir dessa revisão de dados, para um upgrade de análise a partir de 11/2022 até 10/2023. Desta forma, visou-se identificar o modelo de previsão sobre o comportamento da série, empregando a metodologia de Box & Jenkins para as quantidades de produção mensal.

DISCUSSÃO TEÓRICA

Modelagem Box & Jenkins

Técnica quantitativa induzida a partir da década de 70, a metodologia de Box & Jenkins parte da ideia de que cada valor de uma série (temporal) pode ser explicado por valores prévios, a partir do uso da estrutura de correlação temporal, a qual geralmente existe entre valores da série (BOX et al., 1970). Estes modelos têm sido largamente utilizados para modelagem e previsão em aplicações médicas, ambientais, financeiras e de engenharia (PEREIRA et. al, 1986) (FAVA, 2000) (WERNER et al., 2003) (MORETTIN et al., 2006). Essa técnica é fundamentalmente um procedimento em três etapas, sendo elas: a identificação do modelo, estimativa do modelo e a verificação diagnóstica para construir modelos de autorregressivos integrados de médias móveis (BOX et al., 1976), conforme figura 1:

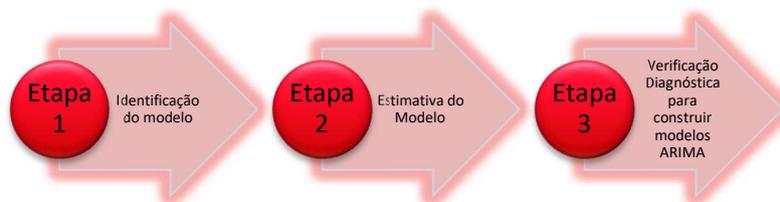


Figura 1: Etapas da Análise de Séries Temporais por Box & Jenkins **Fonte:** Adaptado de Box et al. (1976)

Morettin et al. (2006), consideram uma série temporal qualquer conjunto de observações ordenadas no tempo e definem que essas séries são compostas por quatro 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 à variabilidade intrínseca dos dados, não podendo ser modelado). Por meio da modelagem Box & Jenkins, estudaram-se modelos de séries temporais denominados ARIMA (autorregressivos integrados e de médias móveis) e suas variações.

De acordo com Morettin et al. (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 a identificação do 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 à fase inicial de identificação (MORETTIN et al., 2006).

Modelos Não-Estacionários

Para Werner et al. (2003), a não-estacionariedade acontece 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 et al., 2003).

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 os testes estatísticos de raiz unitária. No caso desse trabalho utilizaram-se os testes de Dickey-Fuller Aumentado (ADF) e KPSS.

Modelos Autoregressivos Integrados de Médias Móveis (ARIMA)

De acordo com Morettin et al. (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

definam-se se $W_t = \Delta^d Z_t$ for estacionária, pode-se representar W_t por um modelo ARMA (p,q), ou seja, $\phi(B)W_t = \theta(B)a_t$ (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, $\phi(B)\Delta^d Z_t = \theta(B)a_t$ (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 $\varphi(B)Z_t = \theta(B)a_t$ (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, $\varphi(B) = \phi(B)\Delta^d = \phi(B)(1 - B)^d$ (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 $d > 1$. Dessa forma, Morettin et al. (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 à oscilação 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 torna-se necessária uma segunda diferença.

Modelos Sazonais

Para Werner et al. (2003) e Namin et al. (2021), 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 que contemplam as séries que apresentam autocorrelação sazonal são chamados de SARIMA (WERNER et al., 2003; NAMIN et al., 2021).

Para Namin et al. (2021), existem características muito específicas da modelagem ARIMA, tais como: São modelos robustos; fáceis de aplicar; existe uma ampla utilização, desde engenharias até economia, contabilidade e administração; representa, de forma compacta, uma pequena quantidade de parâmetros; Difícilmente os parâmetros utilizados se atualizam com o upload de novos dados; e demandam uma grande quantidade de dados (PANKRATZ, 1982) (BISGAARD, 2011) (NAMIN et al., 2021).

O que se refere ao presente estudo, devida a inclusão de cinco anos de observações em dados, ocorreu alteração do modelo, encaixando-se como a exceção da quarta característica citada por Jafariam-Namin et al., (2021).

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 et al. (2006), quando $\{Z_t\}$ exibe um comportamento sazonal determinístico com período 12, um modelo pode ser utilizado: $Z_t = \mu_t + N_t$ (5)

onde μ_t é uma função determinística periódica, satisfazendo $\mu_t - \mu_{t-12} = 0$, ou $(1 - B^{12})\mu_t = 0$ (6) e N_t é um processo estacionário que pode ser modelado por um ARMA(p,q) (MORETTIN et al., 2006).

Assim, N_t satisfaz a equação: $\phi(B)N_t = \theta(B)a_t$ (7). Dessa forma, para

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

$(1 - B^{12})Z_t = (1 - B^{12})\mu_t + (1 - B^{12})N_t$ (8) e de acordo com (6), têm-se $(1 - B^{12})Z_t = (1 - B^{12})N_t$.

(9) Substituindo (7) em (9), obtém-se $\phi(B)(1 - B^{12})Z_t = \theta(B)(1 - B^{12})a_t$ (10) Ou

$\phi(B)W_t = \theta(B)(1 - B^{12})a_t$, (11) onde $W_t = (1 - B^{12})Z_t$ (MORETTIN et al., 2006).

Modelos de Memória Longa: ARFIMA

Essa família de processos, para fins de modelagem das 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 $(1 - B)^d = \sum_{k=0}^{\infty} \binom{d}{k} (-B)^k$ (12) $= 1 - dB + \frac{1}{2!}d(d-1)B^2 - \frac{1}{3!}d(d-1)(d-2)B^3 +$

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

$\{a_t\}$ é ruído branco e $\phi(B)$ e $\theta(B)$ são polinômios em B de graus p e q, respectivamente (MORETTIM et al., 2006) (SOUZA et al, 2010) (SOUZA et al., 2011) (OLIVEIRA, et al., 2015).

METODOLOGIA

O conjunto de dados considerado nesse artigo corresponde a 212 observações mensais de produção de biodiesel em m³ (metros cúbicos) no Brasil, no período de março de 2005 a outubro de 2022. A análise foi realizada por meio dos softwares Microsoft Excel 2010, Eviews 12 e Statistica 10 Enterprise. A coleta e extração dos dados realizaram-se de forma secundária, por meio do site da ANP pela lei de acesso à informação¹.

Inicialmente, analisa-se o gráfico do histograma, a fim de verificar a necessidade de aplicação de transformação dos dados, em casos de alta variabilidade, e existência de efeito sazonal. Posteriormente,

¹ <https://www.gov.br/anp/pt-br/centrais-de-conteudo/dados-estatisticos>

verifica-se o gráfico da série original e da série em primeira diferença para identificar em qual poderia ser realizada a modelagem, através dos testes de estacionariedade de Dickey- Fuller e KPSS. Para identificar a presença de tendência e sazonalidade, realizam-se as análises da função de autocorrelação (FAC) e função de autocorrelação parcial (FACP). Por meio dela, juntamente com o gráfico da série, sugere-se alguns modelos concorrentes.

Após a estimação dos modelos concorrentes, realiza-se a análise de resíduos para seja identificada a existência de comportamento ruído branco, que como pressuposto de um modelo adequado, os resíduos também devem se comportar de forma estacionaria.

RESULTADOS

A análise da figura 2 auxilia a identificação dos possíveis modelos utilizados. A figura mostra a tendência crescente de volume de produção (m³), ao longo dos 12 anos, indicando não-estacionariedade da série. Após, por meio da figura 2, realizou-se a análise das funções de autocorrelação (FAC) e autocorrelação parcial (FACP), da série original, possibilitando a visualização dos modelos possíveis juntamente com a figura 2.

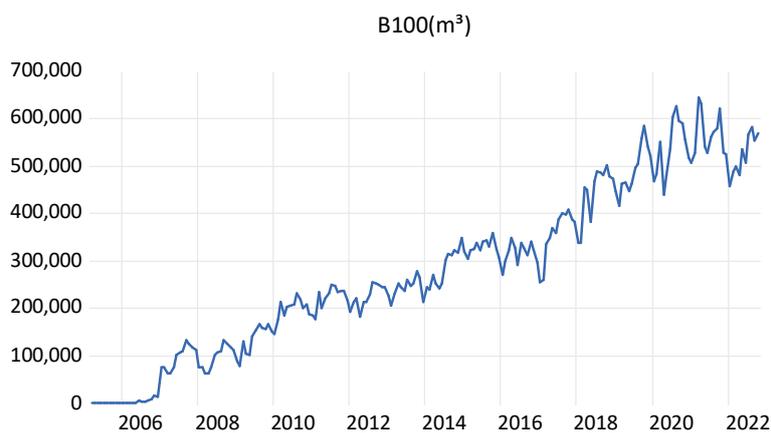


Figura 2: Série de Produção de Biodiesel.

A figura 3, 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 se obteria um modelo AR(1) por haver apenas um lag discrepante e bem definido, indicando que esse seria o modelo um modelo concorrente adequado.

Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob
1	0.973	0.973	203.57	0.000	
2	0.950	0.052	398.36	0.000	
3	0.930	0.072	586.23	0.000	
4	0.912	0.023	767.75	0.000	
5	0.899	0.094	944.94	0.000	
6	0.883	-0.040	1116.8	0.000	
7	0.874	0.118	1285.7	0.000	
8	0.863	-0.011	1451.2	0.000	
9	0.852	0.028	1613.4	0.000	
10	0.843	0.017	1773.0	0.000	
11	0.833	0.012	1929.5	0.000	
12	0.823	-0.004	2083.2	0.000	

Figura 3: FAC e FACP da Série Original de Biodiesel.

Após a análise da FAC e FACP da série original, realizaram-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 tabela 1, o teste ADF indica a não-estacionariedade da série de biodiesel, pois o valor de p foi maior que 5% de significância.

Tabela 1: Teste de Estacionariedade Dickey-Füller da Série Original, indicando não-estacionariedade para $p > 0,05$ de probabilidade.

	t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic	-0.890922	0.7897
Test critical values:		
1% level	-3.461630	
5% level	-2.875195	
10% level	-2.574125	

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

O teste KPSS utiliza-se 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, para $p < 0,05$ de probabilidade. Na tabela 2 é possível compreender que o resultado do teste KPSS não corrobora os resultados obtidos no teste ADF, pois indica estacionariedade na série original do Biodiesel, tendo $p > 0,05$. A relevância dessa observação faz com que não haja unanimidade em considerar a série estacionária, ou não, visto que em ambos os testes são analisados apenas duas hipóteses.

Tabela 2: Teste de Estacionariedade KPSS da Série Original.

	LM-Stat.
Kwiatkowski-Phillips-Schmidt-Shin test statistic	1.806025
Asymptotic critical values*:	
1% level	0.739000
5% level	0.463000
10% level	0.347000

Assim, após a realização dos testes de estacionariedade e, da constatação de que não é possível utilizar a série original para a modelagem, realizou-se a primeira diferença na série, a fim de avaliar a possibilidade de estacionariedade em ambos os testes de raiz unitária. Como se pode visualizar, com a primeira diferença, os testes ADF e KPSS indicaram a estacionariedade da série diferenciada. A tabela 3 evidencia o resultado do teste ADF, a qual apresenta o p -valor zerado, certificando a série diferenciada como estacionária.

O teste KPSS da série diferenciada corroborou com o teste ADF. A tabela 4 composta do teste KPSS indica a estacionariedade da série diferenciada, auxiliando na detecção da estacionariedade para a hipótese alternativa. Dessa maneira, foi possível confirmar a série em primeira diferença como a mais adequada para a análise e desenvolvimento do modelo.

Tabela 3: Teste ADF da série com primeira diferença, resultado positivo para estacionariedade, com $p < 0,05$ de probabilidade.

	t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic	-13.26462	0.0000
Test critical values:		
1% level	-3.461630	
5% level	-2.875195	
10% level	-2.574125	

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

Tabela 4: Teste KPSS da série original, indicando estacionariedade da série, resultando em um $p > 0,05$ de probabilidade.

	LM-Stat.
Kwiatkowski-Phillips-Schmidt-Shin test statistic	0.094049
Asymptotic critical values*:	
1% level	0.739000
5% level	0.463000
10% level	0.347000

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

Existe a necessidade de que os dois testes coincidam na indicação de estacionariedade da série analisada, pois como se pode verificar quando a série original foi testada, indicaram respostas opostas, não auxiliando para uma conclusão sólida dos dados e sua utilização para a modelagem. Para tal, após a apresentação de estacionariedade da série, a etapa seguinte constitui-se do estudo de possíveis modelos e observação numérica de seus coeficientes, para assim considerar a identificação dos modelos capazes de concorrer como os mais adequados ao biodiesel, e encontrar o preditivo ideal. Os modelos concorrentes utilizados, para representar a produção de biodiesel em m^3 (Quadro 1).

Quadro 1: Critérios de comparação para verificar o melhor modelo.

Modelo	p-value dos coeficientes	AIC	BIC
SARIMA(1,1,1)(1,1,0) ₁₂	$\phi_1: +0,986788$ $\Phi_1: +0,354038$ $\theta_1 = -0,201743$	23,37614	23,45531
SARIMA(1,1,2)(1,1,0)₁₂	$\phi_1: +0,995700$ $\Phi_1: +0,358726$ $\theta_1 = -0,202668$ $\theta_2 = -0,284249$	23,32773	23,42232
ARFIMA(1,d,0)	$d = -0,152488$ $\phi_1 = 0,998191$	23,46676	23,53009
ARFIMA(1,d,1)	$d = -0,178356$ $\phi_1 = +0,988376$ $\theta_1 = +0,059487$	23,47496	23,55412

O modelo selecionado foi um SARIMA (1,1,2) (1,1,0)₁₂, destacado no Quadro 1, escolhido por apresentar melhores estatísticas de avaliação. Na análise da série realizaram-se os testes para identificação de modelos ARFIMA, modelos fracionados, os quais foram escolhidos dois modelos concorrentes visíveis, mas devido aos critérios de escolha de modelos, se definiu pelo autorregressivo sazonal. Pode-se visualizar, por meio da figura 3, o comportamento do modelo escolhido através da fac e facp da análise de resíduos (Figura 4).

Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob	
		1	-0.006	-0.006	0.0086	
		2	0.027	0.027	0.1656	
		3	-0.030	-0.030	0.3642	
		4	-0.136	-0.138	4.4252	
		5	0.003	0.002	4.4269	0.035
		6	-0.094	-0.089	6.3888	0.041
		7	-0.024	-0.035	6.5169	0.089
		8	0.017	0.002	6.5785	0.160
		9	0.011	0.007	6.6056	0.252
		10	0.051	0.024	7.1802	0.305
		11	-0.032	-0.040	7.4115	0.387
		12	-0.105	-0.116	9.9204	0.271
		13	0.087	0.089	11.650	0.234
		14	0.001	0.017	11.650	0.309
		15	-0.009	-0.030	11.670	0.389

Figura 4: FAC e FACP dos resíduos.

Com o intuito de realizar previsões para a produção brasileira de biodiesel no período de 11/2022 a 10/2023, fez-se a previsão out-sample, para 12 meses, como temos na figura 5.

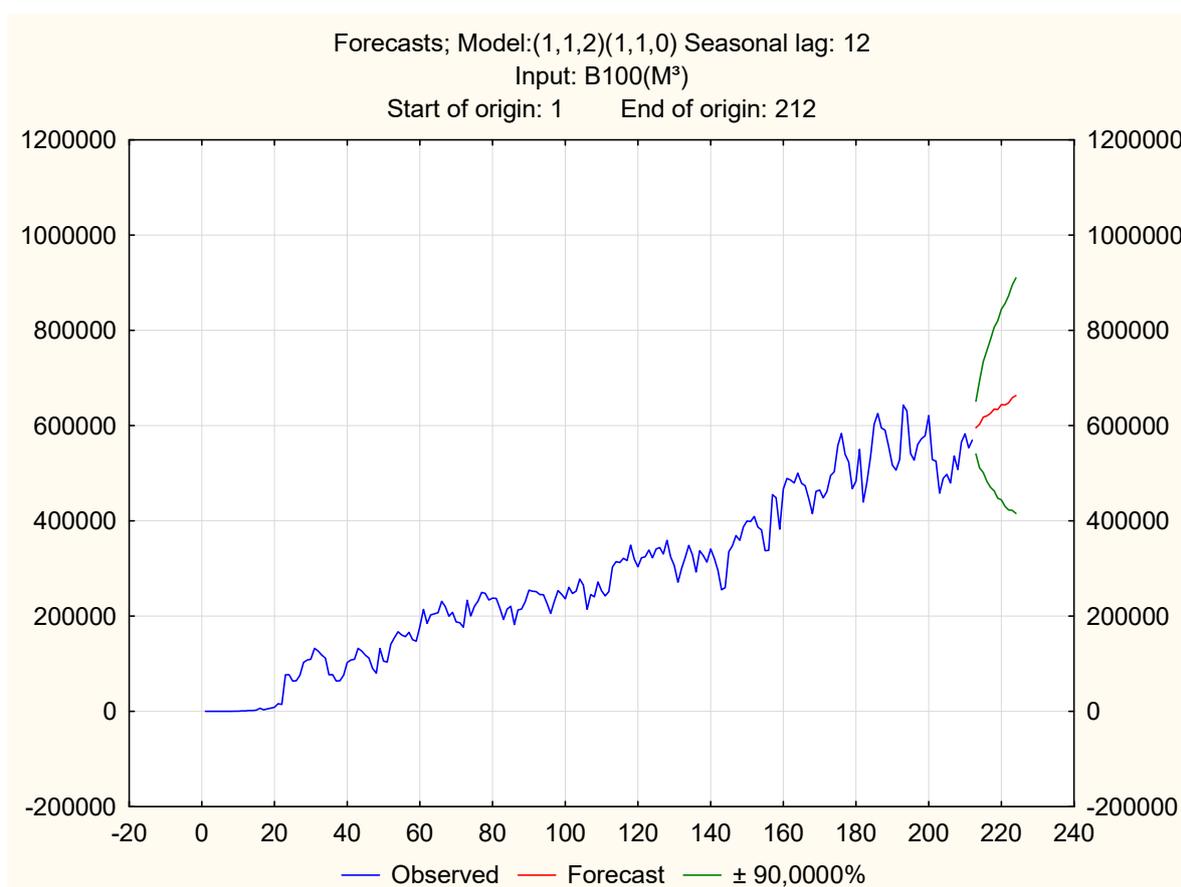


Figura 5: Previsão da produção de Biodiesel para os próximos 12 meses, no período de 11/2022 a 10/2023.

A figura 5, como uma análise final desse trabalho, apresenta-se a previsão para o período de novembro de 2022 a outubro de 2023, notando-se que a produção de biodiesel nesse período tende a um aumento, em torno de 650.000 m³ por mês. Deve-se considerar uma leve queda de produção durante todo o ano de 2021 e 2022, em decorrência da pandemia de Covid-19, em torno de 100.000 m³ em cada mês a

menos que no ano anterior. Esses resultados, se utilizando de um modelo sazonal, essa pesquisa coincide com um estudo realizado por Gonçalves Júnior et al. (2021), em que estudaram a previsão do Imposto sobre Circulação de Mercadorias e Serviços (ICMS) nos estados do sudeste e em um deles, no Espírito Santo, pôde ser modelado por um sarima, sendo que os outros o modelo era Holt-winters. Apesar de os autores analisarem a variável de ICMS, salienta-se a importância de identificação da variabilidade dos dados em relação a um ou mais meses de um ano, o que favorece na modelagem mais adequada a cada série.

CONCLUSÃO

A Alemanha, os Estados Unidos e o Brasil são os maiores mercados mundiais de Biodiesel. O Brasil conta com uma indústria de biodiesel consolidada, possui 70 usinas aptas a produzir e comercializar biodiesel, com uma capacidade instalada superior a 10 bilhões de metros cúbicos.

O Biodiesel sendo um combustível renovável permite uma economia de divisas com a importação de petróleo e óleo diesel, reduzindo a poluição ambiental, gerando alternativas de empregos, em áreas geográficas menos atraentes para outras atividades econômicas e, assim, promovendo a inclusão social. A disponibilização de energia elétrica para comunidades isoladas, na atualidade com elevado custo, em função dos preços do diesel, também deve ser considerada como forma de inclusão, como a inclusão digital, o acesso a bens, serviços, informação, à cidadania e assim por diante.

Considera-se ainda, uma vantagem estratégica, a qual a maioria dos países importadores de petróleo inserem como suas prioridades, a redução da dependência das importações de petróleo. Enfatiza-se também que a introdução do biodiesel aumenta a participação de fontes limpas e renováveis na matriz energética, somando-se principalmente à hidroeletricidade e ao álcool e colocando o Brasil numa posição privilegiada no cenário internacional. Considerando-se um médio prazo, o biodiesel pode se tornar uma importante base de recursos financeiros para o país, somando-se ao álcool como fonte de energia renovável ofertada pelo Brasil à comunidade mundial.

Estando o Brasil entre os maiores produtores e consumidores de biodiesel do mundo, realizar previsão de produção desse biocombustível torna-se uma atividade importante, pois pode revelar as tendências em um momento crítico de mudanças. 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 sua utilização em larga.

A modelagem de séries temporais consiste em um dos métodos utilizados para a realização de previsão e comportamento de dados, mas, além dessa técnica, a compreensão inicial para modelos de predição, vem por meio de regressões lineares, múltiplas, logísticas, etc., evoluindo gradativamente para a análise desse estudo e seguindo a complexidade, realizando análises cross section (estudos transversais) ou até dados em painel (união de séries temporais com cross section), também denominado estudo longitudinal ou de corte.

A aplicação da metodologia Box & Jenkins, apoiada em testes estatísticos subsidiando a modelagem, tornou possível definir modelos apropriados para previsão da produção. Desse modo, a principal

contribuição deste artigo pode ser identificada na perspectiva de produção desse biocombustível, sendo possível uma análise de mercado em relação quantidade produzida no período de um ano. O uso sequencial de gráficos e testes estatísticos substanciando a construção, e a escolha dos modelos mais adequados, bem como a modelagem orientada à previsão da produção de biodiesel no Brasil, atenderam ao objetivo do estudo, sendo necessário a utilização do modelo que melhor se adequou a série de produção, considerando uma diferença entre os termos dessa série, e a sazonalidade no lag 12, ou seja, geralmente nos meses de dezembro, escolhendo-se assim um modelo SARIMA(1,1,2)(1,1,0)₁₂ para previsão da produção nacional do biodiesel. Torna-se importante salientar a possibilidade de continuação das análises da série de produção de biodiesel com modelos mais avançados de estudo, pois foi possível identificar uma variabilidade intrínseca nos dados com o avanço da série no tempo. Recomenda-se estudos com modelos que identifiquem volatilidade nas observações da série, com ou sem transformação, como ARCH-GARCH, SARIMAX ou, ainda, Holt-Winters.

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. In: CONGRESSO BRASILEIRO DE AUTOMÁTICA, 19. **Anais**. 2012.
- ANP. **Agência Nacional do Petróleo, Gás Natural e Biocombustíveis**. 2014.
- ANP. **Agência Nacional do Petróleo, Gás Natural e Biocombustíveis**. 2017.
- ANP. **Agência Nacional do Petróleo, Gás Natural e Biocombustíveis**. 2022.
- BOX, G. E. P.; JENKINS, G. M.. **Time Series Analysis: Forecasting and control**. San Francisco: Holden-Day, 1970.
- BOX, G. E. P.; JENKINS, G. M.. **Time Series Analysis: Forecasting and Control**, Cambridge University Press, Cambridge, 1976.
- BISGAARD, S. K. M.. **Time Series Analysis and Forecasting by Example**, Hoboken, 2011.
- 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**, v.27, n.7, 2011. DOI: <http://dx.doi.org/10.5380/ret.v7i4.25910>
- FAVA, V. L.. **Manual de econometria**. São Paulo: Atlas, 2000.
- GONÇALVES JÚNIOR, G. A.; FEHR, L. C. F. A.; RODRIGUES, R. S.. Previsão do ICMS nos estados da região Sudeste: utilização do modelo sazonal simples e Holt-Winter. **Revista Brasileira de Administração Científica**, v.12, n.4, p.432-440, 2021. DOI: <http://doi.org/10.6008/CBPC2179-684X.2021.004.0030>
- JAFARIAN-NAMIN, S.; GOLI, A.; QOLIPOUR, M.; MOSTAFAEIPOUR, A.; GOLMOHAMMADI, A. M.. Forecasting the wind power generation using Box & Jenkins and hybrid artificial intelligence: A case study. **International Journal of Energy Sector Management**, v.13, n.4, 2019. DOI: <https://doi.org/10.1108/IJESM-06-2018-0002>
- KOHLHEPP, G.. Análise da situação da produção de etanol e biodiesel no Brasil. **Revista Estudos Avançados**, v.24, n.68, p.223-253, 2011. DOI: <https://doi.org/10.1590/S0103-40142010000100017>
- LEE, D. H. Algal biodiesel economy and competition among bio-fuels. **Bioresource Technology**, v.102, n.1, p.43-49, 2011. DOI: <https://doi.org/10.1016/j.biortech.2010.06.034>
- 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**, v.7, n.2, p.166-181, 2011. DOI: <https://doi.org/10.54399/rbgdr.v7i2.427>
- MORETTIN, P. A., TOLOI, C. M. C.. **Análise de Séries Temporais**. 2 ed. São Paulo: Blusher, 2006.
- PANKRATZ, A.. **Forecasting with Univariate Box - Jenkins Models: Concepts and Cases**, Wiley, 1983.
- 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, 1986.
- SOUZA, F. M.; ALMEIDA, S. G.; SOUZA, A. M.. Previsão do preço da gasolina para a região sul do Brasil. **Iberoamerican Journal of Industrial Engineering**, v.3, n.5, 2011. DOI: <https://doi.org/10.13084/2175-8018.v03n05a15>
- 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. DOI: <https://doi.org/10.11221/jima.63.112>

SOUZA, F. M.; ALMEIDA, S. G.; FELICIANE, A. V., SOUZA, A. M. Uso de modelo de memória longa: previsão de equipamentos para a agroindústria. **Iberoamerican Journal of Industrial Engineering**, v.2, n.3, 2011. DOI: <http://doi.org/10.13084/2175-8018.v02n03a08>

TAPANES, N. C; ARANDA, D. A. G.; PEREZ, R. S.; CRUZ, Y. R.. Biodiesel No Brasil: Matérias Primas E Tecnologias De

Produção. **Revista Acta Scientiae & Technicae**, v.1, n.1, p.119-125, 2013. DOI: <https://doi.org/10.17648/uezo-ast-v1i1.11>

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**, v.10, n.1, p.47-67, 2003. DOI: <https://doi.org/10.1590/S0104-530X2003000100005>

Os **autores** detêm os direitos autorais de sua obra publicada. A **CBPC – Companhia Brasileira de Produção Científica (CNPJ: 11.221.422/0001-03)** detêm os direitos materiais dos trabalhos publicados (obras, artigos etc.). Os direitos referem-se à publicação do trabalho em qualquer parte do mundo, incluindo os direitos às renovações, expansões e disseminações da contribuição, bem como outros direitos subsidiários. Todos os trabalhos publicados eletronicamente poderão posteriormente ser publicados em coletâneas impressas ou digitais sob coordenação da Companhia Brasileira de Produção Científica e seus parceiros autorizados. Os (as) autores (as) preservam os direitos autorais, mas não têm permissão para a publicação da contribuição em outro meio, impresso ou digital, em português ou em tradução.