

Área: Sustentabilidade | **Tema:** Produção, Cadeia de Suprimento e Logística Sustentável

**AVALIAÇÃO DA PRODUÇÃO DO BIODIESEL NACIONAL POR MEIO DE UM MODELO DE
PREVISÃO COM VARIÁVEIS EXÓGENAS**

**NATIONAL BIODIESEL PRODUCTION EVALUATION THROUGH A FORECASTING MODEL WITH
EXOGEN VARIABLES**

Adriano Mendonça Souza, Tailon Martins, Bianca Reichert, Cristiano Ziegler e Alisson Castro Barreto

RESUMO

O biodiesel é um combustível renovável produzido por vegetais e gordura animal. No Brasil é obrigatório a adição de 10% do biocombustível ao diesel mineral. A tendência mundial é reduzir os gases emitidos por combustíveis fósseis, nesse sentido o biodiesel vem ganhando espaço no Brasil, com grande viés de expansão para a produção. O objetivo desta pesquisa é encontrar o melhor modelo de previsão para a série de produção do biodiesel no Brasil, por meio da metodologia Box e Jenkins. Utilizou-se a metodologia de Box e Jenkins para analisar a produção mensal de Biodiesel. Além disso, foram utilizadas variáveis exógenas, selecionadas pelo teste de Causalidade de Granger, e testadas em conjunto com o modelo ARIMA, transformando o modelo de previsão em um ARMAX. Os resultados indicam que o melhor modelo para realizar a previsão é o SARIMAX(1,1,1)(1,0,0)12Δ(PIB)1-2(ED)1-3. As duas principais variáveis exógenas estão relacionadas diretamente com a situação econômica do país e as políticas de apoio a biocombustíveis. A pesquisa atinge o objetivo proposto e comprova que modelos estatísticos de previsão podem colaborar em pesquisas relacionadas a biocombustíveis, auxiliando na tomada de decisão em níveis de produção e comercialização.

Palavras-Chave: Previsão, Biodiesel, ARIMA, ARMAX

ABSTRACT

Biodiesel is a renewable product produced by plants and animals oils. In Brazil it is mandatory to add 10% of biofuel to mineral diesel. The global trend is to reduce the gases emitted by fossil fuels, the biodiesel has been gaining ground in Brazil, with great expansion for production. The objective of this research is to find the best forecasting model for biodiesel production in Brazil, using the Box and Jenkins methodology. A Box and Jenkins methodology was used to analyze a monthly production of Biodiesel. In addition, exogenous variables, selected by the Granger causality test, it were identified and tested with the ARIMA model, transforming the prediction model into an ARMAX. The best forecasting model is the SARIMAX(1,1,1)(1,0,0)12Δ(PIB)1-2(ED)1-3. The two main exogenous variables are directly related to the country's economic situation and biofuels support policies. The research achieves the proposed objective and proves that forecasting models can collaborate in biofuel-related research, helping in decision making at production and commercialization levels.

Keywords: Forecast, Biodiesel, ARIMA, ARMAX

AVALIAÇÃO DA PRODUÇÃO DO BIODIESEL NACIONAL POR MEIO DE UM MODELO DE PREVISÃO COM VARIÁVEIS EXÓGENAS

1 INTRODUÇÃO

A demanda global crescente por energia e a consequente poluição ambiental causada pela sua produção, nos últimos anos, aumentaram a preocupação das autoridades e da população na busca pelo desenvolvimento sustentável. Desde a Revolução Industrial, a matriz energética mundial está apoiada sob fontes de energias não renováveis, que são as principais fontes responsáveis pela emissão dos gases de efeito estufa (GEE) (BRASIL, 2019; IEA, 2019).

Em contraste à matriz energética mundial, mais de 40% das fontes de energia brasileira são renováveis, provenientes principalmente da lenha, carvão vegetal, hidráulica e derivados de cana (BRASIL, 2018). Pode-se citar também a utilização de biocombustíveis, como o biodiesel, que por meio do Programa Nacional de Produção e Uso do Biodiesel (PNPB) e da Lei nº 11.097/2005 passou a integrar a matriz energética brasileira (SILVA, 2019).

O biodiesel é um combustível biodegradável proveniente de fontes renováveis como gorduras animais, soja, palma, girassol, babaçu, amendoim, mamona e pinhão-mansão. No Brasil, a soja ainda é a principal matéria-prima do biocombustível (BRASIL, 2018). Dentre as vantagens de sua utilização, pode-se citar vantagens ecológicas como a não emissão de óxidos de enxofre, na combustão de motores; vantagens macroeconômicas como o aumento da demanda por produtos agrícolas; a diversificação da matriz energética; vantagens financeiras através do provável atingimento de metas estabelecidas pelo protocolo de Kyoto e a habilitação no mercado de bônus de carbono; e a diminuição na importação de petróleo e seus derivados (RATHMANN, 2005).

O setor do biodiesel no Brasil é alavancado pela obrigatoriedade da adição do biocombustível ao diesel. Em 2016, o percentual de obrigatoriedade era de 7%, passando para 8% em 2017, 9% em 2018 e 10% em 2019. Caso sejam realizados testes o percentual poderá chegar até 15% e existe ainda a possibilidade de adição voluntária superior ao percentual estabelecido por lei (BRASIL, 2016).

O biodiesel produzido no Brasil tem como destino o mercado nacional e possui regulação e comercialização através da Agência Nacional do Petróleo, Gás Natural e Biocombustíveis (ANP) que realiza os leilões. O Brasil possui uma capacidade instalada de 7,7 bilhões de litros/ano, mostrando haver ainda espaço para crescimento na sua utilização (CARRARO, CÉSAR e CONEJERO, 2019).

A fim de contribuir com as políticas de incentivo à produção nacional de biocombustíveis para diversificar a matriz energética brasileira, tem-se como objetivo encontrar o melhor modelo de previsão para a série de produção do biodiesel no Brasil, por meio da metodologia Box e Jenkins.

2 METODOLOGIA

O foco do estudo foi a produção mensal de Biodiesel - B100 (m³) no Brasil. Os valores analisados foram coletados no site da Agência Nacional do Petróleo, Gás Natural e Biocombustíveis – ANP (2019), no período entre janeiro de 2009 e maio de 2019, contabilizando 125 observações mensais.

A metodologia de Box e Jenkins foi aplicada para analisar a produção do Biodiesel, pois trata a autocorrelação serial dos dados (BOX; JENKINS, 1970). Um dos pressupostos dos modelos Autoregressivos Integrados e de Médias Móveis (ARIMA) é de que as séries devem ser estacionárias. Essas características podem ser observadas por meio do gráfico da série

original, dos gráficos das funções de autocorrelação (FAC) e de autocorrelação parcial (FACP), ou por meio dos testes de raízes unitárias (BOX; JENKINS; REINSEL, 1994).

O gráfico de sequência da série original foi elaborado para identificar comportamento de tendência ou sazonalidade. Os gráficos da FAC e da FACP foram utilizados para verificar a autocorrelação das observações e indicar os possíveis filtros do modelo a ser utilizado na etapa de estimação (AR, MA, ARIMA, SARIMA ou ARFIMA) (MORETTIN, 2016). Para confirmar a estacionariedade da série, foram utilizados os testes de raízes unitárias Dickey-Fuller Aumentado (ADF) e Kwiatkowski–Phillips–Schmidt–Shin (KPSS) (DICKEY; FULLER, 1981; KWIATKOWSKI et al., 1992), os quais indicaram o número de diferenças necessárias para tornar a série estável (BUENO, 2015).

O modelo geral ARIMA (p, d, q) fundamenta-se na força de ligação entre as observações, a qual pode ser utilizada para prever os seus valores futuros (SOUZA, 2016). A equação genérica do modelo ARIMA é apresentada na Equação 1.

$$\phi(B)\Delta^d X_t = \theta(B)\alpha_t \quad (1)$$

Onde B representa o operador de retardo, d é o número de diferenças; ϕ caracteriza o parâmetro autorregressivo, θ é o parâmetro da parte de médias móveis, e α_t representa o resíduo do modelo com valores independentes e identicamente distribuídos (*i.i.d.*).

Uma das alternativas do modelo ARIMA consiste em integrar variáveis exógenas para explicar as oscilações da variável em estudo: o modelo autorregressivo de médias móveis com variável exógena (ARMAX) (GUSTIN; McLEOD; LOMAS, 2018). No modelo ARMAX (p, q, y), adiciona-se um parâmetro na equação geral para representar o efeito da variável exógena (Y_t), conforme a equação 2:

$$\phi(B)\Delta^d X_t = \theta(B)\alpha_t + \phi(B)\Delta^d Y_t \quad (2)$$

Onde X_t representa a série de dados em análise, B é o operador de retardo, d é o número de diferenças da séries, ϕ representa o parâmetro da parte autorregressiva, θ é o parâmetro da parte de médias móveis, α_t caracteriza o resíduo do modelo e Y_t representa a série de dados da variável exógena.

As variáveis exógenas podem ser definidas de acordo com as suas relações de precedência com a variável principal interesse. A variável exógena que preceder a variável principal poderá ser utilizada para realizar a previsão (BUENO, 2015).

Para definir o grau de causalidade realizou-se o teste de Causalidade de Granger que consiste em analisar se as defasagens de uma variável possuem efeito sobre outra variável, por meio de uma equação de regressão (ROCHA; BRILHANTE, 2015). Se ocorre essa precedência, e ela é significativa, então a variável principal terá uma previsão mais efetiva se a variável exógena for adicionada ao seu modelo de regressão (MORETTIN, 2016). Para determinar se Y_t causa X_t no sentido de Granger, utiliza-se a equação 3:

$$X_t = \phi_0 + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \phi_3 Y_{t-3} + \dots + \alpha_t \quad (3)$$

Onde X_t é a variável dependente, Y_t representa a variável explicativa e as suas defasagens e ε_t caracteriza o erro (GRANGER, 2004).

Foram consideradas as variáveis exógenas: PIB Brasileiro (R\$); Consumo de Diesel (barris); Preço de Revenda do Diesel (R\$/l); Importação do Diesel (m³); Exportação do Diesel (m³); Cotação da Soja (cents/bushel) (ANP, 2019; IPEADATA, 2019; INVESTING, 2019).

Após o ajuste dos modelos ARIMA e ARMAX, foi realizada a análise dos resíduos por meio da FAC e FACP, pois o modelo só será apto a realizar previsões se os seus resíduos forem não autocorrelacionados, normalmente distribuídos, com variância constante e média igual a zero. Essas propriedades caracterizam o resíduo como Ruído Branco $\alpha \approx N(0, \sigma^2)$ (BUENO, 2015).

O melhor ajuste será o modelo que apresentar Ruído Branco e os mínimos valores para os critérios de informação Akaike (AIC) e Bayesiano (BIC) (AKAIKE, 1974). Esses critérios são calculados conforme as Equações 4 e 5.

$$AIC = T \ln(SQR) + 2n \quad (4)$$

$$BIC = T \ln(SQR) + n \ln(T) \quad (5)$$

Onde T representa o número de observações, SQR é a soma dos quadrados dos resíduos e p caracteriza o número de parâmetros que foram estimados no modelo.

Com o objetivo de selecionar o melhor modelo, também se estimam as estatísticas de previsão com base nos erros de ajuste do modelo. As estatísticas do Erro Percentual Médio Absoluto (MAPE) e U-Theil serão aplicadas conforme as Equações 6 e 7 (SOUZA et al., 2011).

$$MAPE = \frac{1}{n} \sum \left| \frac{x_t - \hat{x}_t}{x_t} \right| * 100 \quad (6)$$

$$U - Theil = \frac{\sqrt{\sum_{t=1}^N (x_t - \hat{x}_t)^2}}{\sqrt{\sum_{t=1}^N (x_t - x_{t-1})^2}} \quad (7)$$

Onde x_t é a observação da série original no instante t e \hat{x}_t representa a previsão do modelo selecionado no instante t .

Essas estatísticas auxiliam na seleção do melhor modelo, pois o modelo que apresentar os menores valores terá o melhor desempenho ao realizar a previsão da variável em análise.

Após escolha do melhor modelo, elabora-se o gráfico dos valores reais, previstos e dos resíduos com o intuito de verificar se o modelo escolhido tem a capacidade de representar os movimentos e oscilações da série original e realizar a previsão.

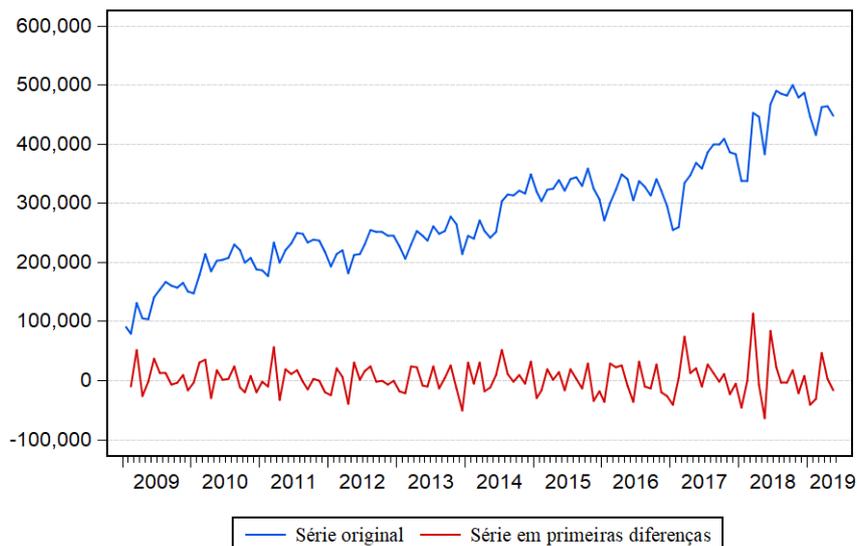
3 RESULTADOS

A análise descritiva da variável principal utilizada nesta pesquisa, a qual se refere a produção nacional de biodiesel, foi realizada por meio das medidas descritivas de tendência central (média, mediana, máximo e mínimo) e de dispersão (desvio-padrão e coeficiente de variação), no período de janeiro de 2009 a maio de 2019. Deste modo, verificou-se que a média mensal da produção de biodiesel no país foi de 282.929,5m³, enquanto a mediana foi de 255.360,7m³.

O maior pico de produção ocorreu em outubro de 2018 em que o país registrou um total de 500.209 m³ de biodiesel produzidos. O menor registro de produção se encontra em fevereiro de 2009, com apenas 80.224,3 m³ produzidos. Os resultados encontrados para as medidas de dispersão são de 95.848,89m³ de desvio-padrão e um coeficiente de variação de 33,87%.

Na análise gráfica da Figura 1 é possível observar a série da produção nacional de biodiesel.

Figura 1 – Produção nacional de biodiesel (m³) 2009-2019



Fonte: os autores.

É possível observar na Figura 1 que a série em análise é não estacionária em nível, apresentando comportamento de tendência crescente, o que evidencia a necessidade da utilização do método de diferenciação para estabilizá-la, a fim de aplicar a modelagem ARIMA. Identificou-se também que a série elevou os níveis de produção no decorrer dos anos, apresentando movimentos repetidos, que denotam a queda da produção para os primeiros e últimos meses do ano, bem como os picos de produção ocasionados na metade do ano. Esses fenômenos podem ser indícios da existência de sazonalidade na série, o qual pode ser captado pelos modelos sazonais (SARIMA) da classe geral ARIMA.

De acordo com Martins e Carneiro (2017) o fenômeno sazonal que pôde ser observado na série da produção do biodiesel é ocasionado pela influência do cultivo das matérias-primas utilizada, sendo a soja a principal. A safra do grão ocorre em maio, mês que antecede os picos de produção do biodiesel. A matéria-prima é fortemente influenciada pelas condições territoriais e climáticas. Essas características têm incentivado o investimento em políticas públicas para o aproveitamento das potencialidades regionais (BRASIL, 2018).

A regiões mais produtivas são a região Centro-Oeste e Sul, pois possuem quase 70% das empresas e 80% do total da capacidade produtiva, enquanto as regiões menos produtivas são as regiões Nordeste e Norte, apresentando o menor número de indústrias e menos de 5% do total de capacidade instalada de produção (MARTINS e CARNEIRO, 2017).

A produção do biodiesel no país surgiu por meio do Programa Nacional de Produção de Biodiesel (PNPB), em meio a um cenário do aumento do preço do petróleo e da pressão mundial, no quesito de reduzir o consumo dos combustíveis fósseis. No Brasil, a Lei 11.097/2005, inseriu o biodiesel como mistura parcial para o diesel mineral, visando promover o desenvolvimento sustentável. Inicialmente o porcentual mínimo estabelecido foi de 2% de adição de biodiesel ao diesel fóssil, e desde então vem sofrendo constantes alterações recorrentes da política nacional dos biodieseis, estabelecida pelo Conselho Nacional de Política Energética (CNPE), que prevê a expansão obrigatório da adição do biodiesel ao diesel fóssil no Brasil de forma gradual e progressiva, a fim de consolidar seu uso.

Os picos produtivos ocasionadores da instabilidade da série se fundamentam em razão do aumento da porcentagem de biodiesel adicionado ao diesel fóssil estabelecido em lei. Nos últimos 10 anos, estas mudanças ocorreram em julho de 2009, em que o percentual de 3% foi alterado para 4%. Em janeiro de 2010 o percentual subiu para 5% e somente em julho de 2014 passou para 6%, e em novembro do mesmo ano atingiu os 7%. O percentual que em março de 2017 passou a 8%, no ano seguinte, no mesmo mês aumentou para 10% (VIEGAS E ARANTES, 2018).

Para verificar a estacionariedade da série foram utilizados os testes de raízes unitárias ADF e KPSS, considerando um nível de significância de 5% ($\alpha=0,05$). A hipótese nula (H_0) do teste ADF é de que a série apresenta raiz unitária, sendo assim é não estacionária. O teste realizado apresentou o *t-statistic* de -2,88, em que rejeitou a hipótese nula após aplicação de primeiras diferenças (*t-statistic* de -11,31), confirmando que a série é estacionária após o método de diferenciação de ordem 1. Para corroborar com o teste ADF, foi realizado o teste KPSS, que apresenta hipótese nula contrária (H_0) de que a série é estacionária. O teste KPSS apresentou *t-statistic* de 0,46 para a série em nível, resultando na rejeição da hipótese nula. Após aplicar primeiras diferenças a hipótese nula foi aceita com *t-statistic* de 0,05, concordando em que a série é estacionária após primeiras diferenças.

Após verificar a estacionariedade da série foram traçados os gráficos da Função de Autocorrelação (FAC) e Função de Autocorrelação Parcial (FACP) para identificar o possível modelo a ser estimado, o qual pode ser visualizado na Figura 2.

Figura 2 – Gráfico da FAC e FACP da produção nacional de biodiesel (m³)

| Autocorrelation | Partial Correlation | AC | PAC | Q-Stat | Prob |
|-----------------|---------------------|--------|--------|--------|------|
| 1 | 0.933 | 0.933 | 111.44 | 0.000 | |
| 2 | 0.873 | 0.018 | 209.76 | 0.000 | |
| 3 | 0.841 | 0.189 | 301.78 | 0.000 | |
| 4 | 0.805 | -0.028 | 386.82 | 0.000 | |
| 5 | 0.761 | -0.039 | 463.50 | 0.000 | |
| 6 | 0.718 | -0.031 | 532.32 | 0.000 | |
| 7 | 0.684 | 0.033 | 595.34 | 0.000 | |
| 8 | 0.653 | 0.004 | 653.15 | 0.000 | |
| 9 | 0.626 | 0.046 | 706.71 | 0.000 | |
| 10 | 0.595 | -0.031 | 755.62 | 0.000 | |
| 11 | 0.574 | 0.064 | 801.43 | 0.000 | |
| 12 | 0.554 | -0.002 | 844.58 | 0.000 | |

Fonte: os autores.

De acordo com a Figura 2, a FACP apresentou apenas um *lag* significativo, indicando que o modelo possa ser um AR(1). O decaimento lento da FAC evidencia que a série original é não estacionária, bem como sugere que o modelo possa ser fracionário.

O passo seguinte consistiu em estimar os possíveis modelos e dentre eles encontrar o melhor modelo utilizado para realizar as previsões.

Na Tabela 1 estão dispostos os parâmetros dos modelos ARIMA e seus respectivos *p-valores*, bem como os critérios penalizadores AIC BIC, e a condição de Ruído Branco-RB.

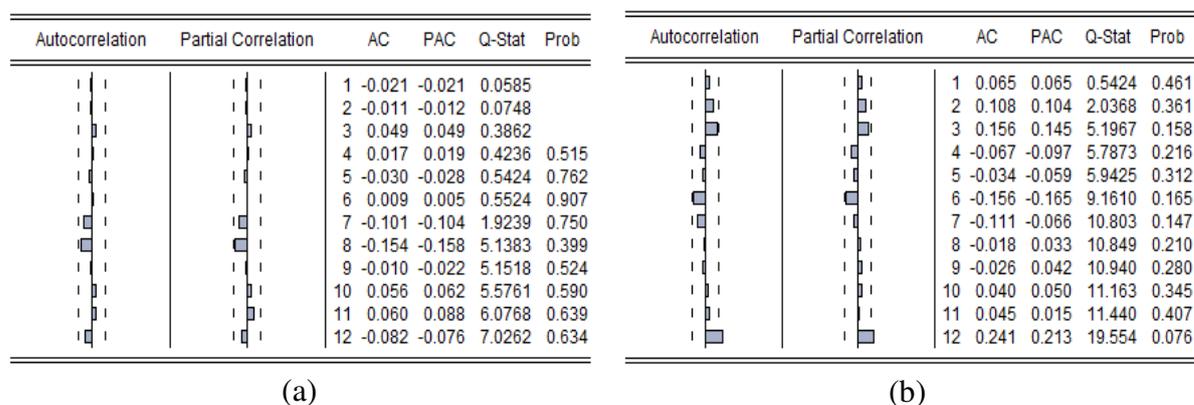
Tabela 1 – Modelos concorrentes ARIMA

| Modelo | Coefficientes | AIC | BIC |
|--------------------------------------|--------------------|-------|-------|
| ARIMA (1,1,1) | $\phi_{12}=0,466$ | 22,98 | 23,06 |
| | $\theta_2=-0,2635$ | | |
| SARIMA (1,1,1) (1,0,0) ₁₂ | $\phi_1=-0,2456$ | 22,94 | 23,03 |
| | $\theta_2=-0,3048$ | | |
| | $\Phi_{12}=0,4964$ | | |

Fonte: os autores.

O melhor modelo encontrado foi o SARIMA (1,1,1)(1,0,0)₁₂, que contém um componente autoregressivo (AR), uma diferença para tornar a série estável, um componente de médias móveis (MA) e um componente sazonal autoregressivo (SAR) de ordem 12. Este modelo apresentou os menores critérios penalizadores (AIC/BIC) entre os concorrentes, e demonstrou ser do tipo ruído branco. A condição de ruído branco foi verificada por meio dos gráficos da Função de Autocorrelação (FAC) e Função de Autocorrelação Parcial (FACP) dos resíduos (a) e dos resíduos ao quadrado (b), que se encontram dentro do limite de 5% de significância, ilustrados na Figura 3.

Figura 3 – FAC e FACP dos resíduos (a) e dos resíduos ao quadrado (b) do modelo SARIMA

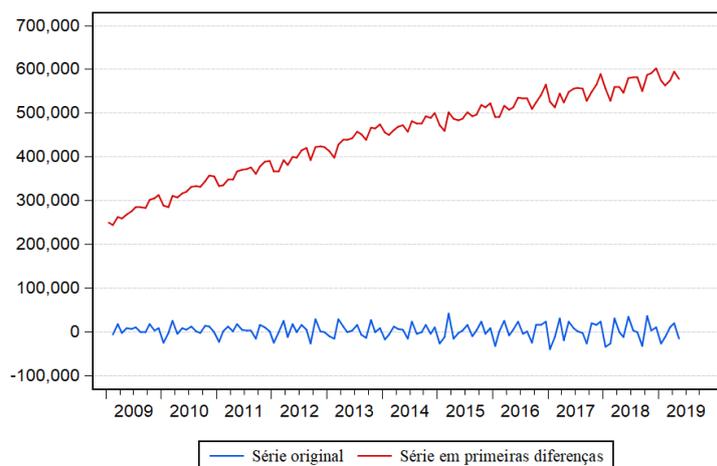


Fonte: os autores.

Buscando aumentar a assertividade das previsões, testou-se o melhor modelo ARIMA encontrado incluindo variáveis exógenas significativas. Este modelo testado é denominado ARMAX. Para isto, foram analisadas as variáveis que de acordo com a literatura podem influenciar na produção do biodiesel.

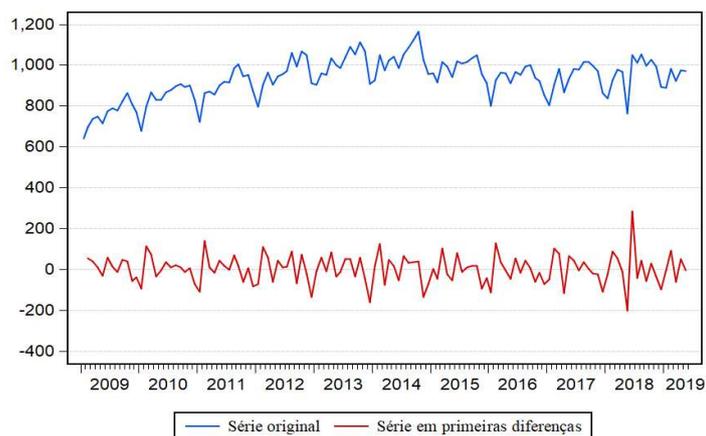
As variáveis mais relevantes foram: PIB Brasileiro (R\$); Consumo de Diesel (barris); Preço de Revenda do Diesel (R\$/l); Importação do Diesel (m³); Exportação do Diesel (m³); Cotação da Soja (cents/bushel). Nas Figura 4 a 9, é possível visualizar as variáveis utilizadas. Para as variáveis que apresentaram comportamento tendencioso e sazonal, utilizou-se a aplicação de primeiras diferenças para torná-las estáveis.

Figura 4 – Produto Interno Bruto – PIB (R\$) 2009-2019



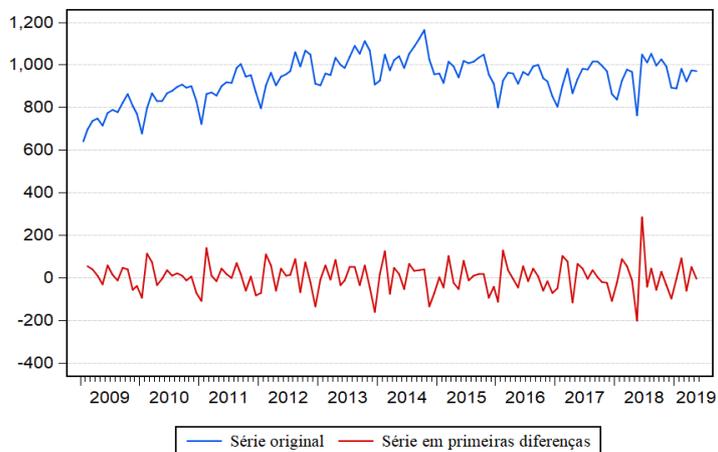
Fonte: os autores.

Figura 5 – Consumo de Diesel (barris) 2009-2019



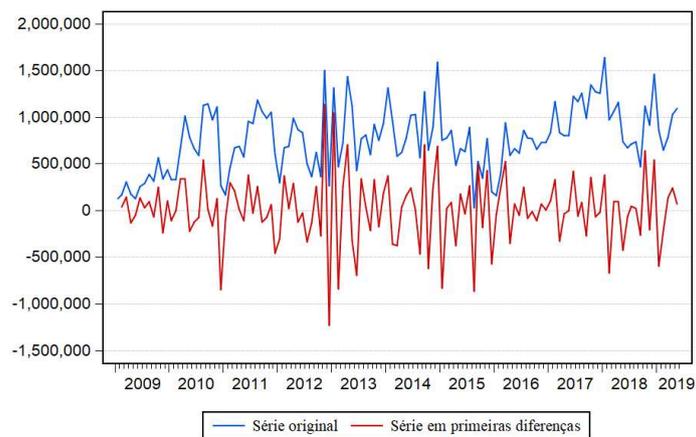
Fonte: os autores.

Figura 6 – Preço de revenda do Diesel (R\$/l) 2009-2019



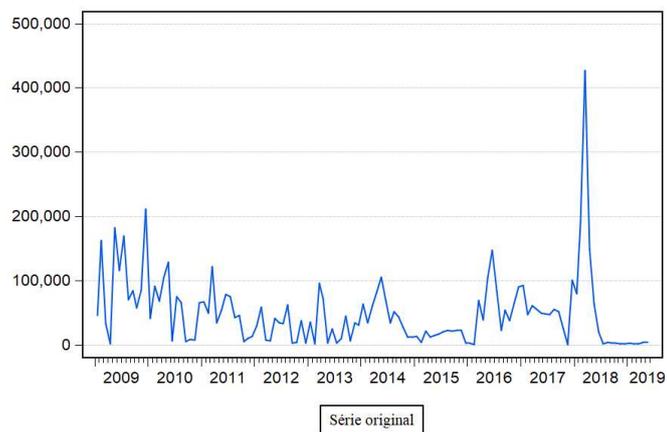
Fonte: os autores.

Figura 7 – Importação de Diesel (m³) 2009-2019



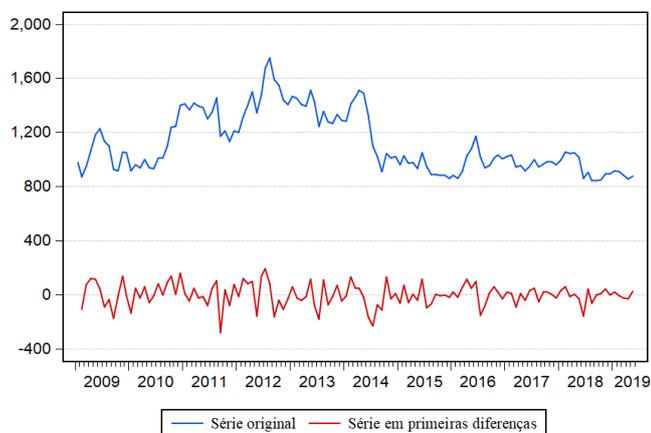
Fonte: os autores.

Figura 8 – Exportação de Diesel (m³) 2009-2019



Fonte: os autores.

Figura 9 – Cotação da Soja (cents/bushel) 2009-2019



Fonte: os autores.

Para confirmar a estacionariedade das variáveis exógenas foram utilizados os testes de raízes unitárias ADF e KPSS, considerando um nível de significância de 5% ($\alpha=0,05$). O teste de estacionariedade pode ser verificado na Tabela 2.

Tabela 2 – Teste de estacionariedade para variáveis exógenas

| Teste de estacionariedade | ADF | | | KPSS | | | Ordem de diferenciação |
|---------------------------|---------------------------------|--------------|---------------|------------------------|-------------|-------------|------------------------|
| | <i>t-Statistic</i> | d = 0 | d = 1 | <i>t-Statistic</i> | d = 0 | d = 1 | |
| Série em análise | | | | | | | |
| Produto Interno Bruto | -3,45 | -1,7 | -12,85 | 0,17 | 0,36 | 0,14 | 1 |
| Exportação de Diesel | -1,94 | -4,96 | -15,52 | 0,46 | 0,30 | 0,20 | 0 |
| Importação de Diesel | -1,94 | -1,17 | -18,92 | 0,46 | 0,57 | 0,23 | 1 |
| Produção de Diesel | -3,44 | -2,97 | -9,08 | 0,14 | 0,26 | 0,14 | 1 |
| Cotação da Soja | -2,88 | -2,01 | -10,97 | 0,46 | 0,52 | 0,13 | 1 |
| Consumo Diesel | -1,94 | 2,34 | -2,59 | 0,46 | 0,71 | 0,36 | 1 |
| Hipótese nula | A série apresenta raiz unitária | | | A série é estacionária | | | |

Fonte: os autores.

Após a aplicação do teste de estacionariedade, foi realizado o teste de Causalidade de Granger para verificar se realmente existe relação entre as variáveis. Verificou-se no sentido de causalidade unidirecional ou bidirecionais, conforme a Tabela 3.

Tabela 3 – Resultado do teste de Causalidade de Granger

| Relações de Causalidade | F-stat | p-valor |
|---|--------|---------|
| $\Delta(\text{Produção de Biodiesel}) \leftrightarrow \Delta(\text{Consumo de Diesel})$ | 2,58 | 0,0797 |
| $\Delta(\text{Consumo de Diesel}) \leftrightarrow \Delta(\text{Produção de Biodiesel})$ | 7,74 | 0,0007 |
| $\Delta(\text{PIB}) \rightarrow \Delta(\text{Produção de Biodiesel})$ | 10,81 | 0,0005 |
| $\Delta(\text{Importação de Diesel}) \rightarrow \Delta(\text{Consumo de Diesel})$ | 4,51 | 0,0129 |
| $\Delta(\text{PIB}) \leftrightarrow \Delta(\text{Consumo de Diesel})$ | 16,01 | 0,0007 |
| $\Delta(\text{Consumo de Diesel}) \leftrightarrow \Delta(\text{PIB})$ | 6,57 | 0,0020 |
| $\Delta(\text{Cotação da Soja na Bolsa}) \rightarrow (\text{Exportação de Diesel})$ | 3,40 | 0,0367 |
| $\Delta(\text{Produção de Biodiesel}) \rightarrow (\text{Exportação de Diesel})$ | 2,93 | 0,0570 |

O símbolo \rightarrow indica causalidade unidirecional; o símbolo \leftrightarrow indica causalidade bidirecional; H_0 : Série “X” não causa série “Y” no sentido de Granger; H_1 : Série “X” causa série “Y” no sentido de Granger. $\alpha=10\%$.

Fonte: os autores.

De acordo com a Tabela, 3 é possível visualizar causalidade bidirecional entre Produção de Biodiesel e Consumo de Diesel, e entre PIB e Consumo de Diesel. Enquanto é evidenciado o comportamento unidirecional entre PIB e Produção de Biodiesel, entre Importação de Diesel e Consumo de Diesel, entre Cotação da Soja na Bolsa e Exportação de Diesel, e entre Produção de Biodiesel e Exportação de Diesel.

As variáveis que implicam diretamente na produção do Biodiesel foram selecionadas para testar o modelo ARMAX, que são o Consumo de diesel, o PIB e a Exportação do Diesel.

Partindo do melhor modelo ARIMA encontrado, SARIMA (1,1,1)(1,0,0)₁₂, foram inseridas as variáveis exógenas de acordo com a Causalidade de Granger: Consumo de diesel, o PIB e a Exportação do Diesel. As séries foram testadas com defasagens de até 12 lags e após diversas combinações de inclusões as variáveis que se mostraram significativas foram: $\Delta(\text{Produto Interno Bruto})(t-1)$, $\Delta(\text{Produto Interno Bruto})(t-2)$, Exportação de Diesel(t-1), Exportação de Diesel(t-2) e Exportação de Diesel(t-3). Deste modo, foi verificado quais variáveis influenciam a produção de Biodiesel em até 12 meses atrás.

Na Tabela 4 apresenta-se o modelo ARIMAX ajustado para realizar as previsões, SARIMAX(1,1,1)(1,0,0)₁₂ $\Delta(\text{PIB})_{1-2}(\text{ED})_{1-3}$. Pode-se observar que todos os parâmetros se encontram entre -1 e 1 com seus respectivos p-valores inferiores a 0,05 e, ainda, demonstrando ser Ruído Branco (RB). Este modelo segue o pressuposto de apresentar os menores critérios penalizadores dentre os modelos concorrentes estimados.

Tabela 4 – Modelo SARIMAX selecionado dentre os modelos concorrentes

| SARIMAX(1,1,1)(1,0,0) ₁₂ $\Delta(\text{PIB})_{1-2}(\text{ED})_{1-3}$ | | | |
|---|---------|-------|-------|
| Parâmetro | p-valor | AIC | BIC |
| $\phi_1=-0,22$ | 0,0313 | 22,79 | 22,97 |
| $\theta_1=-0,28$ | 0,0038 | | |
| $\Phi_{12}=0,39$ | 0,0001 | | |
| PIB(-1)=-0,55 | 0,0021 | | |
| PIB(-2)=-0,34 | 0,0456 | | |
| ED(-1)=-0,7 | 0,0247 | | |
| ED(-2)=-0,10 | 0,0041 | | |
| ED(-3)=-0,12 | 0,0019 | | |

Fonte: os autores.

Para analisar se o modelo é do tipo ruído branco, foram analisados gráficos da Função de Autocorrelação (FAC) e Função de Autocorrelação Parcial (FACP) dos resíduos (a) e dos resíduos ao quadrado (b), ilustrados na Figura 5, o qual mostrou que todos os lags encontram-se dentro do limite de significância de 5%, confirmando que o modelo é do tipo ruído branco.

Figura 10 – FAC e FACP dos resíduos (a) e dos resíduos ao quadrado (b) do modelo SARIMAX

| Autocorrelation | Partial Correlation | AC | PAC | Q-Stat | Prob* | Autocorrelation | Partial Correlation | AC | PAC | Q-Stat | Prob |
|-----------------|---------------------|------------------|--------|--------|-------|-----------------|---------------------|-----------------|--------|--------|------|
| | | 1 0.005 0.005 | 0.0028 | | | | | 1 0.036 0.036 | 0.1639 | 0.686 | |
| | | 2 0.000 0.000 | 0.0028 | | | | | 2 0.051 0.049 | 0.4858 | 0.784 | |
| | | 3 -0.088 -0.088 | 0.9912 | | | | | 3 -0.084 -0.087 | 1.3753 | 0.711 | |
| | | 4 -0.005 -0.005 | 0.9950 | 0.319 | | | | 4 -0.060 -0.057 | 1.8411 | 0.765 | |
| | | 5 -0.066 -0.066 | 1.5577 | 0.459 | | | | 5 -0.012 0.001 | 1.8596 | 0.868 | |
| | | 6 -0.101 -0.110 | 2.8931 | 0.408 | | | | 6 -0.115 -0.117 | 3.5803 | 0.733 | |
| | | 7 -0.207 -0.213 | 8.5311 | 0.074 | | | | 7 -0.090 -0.094 | 4.6511 | 0.702 | |
| | | 8 -0.163 -0.195 | 12.059 | 0.034 | | | | 8 0.019 0.034 | 4.7013 | 0.789 | |
| | | 9 0.057 0.016 | 12.488 | 0.052 | | | | 9 -0.017 -0.031 | 4.7400 | 0.856 | |
| | | 10 0.096 0.048 | 13.743 | 0.056 | | | | 10 0.026 -0.007 | 4.8289 | 0.902 | |
| | | 11 0.056 0.010 | 14.177 | 0.077 | | | | 11 0.084 0.083 | 5.7892 | 0.887 | |
| | | 12 -0.062 -0.106 | 14.710 | 0.099 | | | | 12 0.146 0.132 | 8.7238 | 0.726 | |

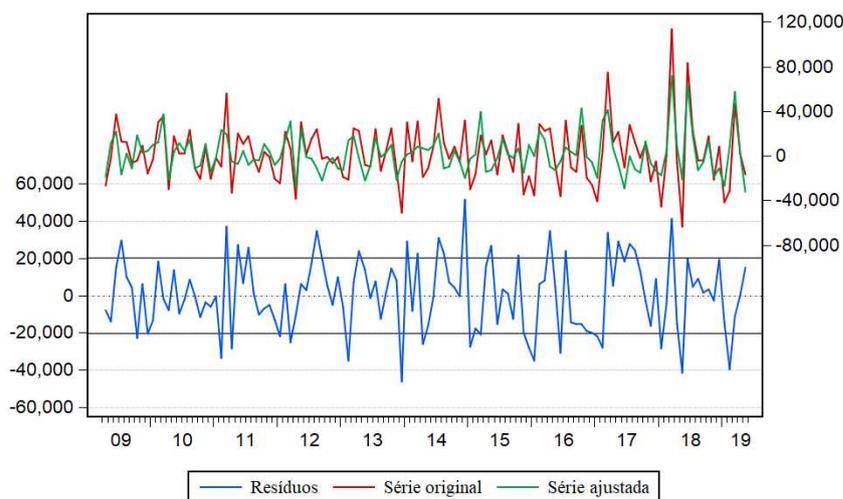
(a)

(b)

Fonte: os autores.

Após analisar os dois melhores modelos encontrados (ARIMA e ARMAX) verificou-se que as estatísticas de previsão para o modelo SARIMA (1,1,1)(1,0,0)₁₂ apresentou U-Theil de 0,035 e MAPE de 5,92%, enquanto o modelo SARIMAX(1,1,1)(1,0,0)₁₂Δ(PIB)₁₋₂(ED)₁₋₃, obteve U-Theil de 0,031 e MAPE de 5,63%, evidenciando que o melhor modelo para realizar as previsões é o modelo ARIMAX. O ajuste do modelo selecionado pode ser visualizado na Figura 11, bem como a série original, valores ajustados e os resíduos.

Figura 11 – Previsão da produção de biodiesel (*in sample*) utilizando o modelo SARIMAX(1,1,1)(1,0,0)₁₂Δ(PIB)₁₋₂(ED)₁₋₃



Fonte: os autores.

O modelo selecionado foi utilizado para realizar a projeção dos valores apresentou estatística de U-Theil menor que 1, indicando que o modelo ajustado é melhor que uma previsão ingênua, o qual é corroborado com o MAPE, que apresentou valores baixos, evidenciando que o modelo seja um bom predictor do comportamento futuro.

É possível afirmar que o modelo ajustado é adequado para realizar previsões futuras, enquanto a produção nacional do biodiesel pode ser explicada pela própria variável em razão dos valores defasados, bem como a existência de variáveis exógenas significativas presentes no modelo, que aumentam ainda mais a assertividade do modelo proposto. O modelo estimado permite realizar previsões acuradas de acordo com o embasamento científico e os critérios propostos na literatura de séries temporais.

As variáveis exógenas PIB e a Exportação do Diesel estão intimamente relacionadas com a variável principal em estudo, a Produção Nacional de Biodiesel. Esse dado foi constatado pelo teste de Causalidade de Granger. Além do teste, as variáveis exógenas impactam a variável resposta de acordo com a situação econômica geral do país e as políticas de apoio a biocombustíveis.

O PIB está relacionado ao consumo de produtos pela população, conseqüentemente, quando aumenta o consumo por produtos industrializados, aumenta o PIB. A maioria dos produtos no Brasil são transportados por caminhões, modal rodoviário, refletindo em um aumento no consumo de biodiesel misturado ao combustível, quando a PIB está em elevação.

Hoje, o diesel consumido nos postos de combustíveis é composto por 10% de biodiesel. Com essa medida, deixam de ser consumidos 10% de diesel mineral, aumentando a oferta desse

produto. Esse excedente pode ser exportado, contribuindo para a balança comercial brasileira. A cada aumento de mistura de biodiesel no diesel mineral, necessita aumentar a produção de biodiesel, o que provoca a diminuição da demanda pelo combustível fóssil, possibilitando maior exportação para suprir a demanda do mercado externo.

4 CONCLUSÕES

A metodologia ARMAX se mostrou eficiente no desenvolvimento de um modelo de previsão da produção do Biodiesel, incorporando as variáveis exógenas à modelagem ARIMA, considerando assim que o objetivo desse estudo foi alcançado. Inicialmente foi desenvolvida uma modelagem ARIMA para a variável principal, Produção do Biodiesel, em que o modelo SARIMA (1,1,1)(1,0,0)₁₂ foi o mais adequado, em que seguidamente se inseriram as variáveis exógenas significativas.

As variáveis exógenas utilizadas nesse estudo foram: PIB Brasileiro (R\$); Consumo de Diesel (barris); Preço de Revenda do Diesel (R\$/l); Importação do Diesel (m³); Exportação do Diesel (m³); Cotação da Soja (cents/bushel). As que mais influenciam a variável principal em estudo foram o PIB e a Exportação do Diesel, isso ficou comprovado pelo teste de Causalidade de Granger. O modelo mais adequado a essa série é o SARIMAX(1,1,1)(1,0,0)₁₂Δ(PIB)₁₋₂(ED)₁₋₃. Esse modelo é apropriado para realizar previsões futuras, por apresentar os valores de U-Theil e MAPE inferiores ao modelo concorrente.

O aumento da produção de biodiesel está diretamente relacionado ao PIB do Brasil, com o aumento do consumo interno, aumenta o consumo de combustível, principalmente para os modais de transporte movidos a diesel. Além disso, devido a potencialidade produtiva do país, se torna possível atender a demanda de mercados externos, o que acaba fomentando a exportação.

Os resultados desse estudo contribuem para a pesquisas sobre biodiesel, comprovando que modelos estatísticos podem colaborar na busca por informações relacionadas a biocombustíveis. Esses dados também podem ser incorporados no planejamento do setor para tomada de decisão, tanto em níveis de produção quanto comercialização.

As limitações desse estudo podem ser atribuídas a quantidade de dados disponíveis, tanto como os diretamente ligadas a cadeia do biodiesel, quanto ao período temporal das series histórias necessárias para esse tipo de pesquisa. Como sugestão para estudos futuros, recomenda-se associar outros combustíveis renováveis, como álcool, para verificar relações de tendências e crescimento de produção desses biocombustíveis, tanto no Brasil como em outros países.

REFERÊNCIAS

- AGÊNCIA NACIONAL DO PETRÓLEO, GÁS NATURAL E BIOCMBUSTÍVEIS. **Biodiesel**. 2018. Disponível em: <<http://www.anp.gov.br/producao-de-biocombustiveis/biodiesel/simp-biodisel>>. Acesso em: 10 ago. 2019.
- AGÊNCIA NACIONAL DO PETRÓLEO, GÁS NATURAL E BIOCMBUSTÍVEIS. ANP. **Dados estatísticos**. 2019. Disponível em: <<http://www.anp.gov.br/dados-estatisticos>>. Acesso em: 06 ago. 2019.
- AKAIKE, H. A new look at the statistical model identification. IEEE Transactions on Automatic Control, v.19, n.6, p.716-723, Dec. 1974.

BRASIL. Empresa de Pesquisa Energética. Ministério de Minas e Energia. **Energia e aquecimento global**. Disponível em: <<http://epe.gov.br/pt/abcdenergia/energia-e-aquecimento-global>>. Acesso em: 10 ago. 2019.

BRASIL. EMPRESA DE PESQUISA ENERGÉTICA. **Matriz Energética e Elétrica**. 2018. Disponível em: <<http://epe.gov.br/pt/abcdenergia/matriz-energetica-e-eletrica>>. Acesso em: 10 ago. 2019.

BRASIL. Lei nº 13.263, de 23 de março de 2016. **Altera a Lei nº 13.033, de 24 de setembro de 2014, para dispor sobre os percentuais de adição de biodiesel ao óleo diesel comercializado no território nacional**. 2016. <http://www.planalto.gov.br/ccivil_03/_Ato2015-2018/2016/Lei/L13263.htm>. Acesso em: 10 ago. 2019.

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.; REINSEL, G.C. **Time series analysis: forecasting and control**. 3. ed. New Jersey: Printice Hall, 1994.

BUENO, R. de L. da S. **Econometria de séries temporais**. 2. ed. São Paulo: Cengage Learning, 2015.

CARRARO, Alecyr Reis; CÉSAR, Aldara da Silva; CONEJERO, Marco Antonio. **Potencial para produção de biodiesel no Brasil**. *AgroANALYSIS*, v. 38, n. 5, p. 21-22, 2019.

DICKEY, D. A.; FULLER, W. A. Likelihood ratio statistics for autoregressive time series with a unit root. *Econometrica*, v. 49, n. 4, p. 1057-1072, 1981.

GRANGER, C. W. J. Time Series Analysis, Cointegration, and Applications. *American Economic Review*, v. 94, n. 3, p. 421-425, 2004.

GUSTIN, M.; MCLEOD, R.S.; LOMAS, K.J. Forecasting indoor temperatures during heatwaves using time series models. *Building and Environment*, v. 143, p. 727-739, 2018.

INTERNATIONAL ENERGY AGENCY. **World Energy Balances**. IEA, 2019. Disponível em: <www.iea.org>. Acesso em: 10 ago. 2019.

INVESTING. **Soja**. Disponível em: <https://br.investing.com/commodities/us-soybeans> Acesso em: 10 ago. 2019.

IPEADATA. **Produto interno bruto (PIB)**. 2019. Disponível em: <<http://www.ipeadata.gov.br/Default.aspx>>. Acesso em: 06 ago. 2019.

KWIATKOWSKI, D. et al. Testing the null hypothesis of stationarity against the alternative of a unit root. *Journal of Econometrics*, v. 54, p. 159-178, 1992.

MARTINS, L. O. S.; CARNEIRO, R. A. F. O Brasil está pronto para o B20? Um estudo sobre a capacidade e potencialidade de produção de biodiesel. *Revista Textura*, v.10, n.19, p.123-129, 2017. Disponível em: Acesso em: 25 Jul. 2019.

MORETTIN, P. A. **Econometria financeira: um curso em séries temporais financeira**. 3. ed. ed. São Paulo: Blucher, 2016.

RATHMANN, Régis et al. **Biodiesel: uma alternativa estratégica na matriz energética brasileira**. II Seminário de Gestão de Negócios, v. 1, 2005.

ROCHA, F. J. S.; BRILHANTE, Á. A. **Introdução à metodologia de Vetor Autorregressivo (VAR): teoria e aplicação**. São Paulo: Iglu, 2015.

SILVA, Marcelo Santana et al. **Diagnósticos econômico e estrutural da cadeia produtiva do biodiesel no Brasil: um estudo de caso no estado da Bahia**. *RDE-Revista de Desenvolvimento Econômico*, v. 3, n. 41, 2019.

SOUZA, A. M. et al. Electrical energy supply for Rio Grande do Sul, Brazil, using forecast combination of weighted eigenvalues. *Gestão da Produção, Operações e Sistemas*, v. 6, n. 3, p. 23-39, jul./set. 2011.

SOUZA, F. M. **Modelos de previsão: aplicações à energia elétrica ARIMA-ARCH-AI e ACP**. 1 ed. Curitiba: Appris, 2016.

VIEGAS, T. O. C.; ARANTES, L. S. Políticas públicas para a ampliação da produção de biodiesel no Brasil. 2018. Revista Observatorio de la Economía Latinoamericana. Disponível em: <<https://www.eumed.net/rev/oel/2018/10/producao-biodiesel-brasil.html//hdl.handle.net/20.500.11763/oel1810producao-biodiesel-brasil>>. Acesso em: 25 Jul. 2019.