

**UNIVERSIDADE FEDERAL DE JUIZ DE FORA
FACULDADE DE CIÊNCIAS ECONÔMICAS
CURSO DE CIÊNCIAS ECONÔMICAS**

Carina Pinto Rodrigues

Modelo de previsão de mercado para o consumo de óleo diesel no Brasil

Juiz de Fora

2023

Carina Pinto Rodrigues

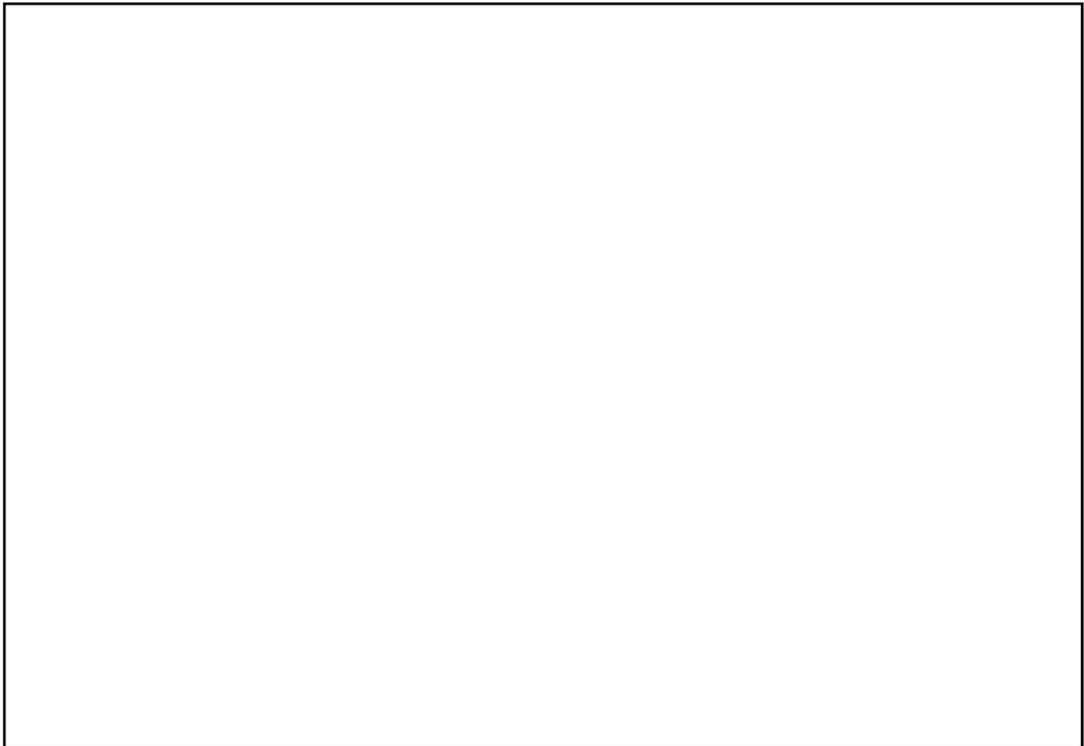
Modelo de previsão de mercado para o consumo de óleo diesel no Brasil

Monografia apresentada ao Curso de Ciências Econômicas da Universidade Federal de Juiz de Fora, como pré-requisito para a obtenção do título de Bacharel em Ciências Econômicas.

Orientador: Dr. Alexandre Zanini

Juiz de Fora

2023





UNIVERSIDADE FEDERAL DE JUIZ DE FORA
REITORIA - FACECON - Depto. de Economia

FACULDADE DE ECONOMIA / UFJF

ATA DE APROVAÇÃO DE MONOGRAFIA II (MONO B)

Na data de 04/07/2023, a Banca Examinadora, composta pelos professores

1 – Alexandre Zanini - orientador; e

2 – Fernando Salgueiro Perobelli,

reuniu-se para avaliar a monografia da acadêmica **CARINA PINTO RODRIGUES**, intitulada: **MODELO DE PREVISÃO DE MERCADO PARA O CONSUMO DE ÓLEO DIESEL NO BRASIL**.

Após primeira avaliação, resolveu a Banca sugerir alterações ao texto apresentado, conforme relatório sintetizado pelo orientador. A Banca, delegando ao orientador a observância das alterações propostas, resolveu **APROVAR** a referida monografia.

ASSINATURA ELETRÔNICA DOS PROFESSORES AVALIADORES



Documento assinado eletronicamente por **Alexandre Zanini, Professor(a)**, em 05/07/2023, às 09:25, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Fernando Salgueiro Perobelli, Professor(a)**, em 05/07/2023, às 10:00, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



A autenticidade deste documento pode ser conferida no Portal do SEI-Ufjf (www2.ufjf.br/SEI) através do ícone Conferência de Documentos, informando o código verificador **1352672** e o código CRC **98D390C0**.

RESUMO

O principal objetivo desse trabalho é prever e analisar o consumo do óleo diesel no Brasil. Os resultados obtidos no estudo podem auxiliar os agentes econômicos em relação à demanda futura, visto que é um combustível importante e utilizado em diversas áreas do país. A base de dados é composta por dados mensais da média diária do consumo de óleo diesel, medidos em milhares de barris, entre os períodos de janeiro de 1979 e janeiro de 2023. Essas informações representam o consumo aparente correspondente à produção industrial doméstica acrescida das importações e diminuída as exportações. Os métodos utilizados foram o Método Amortecimento Exponencial (MAE) e o Método de Box & Jenkins (MBJ). Foi feita uma análise recursiva fora da amostra para escolher o melhor método realizando-se uma competição entre os dois. O método vencedor foi o MBJ, estimando-se, ao final da análise de competição, um ARIMA(1,1,1)(1,0,2) com transformação de raiz quadrada. Este foi o modelo utilizado para calcular as previsões de consumo de óleo diesel no Brasil até dezembro de 2024. Através dos valores estimados verificou-se que é possível ocorrer um aumento do consumo em 2023 e 2024.

Palavras-chave: Óleo Diesel, Amortecimento Exponencial, Box & Jenkins, Competição de Métodos, Séries Temporais.

ABSTRACT

The main objective of this work is predict and analyze the consumption of diesel in Brazil. The results obtained in the study can assist the economic agents in relation to the future demand, since it is an important fuel and used in several areas of the country. The database consists of monthly data of the daily average of diesel fuel consumption, measured in thousands of barrels, between the periods of January 1979 and January 2023. This information represents the apparent consumption corresponding to domestic industrial production increased imports and decreased exports. The methods used were the exponential smoothing (ES) and Box & Jenkins (BJ) method. It was made a recursive analysis out of the sample to choose the best method by performing a competition between the two. The winning method was BJ, at the end of the competition analysis, an ARIMA(1,1,1)(1,0,2) with square root transformation. This was the model used to calculate diesel oil consumption forecasts in Brazil until December 2024. Through the estimated values it was found that it is possible to increase consumption in 2023 and 2024.

Key words: Diesel, Exponential Smoothing, Box & Jenkins, Competition of Methods, Time series.

LISTA DE FIGURAS

Figura 1 – Distribuição percentual da produção de derivados energéticos de petróleo - 2021.....	11
Figura 2 – Demanda mensal de óleo Diesel.....	12
Figura 3 – Demanda mensal de óleo Diesel S-10.....	13
Figura 4 – Procedimento de out-of-sample rolling evaluation	28
Figura 5 – Consumo de óleo diesel no Brasil (em milhares de barris)	32
Figura 6 – Função de Autocorrelação dos Erros – Método de Amortecimento Exponencial.....	35
Figura 7 – Função de Autocorrelação dos Erros – Método de Box & Jenkins	36
Figura 8 – Consumo de óleo diesel (em milhares de barris) – período de 1979 a 2024. Valores reais, ajustado e previsões com intervalo de confiança de 95%	39

LISTA DE QUADROS

Quadro 1 – Resumo das características teóricas da FAC e da FACP dos modelos AR(p), MA(q) e ARMA (p,q).....	26
---	----

LISTA DE TABELAS

Tabela 1 – Produção de derivados de petróleo, energéticos e não energéticos – 2012-2021	11
Tabela 2 – Parâmetro do Método de Amortecimento Exponencial	33
Tabela 3 – Fatores Sazonais.....	33
Tabela 4 – Parâmetros do Modelo de Box & Jenkins.....	35
Tabela 5 – Competição entre os Métodos.....	37
Tabela 6 – Valores da Previsão do Consumo de Óleo Diesel (em milhares de barris) – Intervalo de Confiança de 95%	39
Tabela 6 – Valores da Previsão do Consumo de Óleo Diesel (em milhares de barris) – Intervalo de Confiança de 95% - Continuação	40

SUMÁRIO

1	INTRODUÇÃO.....	7
2	MERCADO DE COMBUSTÍVEL	10
3	REVISÃO DE LITERATURA.....	14
4	METODOLOGIA.....	17
4.1	Séries Temporais e Processos Estocásticos	17
4.2	Método Amortecimento Exponencial (MAE)	18
4.3	Método Univariado de Box & Jenkins (MBJ)	24
4.4	Competição de Métodos	27
4.4.1	Desempenho Preditivo.....	29
5	ANÁLISE DE DADOS E RESULTADOS	32
5.1	Base de Dados.....	32
5.2	Equações de Previsão	33
5.3	Competição de Métodos	37
5.3	Previsão do Consumo de Óleo Diesel.....	38
6	CONCLUSÃO.....	41
	REFERÊNCIAS.....	43

1 INTRODUÇÃO

O combustível mais utilizado no Brasil é o óleo diesel. Suas aplicações vão além do setor automotivo, estão presentes na área industrial, agrícola, marítima, ferroviária e para geração de energia. No país, o setor rodoviário é o que mais utiliza o diesel, seu consumo está acima de 80%, segundo BRUNETTI (2012). Um aumento do consumo desse combustível se deu a partir do crescimento de demanda por serviços de transporte de passageiros, da sua utilização na geração de energia elétrica, e no crescimento do transporte de cargas (SOUZA; RODRIGUES; FURLAN, 2021).

Além disso, a Lei nº 11.097, de 13 de janeiro de 2005, foi criada com o intuito de acrescentar o biodiesel na matriz energética do Brasil, com o Programa Nacional de Produção e Uso de Biodiesel. Inicialmente, era necessário utilizar um percentual mínimo de 2% sobre o diesel comercial. A partir de 2008 esse volume aumentou. A expansão da demanda pelo derivado do diesel e uma estagnação na produção e para o refino do óleo, causou um aumento considerável na importação para o abastecimento doméstico (SOUZA; RODRIGUES; FURLAN, 2021).

Segundo Silva (2010), a grande procura por diesel altera de forma substancial a indústria de refino e a dependência energética do Brasil. Atualmente, não existe um substituto direto desse óleo, o que o torna fundamental para o desenvolvimento econômico.

Neste sentido, a demanda por este combustível é função de variações do PIB, por exemplo, expansão da fronteira produtiva e do grau de interação entre as esferas econômicas; urbanização; e, do ganho de eficiência da frota veicular (SILVA, 2010, p.52).

Devido à grande dependência em relação ao diesel, ocorreram avanços em diversos setores, como no setor de transporte, que contribuiu na propagação de inovações, alargamento de mercados e mobilidade de fatores de produção, no qual colaborou para o aumento de renda e da produção. (DELÁZARO; RODRIGUES; COSTILHES, 2013).

Outra área que pode ser destacada é a do agronegócio. Segundo o Centro de Estudos Avançados em Economia Aplicada (Cepea) em parceria com a Confederação da Agricultura e Pecuária do Brasil, estima-se que o setor fique por volta de 25,5% do Produto Interno Bruto (PIB) do país (CEPEA, 2022).

A pecuária e a agricultura são responsáveis por uma cadeia de produção que abrange fabricantes de maquinário, veterinários e transportadores a lojistas e consumidor. Além disso, não se pode mencionar fornecimento de alimentos e geração de emprego e renda sem a atuação do agronegócio. (CENTENARIO DIESEL, 2020)

O óleo diesel é fundamental para este cenário, pois é o combustível mais usado nos motores pesados e grupos geradores, sendo essencial para o funcionamento das áreas rurais. (CENTENARIO DIESEL, 2020).

Porém, o aumento no consumo e as alterações na estrutura de suprimentos do diesel vêm causando preocupações aos agentes econômicos. Por esse motivo, faz-se necessário uma investigação na estrutura de demanda do óleo, principalmente no que diz respeito a gestão de importação, distribuição e precificação (SOUZA; RODRIGUES; FURLAN, 2021).

Ao observar a importância do óleo diesel para o desenvolvimento do país, justifica-se elaborar um estudo que pode gerar informações importantes para subsidiar as decisões de agentes econômicos em relação à demanda de consumo futuro, com foco na produção industrial doméstica somada as importações e diminuída das exportações.

Desta forma, este trabalho monográfico tem como objetivo geral estudar e modelar a evolução do consumo de óleo diesel no Brasil do período de janeiro de 1979 a janeiro de 2023, no qual o ano inicial escolhido representa uma data muito importante, pois foi quando ocorreu o segundo choque do petróleo, que levou ao aumento do preço do petróleo, deterioração da balança comercial, processo inflacionária, desemprego, entre outros problemas (MELO, 2008). Através dessas informações, foi possível estimar o consumo para o intervalo de fevereiro de 2023 a dezembro de 2024. As previsões do modelo foram realizadas com dados mensais retirados do site Ipeadata¹. Este trabalho também tem como propósito contribuir para os setores de combustíveis e gerar informações para auxiliar um melhor planejamento de políticas de demanda e da oferta do óleo diesel. Além disso, ele tem o intuito de ampliar o conhecimento disponível na literatura sobre o tema.

Os objetivos específicos foram: pesquisar utilizações e segmentos do óleo diesel no Brasil; capturar a base de dados que foi utilizada no trabalho; estabelecer o modelo estatístico que melhor se adaptou para a realização das previsões; e verificar

¹ www.ipeadata.gov.br

se o método utilizado possuía um desempenho preditivo que permitisse a utilização do modelo, gerando as previsões de consumo de óleo diesel.

Feitas estas considerações, atenta-se que este trabalho está dividido da seguinte forma: Introdução, revisão de literatura, com o intuito de apresentar contribuições de pesquisas e discussões importantes sobre o tema, metodologia utilizada, análise de dados e resultados e conclusão.

2 MERCADO DE COMBUSTÍVEL

Devido à importância dos combustíveis no Brasil, fez-se necessário a criação de um órgão capaz de regulamentá-los. No dia 6 de agosto de 1997, foi criada a Lei nº 9.478, mais conhecida como a Lei do Petróleo, que logo após o decreto nº 2.455 em 14 de janeiro de 1998, tornou-se a Agência Nacional do Petróleo. Esse decreto tinha como objetivo direcionar o monopólio do petróleo e do gás natural à ANP, na qual ficou responsável em fiscalizar e regular o setor (BRASIL,1997).

Além das atribuições de regular, contratar e fiscalizar a indústria do petróleo, a ANP também controla a qualidade dos derivados do petróleo e autoriza a importação e exportação desses produtos (BRASIL,1997). É um órgão importante para se obter dados e informações.

O petróleo e seus derivados têm uma grande importância na economia do país, e seus bens são gerados a partir de insumos de difícil substituição na matriz produtiva. Exemplos disso são a gasolina e o diesel, utilizados como combustíveis de motores de combustão interna dos veículos de transporte e serviços de infraestrutura. A energia e o transporte são essenciais na produção de bens e serviços. (CANELAS, 2007).

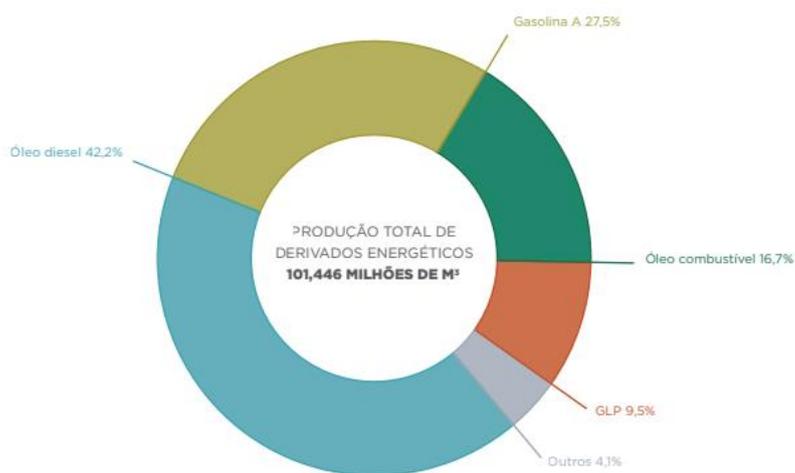
No ano de 2021, os derivados de petróleo tiveram uma produção 3,8 milhões de m³ a mais do que no ano de 2020. Os derivados energéticos equivaleram a 87,6% do total produzido, com um aumento de 5,6% em relação a 2020, e o óleo diesel representou uma variação positiva de 1,5% comparada ao ano anterior, como pode ser visto na tabela 1. A figura 1, por sua vez, mostra a porcentagem da produção dos derivados energéticos. Como já era esperado, o diesel teve uma maior variação (42,2%), visto que é o combustível mais utilizado no país (ANP, 2022).

Tabela 1 – Produção de derivados de petróleo, energéticos e não energéticos – 2012-2021

DERIVADOS DE PETRÓLEO	PRODUÇÃO (M³)										21/20 %
	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	
TOTAL	120.203.986	127.299.882	130.152.722	122.120.793	114.418.187	110.358.366	108.309.562	108.335.341	112.024.495	115.778.241	3,35
Energéticos	102.528.383	110.160.883	112.717.562	106.717.844	100.230.490	96.725.331	94.285.616	94.009.496	96.047.553	101.446.027	5,62
Gasolina A	27.061.075	29.720.707	30.078.550	26.923.072	27.719.573	27.774.987	25.169.472	25.257.431	23.395.022	27.903.803	19,27
Gasolina de aviação	77.606	93.685	93.762	72.486	53.902	59.662	46.220	-	2.216	55.899	..
GLP	10.361.616	10.228.151	10.050.965	9.897.467	9.663.122	10.336.423	9.979.417	9.815.071	9.858.567	9.600.855	-2,61
Óleo combustível	13.691.084	14.761.276	16.267.891	14.339.295	11.506.738	11.691.291	10.751.048	11.866.979	17.238.550	16.952.832	-1,66
Óleo diesel	45.504.004	49.539.186	49.675.057	49.457.609	45.369.807	40.625.830	41.957.390	40.998.507	42.215.122	42.852.980	1,51
QAV	5.422.769	5.554.391	6.079.114	5.656.859	5.789.278	6.168.600	6.376.333	6.066.674	3.332.770	4.074.103	22,24
Querosene iluminante	23.885	15.393	12.005	7.396	7.668	5.830	5.736	4.834	5.306	5.554	4,66
Outros	386.345	248.094	460.217	363.660	120.403	62.709	-	-	-	-	..
Não energéticos	17.675.603	17.138.999	17.435.160	15.402.949	14.187.697	13.633.035	14.023.947	14.325.845	15.976.942	14.332.214	-10,29
Asfalto	2.569.635	2.653.348	3.248.853	2.015.366	2.152.075	1.955.427	1.899.816	1.721.537	2.318.275	1.875.770	-19,09
Coque	4.452.350	4.810.510	4.748.864	4.958.620	5.076.586	4.928.529	4.468.572	4.411.889	4.332.411	4.367.408	0,81
Nafta	6.440.115	5.354.014	5.074.640	4.608.816	3.175.691	3.086.305	4.046.483	4.480.133	6.194.691	4.623.048	-25,37
Óleo lubrificante	607.979	689.214	682.053	640.490	616.529	593.536	602.881	566.063	454.679	621.662	36,73
Parafina	123.445	122.647	134.636	136.934	162.366	120.051	126.197	124.131	73.600	57.115	-22,40
Solvente	290.241	454.262	384.262	358.134	336.158	326.117	330.009	312.790	317.674	411.408	29,51
Outros	3.191.837	3.055.004	3.161.852	2.684.589	2.668.293	2.623.069	2.549.988	2.709.302	2.285.612	2.375.803	3,95

Fonte: ANP (2022)

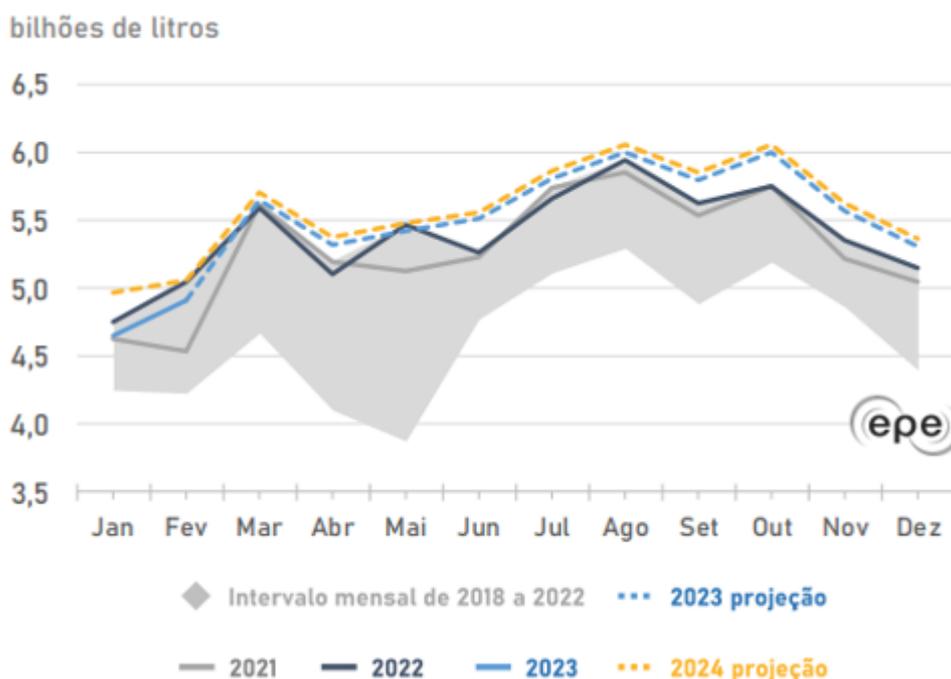
Figura 1 – Distribuição percentual da produção de derivados energéticos de petróleo - 2021.



Fonte: ANP (2022)

Estes dados mostram o aumento da demanda de combustíveis no Brasil, principalmente do óleo diesel, no qual teve um acréscimo de 2% em 2022, marcando um recorde histórico de 64,7 bilhões de litros. De acordo com a Empresa de Pesquisa Energética (EPE) (2023), o aumento do consumo em 2022 se deu em grande parte devido ao escoamento da produção agrícola e o bom desempenho das exportações. Segundo o Instituto Brasileiro de Geografia e Estatística (IBGE) (2023), pesquisas apontam que o país poderá alcançar um recorde no transporte de cargas. Além disso, outras fontes responsáveis no crescimento da demanda de óleo diesel são os transportes coletivos de ônibus. Na figura 2 é possível ver o crescimento ao longo dos meses em relação a demanda do diesel e a projeção de como será em 2024.

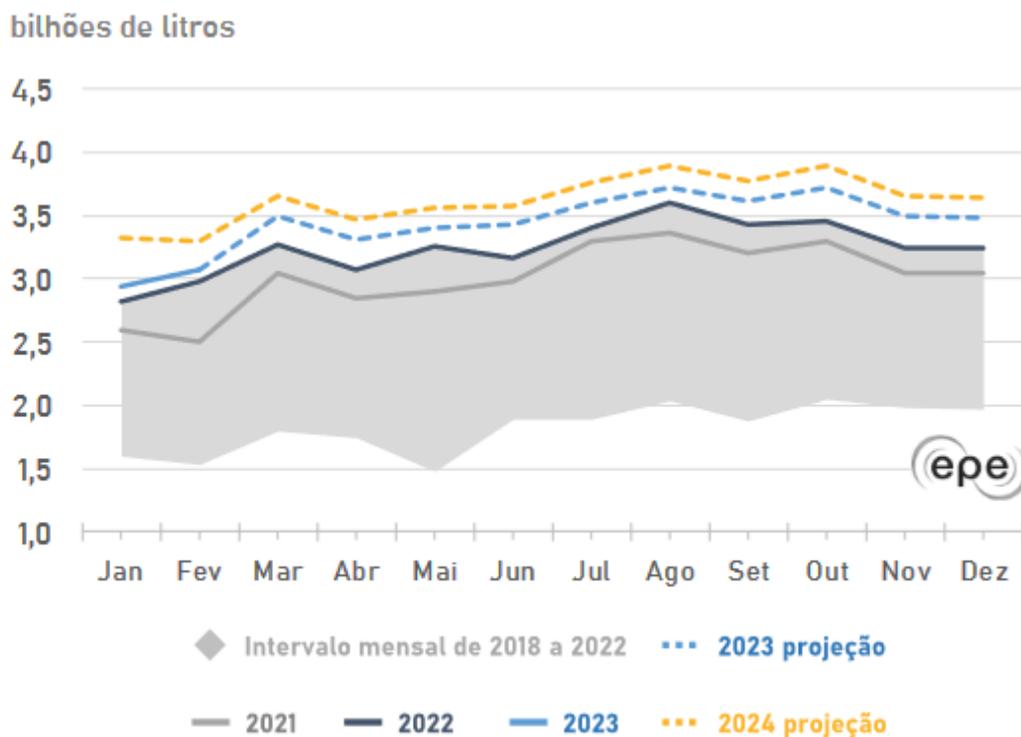
Figura 2 – Demanda mensal de óleo Diesel.



Fonte: EPE (2023a)

A figura 3 representa a demanda por óleo diesel S10, cujo uso é o transporte rodoviário. A Demanda em 2023 teve um aumento de 6%, o que equivale a 41,3 bilhões de litros, e a projeção para 2024 é que ocorra um aumento de 5,4% em relação ao ano anterior, o que equivale a 43,5 bilhões de litros. (EPE, 2023).

Figura 3 – Demanda mensal de óleo Diesel S-10



Fonte: EPE (2023b)

Evidenciada a importância do óleo diesel, no próximo capítulo abordar-se-á alguns trabalhos que se dedicaram ao estudo deste combustível no decorrer do tempo.

3 REVISÃO DE LITERATURA

A partir da análise de diversas pesquisas, foi possível encontrar alguns projetos que utilizaram diferentes abordagens em relação à previsão do óleo diesel e da sua importância para a economia do país. Bastos (2014) aplicou diferentes tipos de métodos para prever a produção nacional de óleo diesel para o período de 2014 a 2015, analisando o intervalo de janeiro de 2000 a agosto de 2014. Foi feito um estudo da série temporal em relação à produção mensal e uma análise dos erros das previsões. De acordo com as características da base de dados, os modelos aplicados foram: o de decomposição clássica, suavização exponencial de Holt, sazonal de Holt-Winter e o modelo ARIMA de Box e Jenkins. O objetivo final do pesquisador foi identificar qual desses métodos era o melhor, que mais se aproximou da realidade, além de apresentar o menor erro possível.

Em outra perspectiva de análise, Moreira (1996) projetou um modelo a nível nacional e outro a nível regional do consumo do óleo diesel no Brasil a longo prazo. Alguns dos objetivos do projeto foi buscar os determinantes do consumo e prever as variáveis envolvidas. Em nível nacional, ele determinou que as variáveis de preço e nível global de atividades eram as mais relevantes para determinar o consumo do diesel, e em nível regional eram as variáveis envolvendo o consumo nacional de óleo diesel projetado e as tendências de crescimento do consumo de cada estado. O autor utilizou uma amostra disponível do consumo e do preço do combustível com dados mensais de janeiro de 1980 a dezembro de 1994 para a previsão no nível nacional e o período de 1970 a 1992 para o nível regional. O modelo empregado foi o de correção de erros (VEC). Os principais resultados do trabalho foram a existência de alguma relação do consumo de óleo *per capita* com o PIB *per capita* e que o consumo das regiões é afetado pelo nível global de atividades do país.

Caetano (2013) apresenta um mapeamento do mercado e a projeção do consumo do diesel em um horizonte de 4 anos. Além disso, através de ferramentas estatísticas e análises descritivas, foi possível comprovar a importância do óleo diesel para a economia do país e a relevância do setor de transporte rodoviário para o consumo do diesel. O método utilizado é o de *forward* de Regressão Múltipla e foram realizados também testes de correlação entre as variáveis importantes para o consumo. Além de mapear o mercado de óleo diesel no Brasil, este trabalho teve

como objetivo apresentar tomadas de decisões, como previsões de crescimentos setoriais, de capacidade e de demanda, e aspectos técnicos que pudessem auxiliar ações governamentais e investimentos estruturais e de infraestrutura. Os principais fatores para a escolha do método foram o número reduzido de dados confiáveis e estruturados, a robustez da metodologia e a sua simplicidade de compreensão e implementação. Na primeira etapa, foi possível verificar as variáveis que poderiam influenciar o consumo do óleo diesel. Em seguida, procurou-se uma forma de comprovar estatisticamente a relação dessas variáveis com o consumo. A importância do setor rodoviário no consumo do óleo diesel foi relatada de duas formas: a primeira foi quando o consumo se dividiu por setor e modal, que mostrou a relevância do setor de transporte no consumo entre os setores, e do modal rodoviário no consumo do setor de transportes; a segunda ocorreu na construção do modelo em que a variável movimentação de carga no modal rodoviário era a mais significativa.

Brafman (2009) propôs um modelo para projetar o consumo de ciclo otto (gasolina, álcool e GNV) e diesel no Brasil e comparou com um modelo ARMA. O objetivo era apurar qual deles seria o melhor em relação ao processo de distribuição de renda e venda de combustíveis. Através dessa análise, seria possível desenvolver planos de logísticas e orçamentos para o mercado. Foram utilizados dados de janeiro de 2001 a abril de 2008 para estimar o modelo de maio de 2008 a abril de 2009, totalizando um período de 12 meses. O autor escolheu uma metodologia específica de regressão linear múltipla, mas obteve um modelo de regressão simples, pois houve uma forte correlação do consumo do óleo diesel com uma das variáveis (cargas movimentadas pelo modal rodoviário), dispensando todas as outras envolvidas. Por fim, o autor concluiu que o seu modelo proposto foi mais efetivo do que o modelo ARMA.

Rodrigues e Losekann (2015) analisou os determinantes da demanda por óleo diesel do transporte terrestre no Brasil, realizando projeções de demanda nacional por diesel para o período entre 2014 e 2022. Foi utilizado um modelo de painel dinâmico e depois um modelo econométrico para realizar as previsões em dois cenários de crescimento econômico: cenário base e otimista da evolução do PIB nacional. A estimação foi feita com dados estaduais anuais a partir de uma especificação de painel dinâmico. O objetivo do trabalho foi contribuir para o planejamento do setor de combustível, a fim de subsidiar a elaboração de políticas de planejamento de gestão de demanda e oferta do setor energético. Os resultados indicaram que, nos dois

casos, haveria um significativo aumento da demanda por óleo diesel no período em questão, os consumidores são mais sensíveis a preços do que a renda, o que significa que, as políticas voltadas para o controle de demanda serão melhores através de variações no preço do combustível.

Uchôa et al. (2020) também desenvolveu um estudo para estimar a demanda por diesel, no qual seu objetivo era fornecer elasticidades de demanda para os três combustíveis mais usados no Brasil: gasolina, etanol e diesel. Foram utilizados dados a nível municipal para o período de 2007 a 2016. Algumas das conclusões foram que as demandas dos três combustíveis são elásticas de preço, o aumento do PIB afeta positivamente a demanda por gasolina e diesel, mas não afeta o etanol e que a relação preço etanol/gasolina é uma variável relevante para a demanda tanto de gasolina quanto de etanol.

Apresentados alguns trabalhos que se dedicaram ao estudo do óleo diesel, e o que os diferencia deste projeto é o uso de dois diferentes métodos e a realização da competição entre eles para encontrar qual possui o melhor desempenho. Atenta-se que, no próximo capítulo, será apresentada a metodologia aplicada no presente trabalho monográfico e que consiste basicamente numa competição de dois métodos autoprojativos, selecionando o melhor deles para gerar as previsões de consumo.

4 METODOLOGIA

Atenta-se que toda esta seção é baseada nos conceitos metodológicos do trabalho de Zanini (2023). Neste projeto serão utilizados os métodos estatísticos de Amortecimento Exponencial (MAE)² e de Box & Jenkins (MBJ)³ para selecionar o melhor modelo, o qual será utilizado para calcular as previsões de consumo diesel no Brasil.

4.1 Séries Temporais e Processos Estocásticos

Uma série temporal são observações de uma dada variável, ordenadas no tempo em intervalos iguais. Se Z_t representa o valor da variável aleatória Z no instante t , a série temporal por Z_1, Z_2, \dots, Z_N onde N é o número de observações seriais da variável. Para as séries temporais discretas, os períodos de coleta de dados são distribuídos em dados diários, semanais, mensais e anuais.

Um modelo que representa a probabilidade de uma sequência de observações é chamado de processo estocástico, ou seja, é uma sequência no tempo de variáveis aleatórias. Este processo é uma família $Z=\{Z(t), t \in \mathbf{N}\}$ tal que para cada $t \in \mathbf{R}$, $Z(t)$ é uma variável aleatória. Então, com uma série temporal discreta, onde $\mathbf{N} \equiv \mathbf{Z}=\{1, \dots, t\}$, $\mathbf{N} \in \mathbf{R}$, o Z_t é uma denotação do valor de Z no período t . Supondo que Z_t faz parte de um experimento que pode ser repetido de idênticas condições, a cada experimento é possível obter um registro dos valores de Z_t no tempo.

Segundo Souza & Camargo (1996), o processo estocástico está relacionado estatisticamente a funções de distribuição que conhecem suas funções até a n -ésima ordem. Existem algumas dificuldades devido à falta de conhecimento de todas as funções de distribuição até a n -ésima ordem e por possuir somente uma realização do processo estocástico. Desta forma, podem ocorrer algumas situações problemáticas. Para isso, existem duas restrições a serem consideradas: Estacionaridade, que ocorre caso o processo estocástico gere uma série de

² Montgomery & Johnson (1990).

³ Box & Jenkins (1994).

observações invariantes no tempo; e a Ergodicidade, no qual o processo estocástico ocorre apenas uma vez e é suficiente para obter todas as suas estatísticas.

Para este trabalho, o objetivo é estudar e manipular métodos de análises de séries temporais, a fim de, atingir o modelo estocástico existente. Os modelos de previsão podem ser autoprojativos e/ou causais. Nos modelos autoprojativos (amortecimento exponencial e univariado de Box & Jenkins) utilizam-se valores passados da amostra. Já nos modelos causais (Regressão Dinâmica), outras informações possivelmente relevantes são consideradas. Utilizam-se a análise do histórico de outras variáveis. Os métodos de análises de séries temporais utilizados serão apresentados abaixo.

4.2 Método Amortecimento Exponencial (MAE)

No método de amortecimento exponencial pode-se dar um peso maior para informações atuais nas séries que possuem maior variância. Já as séries com menor variância, é possível ponderar de forma igual os dados presentes e os dados passados. Neste modelo, pode-se incluir parâmetros de tendência e sazonalidade para definir o comportamento da série temporal. Além disso, são desenvolvidos métodos de atualização destes parâmetros sem interferir nos pesos que são dados a informações no presente e no passado.

A partir disso, será feito uma breve descrição do método de amortecimento exponencial. Considerando o seguinte conjunto de dados Z_1, Z_2, \dots, Z_T seja uma série temporal de tamanho "T".

Supondo que esta série represente um produto de determinada empresa cuja demanda mensal não apresenta uma variação significativa no seu nível ao longo do tempo, ou seja, não ocorrem mudanças no nível de venda com o tempo ou, se ocorrem, são variações pouco significativas. Então, para este produto, a equação de previsão pode ser representada por:

$$Z_t = a(T) + \varepsilon_t \quad (4.2.1)$$

Onde:

Z_t = venda no período t ($t = 1, 2, \dots, T$);

$a(T)$ = parâmetro representativo do nível médio das vendas no instante T ;

ε_t = erro de previsão e $\varepsilon_t \sim N(0, \sigma^2)$.

Pode-se observar que, dada as particularidades deste produto (nível de venda mais ou menos constante), a estimativa ($\hat{a}(T)$) para o parâmetro “a” na equação (4.2.1) fornece exatamente uma previsão para o produto observado. Por este motivo, o parâmetro pode ser estimado de diversas maneiras. Dentro de uma lógica “autoprojetiva”, poder-se-ia estimá-lo, por exemplo, através do modelo ingênuo (ou *naive*) e que utiliza como predictor o último dado conforme indicado pela equação (4.2.2) abaixo:

$$\hat{Z}_T(\tau) = Z_T \quad (4.2.2)$$

Onde:

$\hat{Z}_T(\tau)$: previsão para Z_T , τ passos-à-frente (ou previsão de $Z_{T+\tau}$ feita no instante T);

Z_T : último dado disponível

τ : horizonte de previsão

O modelo tem como finalidade fazer um sistema no qual possa reestimar os parâmetros a cada período, introduzindo a novas informações. No final do período T , é possível obter duas informações:

1 - A estimativa de “a” feita no final do período anterior $\hat{a}(T-1)$;

2 - O último dado disponível $Z(T)$;

Assim sendo, com estas informações será possível calcular uma estimativa atualizada do nível de venda: $\hat{a}(T)$. A sugestão para resolver o sistema é fazer uma alteração na estimativa velha ($\hat{a}(T-1)$) do nível por uma fração do erro de previsão causado pelo uso desta estimativa para o resultado mais atual. Considerando que o erro de previsão no último período pode ser definido como $\varepsilon(T) = Z(T) - \hat{a}(T-1)$, a formulação matemática para esta proposta está representada nas equações abaixo:

$$\hat{a}(T) = \hat{a}(T-1) + \alpha * [Z(T) - \hat{a}(T-1)] \quad (4.2.3)$$

$$\hat{a}(T) = \alpha * Z(T) + (1 - \alpha) * \hat{a}(T-1) \quad (4.2.4)$$

Para simplificar a notação, define-se que $\hat{a}(T) \equiv S_T$. Desta forma, temos que a equação (4.2.4) acima pode ser representada como:

$$S_T = \alpha * Z_T + (1 - \alpha) * S_{T-1} \quad (4.2.5)^4$$

Onde:

S_T = valor amortecido da série;

α = constante de amortecimento ou hiperparâmetro (número entre 0 e 1).

Pode-se concluir que no método de amortecimento exponencial será feita uma atualização automática no parâmetro "a", que representa o nível de vendas na equação (4.2.1), no qual é dado um peso α para o "presente" (último dado) e um peso $(1 - \alpha)$ para o "passado" (estimativa anterior para o nível e que pode ser obtida de várias formas como, por exemplo, uma média, uma média móvel, dentre outras)⁵.

Porém, este modelo não é ideal na presença de alterações do nível da série, ou seja, na presença de um componente de tendência. Para este tipo de série, o modelo mais adequado é o seguinte⁶:

$$Z_t = (a_1(T) + a_2(T) * t) + \varepsilon_t \quad (4.2.6)$$

Onde:

$a_1(T)$: parâmetro de nível no instante T;

$a_2(T)$: parâmetro de tendência no instante T;

t: variável tempo (t = 1, 2, ..., T sendo T é a quantidade de dados existentes);

⁴ A equação (4.2.5) é também conhecida como Modelo de Brown.

⁵ Para mais detalhes, ver MONTGOMERY & JOHSON (1990). Como um exemplo, imagine que se esteja trabalhando com dados mensais no período de janeiro de 2001 a dezembro de 2005. Neste caso, a equação (2.6) informa que a atualização do parâmetro de nível (portanto a previsão para o dado de janeiro de 2006, dado que o modelo para as vendas é constante) será feita dando-se um peso α para o dado de dezembro e um peso $(1 - \alpha)$ para a última estimativa feita para o dado de dezembro de 2005 (feita obviamente em novembro de 2005).

⁶ Verifique que, diferentemente da equação (4.2.1), existem agora dois parâmetros e por isto é feita a diferença entre a_1 e a_2 .

ε_t é o erro de previsão e $\varepsilon_t \sim N(0, \sigma^2)$.

É importante destacar que um sistema de atualização paramétrica similar a equação (4.2.5) será aplicada sobre os parâmetros da equação (4.2.6). Este modelo é o Modelo de Holt-2Parâmetros. A atualização dos parâmetros pode ser feita como nas equações (4.2.7) e (4.2.8) a seguir:

$$\hat{a}_1(T) = \alpha * Z_T + (1 - \alpha) * \left[\hat{a}_1(T-1) + \hat{a}_2(T-1) \right] \quad (4.2.7)$$

$$\hat{a}_2(T) = \beta * \left[\hat{a}_1(T) - \hat{a}_1(T-1) \right] + (1 - \beta) * \left[\hat{a}_2(T-1) \right] \quad (4.2.8)$$

Pode-se observar que as equações acima possuem a mesma ideia de se ponderar o "presente" e "passado" para realizar a atualização dos parâmetros. Porém, são usadas duas constantes de amortecimento (α e β), uma para o parâmetro de nível (a_1) e outra para o parâmetro de tendência (a_2).

Na equação (4.2.7), a atualização do parâmetro de nível é feita dando um peso α para o dado real mais recente e um peso $(1-\alpha)$ para a última estimativa feita para o nível que é composta por: $\hat{a}_1(T-1)$, estimativa feita para o nível no instante $(T-1)$, mais $\hat{a}_2(T-1)$, que é a estimativa feita para a tendência também no instante anterior $(T-1)$.

Até o presente momento, foi possível abordar a estimativa de dois parâmetros (nível e tendência). Porém, existe um certo comportamento que pode influenciar no comportamento periódico das vendas, ou seja, dependendo da época do ano, ocorre um incremento ou decréscimo nas vendas. Em séries temporais, este comportamento é conhecido como "sazonalidade", ou seja, um movimento periódico (cíclico) da série no decorrer do tempo.

Neste caso, o modelo mais adequado pode⁷ ser o expresso na equação (4.2.9):

$$Z_t = (a_1(T) + a_2(T) * t) * \rho_t + \varepsilon_t \quad (4.2.9)$$

$a_1(T)$: parâmetro de nível no instante T;

$a_2(T)$: parâmetro de tendência no instante T;

⁷ É dito "pode", pois a equação apresenta uma sazonalidade multiplicativa, entretanto, pode-se encontrar também uma sazonalidade aditiva

t: variável tempo (t = 1, 2, ..., T sendo T é a quantidade de dados existentes);
 ρ_t : fator sazonal referente ao período t;
 ε_t é o erro de previsão e $\varepsilon_t \sim N(0, \sigma^2)$.

Este modelo é conhecido como **Modelo de Holt-Winters**. Na equação (4.2.9) foi incluído um novo parâmetro com o objetivo de “captar” o efeito da sazonalidade sobre as vendas⁸. Atenta-se que um procedimento de atualização paramétrica similar ao expresso na equação (4.2.5) será utilizado para atualizar sequencialmente os parâmetros deste modelo. Este procedimento pode ser visualizado nas equações (4.2.10) a (4.2.12) a seguir:

$$\hat{a}_1(T) = \alpha * \frac{Z_T}{\hat{\rho}_t(T-1)} + (1-\alpha) * \left[\hat{a}_1(T-1) + \hat{a}_2(T-1) \right] \quad (4.2.10)^9$$

$$\hat{a}_2(T) = \beta * \left[\hat{a}_1(T) - \hat{a}_1(T-1) \right] + (1-\beta) * \left[\hat{a}_2(T-1) \right] \quad (4.2.11)^{10}$$

$$\hat{\rho}_t(T) = \gamma * \frac{Z_T}{\hat{a}_1(T)} + (1-\gamma) * \left[\hat{\rho}_t(T-1) \right] \quad (4.2.12)$$

Pela equação (4.2.10) acima, percebe-se que o nível estimado é função da última observação, do fator sazonal estimado no instante anterior (mesmo mês do ano anterior) e estimativas anteriores do nível e tendência. Vê-se que o termo $\frac{Z_T}{\hat{\rho}_t(T-1)}$ representa a observação no instante T dessazonalizada. Já pela equação (4.2.11),

⁸ Existem uma restrição de normalização que faz com que $\sum_{i=1}^L \rho_i = L$, onde L é o comprimento do período sazonal. Isto é, caso se esteja trabalhando com dados mensais e um parâmetro de sazonalidade multiplicativo, a soma dos fatores sazonais precisa ser igual a 12 (obviamente pois existem 12 meses em um ano).

⁹ Onde $\hat{\rho}_t$ é o fator sazonal correspondente ao período (mês, trimestre, etc) t.

¹⁰ Trata-se da mesma equação (4.2.8).

observa-se que a taxa de crescimento (tendência) estimada é função dos níveis estimados em T e T-1 e da taxa de crescimento estimada anteriormente¹¹. Por fim, pela equação (4.2.12), vê-se novamente a ideia básica do método de amortecimento exponencial de atualizar os parâmetros do modelo atribuindo pesos diferenciados à “idade da informação”, ou seja, ponderando “presente” e “passado” de forma diferenciada. Esta equação indica que o fator sazonal correspondente ao período T é função do fator sazonal correspondente ao mesmo período no ano anterior e da última observação. Logo, observa-se que o fator sazonal correspondente a um certo “mês”, por exemplo, só é atualizado uma vez por ano, ao se receber o dado referente àquele “mês”. Pode-se observar também que existem três constantes de amortecimento (α , β e γ), uma para cada parâmetro (nível, tendência e sazonalidade) a ser atualizado.

A constante de amortecimento pode ser definida como uma quantidade fixa que é utilizada para fazer a atualização sequencial dos parâmetros. Este conceito caracteriza o modelo obtido através do método de amortecimento exponencial como um modelo com “validade local”, ou seja, a cada instante de tempo, a cada dado real que chega, é feita uma atualização dos parâmetros do modelo com base nas equações descritas anteriormente, sendo que os fatores responsáveis por esta atualização são exatamente as constantes de amortecimento ou hiperparâmetros.

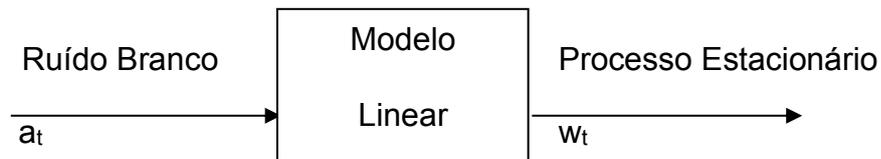
Em determinadas situações, faz-se necessário intervir nos valores projetados por um modelo do tipo Holt (equação (4.2.6)). No modelo de Holt, vê-se que são estimados dois parâmetros, um para o nível e outro para a tendência. Dependendo da “magnitude” do parâmetro de tendência, pode-se gerar previsões “explosivas” (muito altas) no horizonte de previsão. Uma maneira de se corrigir este problema, é realizar o que se denomina de *damped trend*. Neste procedimento, inclui-se mais um hiperparâmetro no modelo conforme a equação 4.2.13 a seguir. Esta equação já representa a equação de previsão. Vê-se que a função deste hiperparâmetro é reduzir a tendência no horizonte de previsão.

$$Z_t = \hat{a}_1(T) + \sum_{j=1}^{\tau} \varphi^{j-1} * \hat{a}_2(T) * \tau + \varepsilon_t \quad (4.2.13)$$

¹¹ Como visto anteriormente, a constante de amortecimento (β) é diferente daquela (α) usada para atualização do nível.

4.3 Método Univariado de Box & Jenkins (MBJ)

O modelo de Box & Jenkins tem como fundamento teórico a Teoria Geral de Sistemas Lineares, que diz que um processo estacionário de segunda ordem¹² é gerado através de uma passagem de um ruído branco por um filtro linear de memória infinita.



Definindo um operador de atraso "B" (*backward shift operator*)

$$B^k Z_t = Z_{t-k}$$

Tem-se que:

$$w_t = a_t - \Psi_1 a_t B - \Psi_2 a_t B^2 - \dots$$

$$w_t = (1 - \Psi_1 B - \Psi_2 B^2 - \dots) a_t = \Psi(B) a_t$$

$$w_t = \Psi(B) a_t \Leftrightarrow a_t = \Psi(B)^{-1} w_t$$

Sendo que:

$$\Psi(B)^{-1} = \pi(B)$$

$$\text{onde } \pi(B) = 1 - \pi_1 B - \pi_2 B^2 - \dots$$

Como $\Psi(B)$ [ou $\pi(B)$] possuem infinitos parâmetros, causando um problema, Box & Jenkins utilizam o fato de que, sob certas condições, pode-se dizer que todo polinômio infinito pode ser expresso pelo quociente de dois polinômios finitos.

Então:

$$\Psi(B) = \theta(B) / \phi(B), \text{ onde:}$$

$$\theta(B) = 1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q \Rightarrow \text{Polinômio MA (q)}$$

$$\phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p \Rightarrow \text{Polinômio AR (p)}$$

¹² Em termos gerais, processo estacionário de 2ª ordem é um processo estocástico com média e variância constantes.

Desta forma, surgem os denominados Modelos ARMA (p,q) que assumem a seguinte forma:

$$\phi(B)w_t = \theta(B)a_t \quad (4.3.1)$$

Porém, para formar séries não estacionárias na média, a estacionariedade é produzida através da diferenciação da série original, construindo uma série não estacionária homogênea¹³. Logo, se Z_t é uma série não estacionária, procurar-se-á transformá-la na série X_t não estacionária homogênea. Então:

$$X_t = Z_t - Z_{t-1} = Z_t - BZ_t = (1 - B) Z_t = \nabla Z_t$$

Onde: $\nabla = (1 - B) \Rightarrow$ operador de diferença

Pode-se dizer que, serão aplicadas tantas diferenças forem precisas para produzir estacionariedade (na média) da série resultante. Considerando uma série W_t estacionária, ela será obtida pela aplicação de "d" ($d = 0, 1, 2, \dots$)¹⁴ diferenças na série original: $w_t = \nabla^d Z_t$. Desta forma, surgem os chamados modelos ARIMA(p,d,q), que possuem a seguinte forma:

$$\phi(B)\nabla^d Z_t = \theta(B)a_t \quad (4.3.2)$$

Box & Jenkins têm como fundamento o princípio da parcimônia (a escolha de um modelo com o menor número de parâmetros possíveis para uma representação matemática adequada) e a construção de modelos através de um ciclo iterativo (estratégia de seleção de modelos até a obtenção de um modelo satisfatório). Portanto, várias etapas de análise, como a identificação da estrutura do modelo, abrangendo a estimação paramétrica e por fim, realizando vários testes de validação do modelo fazem parte da modelagem da metodologia de Box & Jenkins.

O primeiro passo metodológico tem como finalidade a identificação da ordem de homogeneidade "d", determinando o número de vezes em que a série original deve ser diferenciada a fim de, tornar-se uma série estacionária. Este processo acontece através do gráfico da série ou da função de autocorrelação (FAC)¹⁵.

O segundo passo ocorre com a identificação da ordem do modelo (identificação de p e q). São analisados as funções de autocorrelação (FAC) e autocorrelação parcial

¹³ A não estacionariedade homogênea exclui os processos de comportamento explosivos e/ou altamente não-lineares (Souza & Camargo, 1996).

¹⁴ Na prática, normalmente o grau máximo de diferenciação é 2.

¹⁵ A Função de Autocorrelação para uma série não estacionária apresenta um lento decréscimo.

(FACP). No Quadro 1 abaixo está um resumo das características das funções para o modelos AR(p), MA(q) e ARMA(p,q):

Quadro 1 – Resumo das características teóricas da FAC e da FACP dos modelos AR(p), MA(q) e ARMA (p,q)

Modelo	Função de Autocorrelação (ρ_k)	Função de Autocorrelação Parcial (ϕ_{kk})
AR(p)	Infinita (Exponencial e/ou senóides amortecidas)	Finita (Corte após o lag “p”)
MA (q)	Finita (Corte após o lag “q”)	Infinita (Exponencial e/ou senóides amortecidas)
ARMA (p,q)	Infinita (Exponencial e/ou senóides amortecidas após o lag “q-p”)	Infinita (Exponencial e/ou senóides amortecidas após o lag “p-q”)

Fonte: Zanini, 2023.

Logo após a identificação da ordem do modelo, é preciso estimar os parâmetros através da técnica da máxima verossimilhança. Para finalizar, utiliza-se os testes de aderência para verificar a adequação final do modelo. Nestes testes, encontram-se testes para resíduos, os testes de sobrefixação, entre outros.

Caso as séries temporais apresentem sazonalidades, tem-se o modelo SARIMA(p,d,q)x(P,D,Q). Este procedimento segue a mesma ideia para a obtenção do modelo ARIMA não sazonal.

Nesse caso a modelagem segue a seguinte equação:

$$\varphi(B)\Phi(B^s)\nabla^d S D \nabla^D Z_t = \theta(B)\Theta(B^s) a_t \quad (4.3.4)$$

Onde:

$\phi(B)$: operador não sazonal auto-regressivo

ϕ_i : parâmetros auto-regressivo não-sazonais

$\nabla^d = (1-B)^d$: operador diferença não sazonal de ordem d

$\Phi(B^s)$: operador sazonal auto-regressivo

Φ_i : parâmetros auto-regressivo sazonais

$\nabla^{D_s} = (1-B^s)^D$: operador diferença sazonal de ordem D

$\theta(B)$: operador não sazonal de médias móveis

θ_i : parâmetros de médias móveis não sazonais

- $\Theta(B^s)$: operador sazonal de médias móveis
 Θ_i : parâmetros de médias móveis sazonais

4.4 Competição de Métodos

Como dito anteriormente, este projeto realizou uma competição entre os dois métodos utilizados, a fim de comparar qual é o melhor. Para isso, foram utilizadas as análises dentro da amostra (*in-sample*) e fora da amostra (*out-of-sample*). Esta comparação tem como objetivo verificar os valores reais e os valores "ajustados". Fornece o nível de erro de previsão gerado pelos dois modelos. Este nível de erro é calculado um período à frente, indicando um desempenho preditivo para o horizonte futuro. Neste procedimento, parte dos dados são selecionados (por exemplo, os últimos 12 meses) para validar o poder de previsão dos modelos ajustados com os dados restantes.

O procedimento denominado análise recursiva fora da amostra (*out-of-sample rolling evaluation*) também será analisado. Neste estudo, são feitas previsões para cada origem. Caso sejam retirados os 12 últimos dados¹⁶, então serão feitas no período *out-of-sample* um total de 78 previsões (12 previsões para um passo (mês) à frente, 11 previsões para 2 passos à frente, 10 previsões para 3 passos à frente e assim sucessivamente). Pode-se definir como critério, por exemplo, que o método vencedor a ser selecionado será aquele que minimizar o MAD (*Mean Absolute Deviation*) acumulado ou erro médio absoluto acumulado fora da amostra.

Para entender melhor o procedimento de *rolling evaluation* observar a figura 4 a seguir.

¹⁶ Obviamente retirar os últimos dados é apenas um procedimento de análise para seleção de modelos. Para se obter as previsões desejadas, depois de feita análise *in-sample* e *out-of-sample*, os dados retirados são novamente incorporados e os parâmetros atualizados. Foi dado como exemplo a retirada dos últimos 12 dados o que geraria um total de 78 previsões fora da amostra. Entretanto, este "período de corte" pode ser definido de acordo com cada problema. Como exemplo, se são retirados os últimos 6 dados, geram-se 21 previsões fora da amostra.

Figura 4 – Procedimento de *out-of-sample rolling evaluation*

		HORIZONTE											
		jan	fev	mar	abr	mai	jun	jul	ago	set	out	nov	dez
ORIGEM	dez	A	A	A	A	A	A	A	A	A	A	A	A
	Jan	F	F	F	F	F	F	F	F	F	F	F	F
	Fev	A	A	F	F	F	F	F	F	F	F	F	F
	mar	A	A	A	F	F	F	F	F	F	F	F	F
	abr	A	A	A	A	F	F	F	F	F	F	F	F
	mai	A	A	A	A	A	F	F	F	F	F	F	F
	Jun	A	A	A	A	A	A	F	F	F	F	F	F
	Jul	A	A	A	A	A	A	A	F	F	F	F	F
	ago	A	A	A	A	A	A	A	A	F	F	F	F
	set	A	A	A	A	A	A	A	A	A	F	F	F
	out	A	A	A	A	A	A	A	A	A	A	F	F
	nov	A	A	A	A	A	A	A	A	A	A	A	F

Fonte: Zanini (2023)

Nota: F = Forecast; A = Actual.

A partir das previsões *out-of-sample*, algumas métricas de desempenho podem ser obtidas como, por exemplo, o MAPE (*Mean Absolute Percentual Error*) e o MAD (*Mean Absolute Deviatio*) para cada horizonte de previsão de forma acumulada. Pode-se calcular também o GMRAE (*Geometric Mean Relative Absolute Error*) que compara o erro do modelo escolhido com o erro do modelo ingênuo (que utiliza o último dado disponível na previsão). O GMRAE é a média geométrica da razão (quociente) entre o erro absoluto (portanto, em módulo) do modelo estimado e o erro absoluto do método ingênuo. Portanto, é desejável que o GMRAE seja igual ou menor do que 1.

É importante ressaltar que, assim que o melhor método for selecionado de acordo com algum critério previamente estabelecido, os dados retirados da amostra são novamente inseridos, atualizando os parâmetros e realizando as projeções. Em

síntese, a análise *in-sample* e *out-of-sample* é apenas um procedimento que visa definir o melhor dentre dois ou mais métodos de previsão.

4.4.1 Desempenho Preditivo

Várias métricas podem ser calculadas para medir o desempenho a partir da comparação dos valores reais e dos valores "ajustados" pelo modelo. A seguir estão algumas delas que serão utilizadas no trabalho:

1) MAPE (*Mean Absolute Percentual Error*)

O MAPE (erro médio absoluto percentual) é calculado através da diferença entre valores estimados e reais e equivale às previsões um passo-à-frente (por exemplo, para o mês seguinte). Veja a equação (4.4.1) seguir:

$$MAPE = \frac{\sum_{t=1}^N \frac{|Y(t) - \hat{Y}(t)|}{Y(t)}}{N} \times 100 \quad (4.4.1)$$

Onde:

$Y(t)$ = valor da série temporal no período (t) ;

$\hat{Y}(t)$ = valor ajustado da série temporal para o período (t);

N = total de dados utilizados (total de observações).

2) Coeficiente de Explicação (R^2)

O coeficiente de explicação (R^2) indica o quanto da variação total dos dados (série dependente) é explicada pelo modelo. Este coeficiente é calculado através da comparação do erro do modelo e a variação dos dados da série dependente (série a ser prevista) em torno de sua média. Importante salientar que o coeficiente de explicação (ou determinação) varia entre 0 e 100%. Veja a equação (4.4.2) a seguir:

$$R^2 = \left(1 - \frac{\sum_{t=1}^N (Y(t) - \hat{Y}(t))^2}{\sum_{t=1}^N (Y(t) - \bar{Y})^2} \right) \times 100 \quad (4.4.2)$$

Onde:

$Y(t)$ = valor da série temporal no período (t);

$\hat{Y}(t)$ = valor ajustado da série temporal para o período (t);

\bar{Y} = média das observações (média da série temporal) ;

N = total de dados utilizados (total de observações).

3) Coeficiente de Explicação Ajustado (R^2 ajustado)

O coeficiente de explicação ajustado (R^2 ajustado) tem interpretação similar ao coeficiente de explicação, ou seja, indica o quanto da variação total dos dados é explicada pelo modelo. Entretanto, faz-se uma correção tendo em vista a quantidade de parâmetros no modelo. Veja a equação (4.4.3) a seguir:

$$R^2_{ajust} = \left(1 - \frac{\sum_{t=1}^N \frac{(Y(t) - \hat{Y}(t))^2}{N - k}}{\sum_{t=1}^N \frac{(Y(t) - \bar{Y})^2}{N - 1}} \right) \times 100 \quad (4.4.3)$$

Onde:

$Y(t)$ = valor da série temporal no período (t);

$\hat{Y}(t)$ = previsão da série temporal para o período (t);

\bar{Y} = média das observações (média da série temporal) ;

N = total de dados utilizados (total de observações);

k = número de parâmetros do modelo

4) GMRAE

Pode-se calcular também o GMRAE (*Geometric Mean Relative Absolute Error*) que compara o erro do modelo selecionado com o erro do modelo ingênuo (que usa como previsão o último dado disponível). O GMRAE é a média geométrica da razão (quociente) entre o erro absoluto (portanto, em módulo) do modelo estimado e o erro absoluto do método ingênuo. Portanto, é desejável que o GMRAE seja igual ou menor do que 1.

$$GMRAE = \sqrt[N]{\prod_{t=1}^N \left(\frac{|Y(t) - \hat{Y}(t)|}{|Y(t) - Y(t-1)|} \right)} \quad (4.4.4)$$

Onde $Y(t)$ é o valor da série temporal no período (t);

$\hat{Y}(t)$ é a previsão da série temporal para o período (t);

$Y(t-1)$ é o valor da série temporal no período (t-1);

N é o número de observações.

Apresentada a metodologia de análise, no próximo capítulo pode-se ver a análise de dados e resultados deste trabalho monográfico.

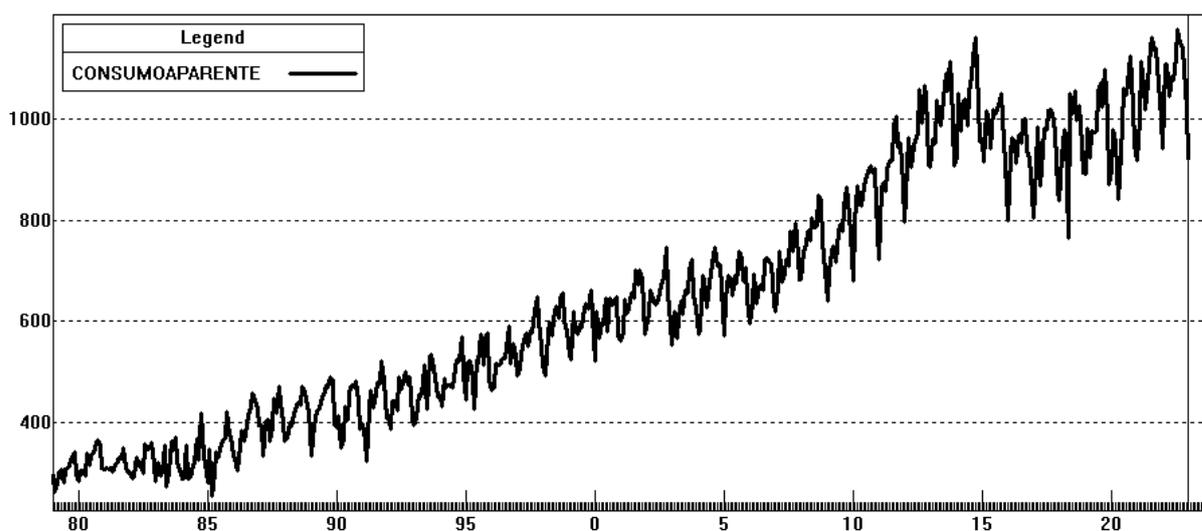
5 ANÁLISE DE DADOS E RESULTADOS

5.1 Base de Dados

Os dados utilizados para a previsão do modelo de consumo de óleo diesel no Brasil foram coletados do Instituto de Pesquisa Econômica Aplicada (IPEA)¹⁷. As estimativas foram feitas através do *software Forecast Pro for Windows* (FPW).

Os dados são mensais da média diária do consumo de óleo diesel, medidos em milhares de barris entre os períodos de janeiro de 1979 e janeiro de 2023. Essas informações representam o consumo aparente correspondente à produção industrial doméstica acrescida das importações e diminuída as exportações. Na figura 5 abaixo pode-se ver o gráfico com a evolução do consumo de diesel no período em análise:

Figura 5 – Consumo de óleo diesel no Brasil (em milhares de barris)



Fonte: Elaboração própria.

Na próxima seção serão apresentados os resultados obtidos dos métodos e a competição entre eles, determinando qual o melhor. O período para a previsão do consumo do óleo diesel foi entre fevereiro de 2023 a dezembro de 2024.

¹⁷ Os dados primários foram publicados através da Agência Nacional do Petróleo, Gás Natural e Biocombustíveis (ANP) e disponibilizados no Instituto de Pesquisa Econômica Aplicada (IPEA).

5.2 Equações de Previsão

Utilizando o método MAE, o modelo que melhor se explica a evolução do consumo de óleo diesel é o Modelo de Holt-Winters apresentado no capítulo 4 (equação 4.2.9). Na tabela 2 abaixo estão os parâmetros estimados e os hiperparâmetros. Na tabela 3 estão os valores dos fatores sazonais estimados:

Tabela 2 – Parâmetro do Método de Amortecimento Exponencial

Componente	Parâmetro	Hiperparâmetro
Nível	1077,9	0,23457
Tendência	1,4525	0,00313
Sazonalidade	*	0,17675

Fonte: Elaboração própria.

* Nota: Fatores sazonais estão na tabela 3.

Tabela 3 – Fatores Sazonais

Meses	Fator	Meses	Fator
Janeiro	0,88447	Julho	1,04252
Fevereiro	0,98253	Agosto	1,07079
Março	1,00817	Setembro	1,06057
Abril	0,97349	Outubro	1,07180
Maiο	0,96160	Novembro	1,01711
Junho	1,01677	Dezembro	0,92915

Fonte: Elaboração própria.

De acordo com os valores encontrados acima, a estrutura do modelo de Holt-Winters apresenta a seguinte equação:

$$Z_t = (1.077,9 + 1,4525 * \tau) * \rho_t \quad (5.2.1)$$

Z_t representa a quantidade de óleo diesel consumido no período τ (horizonte de previsão). O valor de 1.077,9 representa o parâmetro de nível e o valor de 1,4525 é o parâmetro de tendência. Além disso, ρ_t é o parâmetro de sazonalidade, cujos valores (um para cada mês) são apresentados na tabela 3. Logo, os parâmetros estimados indicam um nível atual de 1.077,9 em milhares de barris de óleo diesel

consumidos e um acréscimo logo para o próximo período de 1,4525 em milhares de barris do consumo neste nível. O maior fator sazonal estimado foi 1,07180 para o mês de outubro, o que significa que este mês incrementa em cerca de 7% o consumo. Já para dezembro estimou-se um fator de 0,92915, o que significa uma queda de cerca de 7% no consumo.

Pode-se perceber no Modelo de Holt-Winters que a estimativa do hiperparâmetro de nível significa que se atribui um peso menor para o presente (0,23457) do que para o passado (0,76543) na estimação do parâmetro de nível. O mesmo ocorreu com os hiperparâmetros de tendência e de sazonalidade.

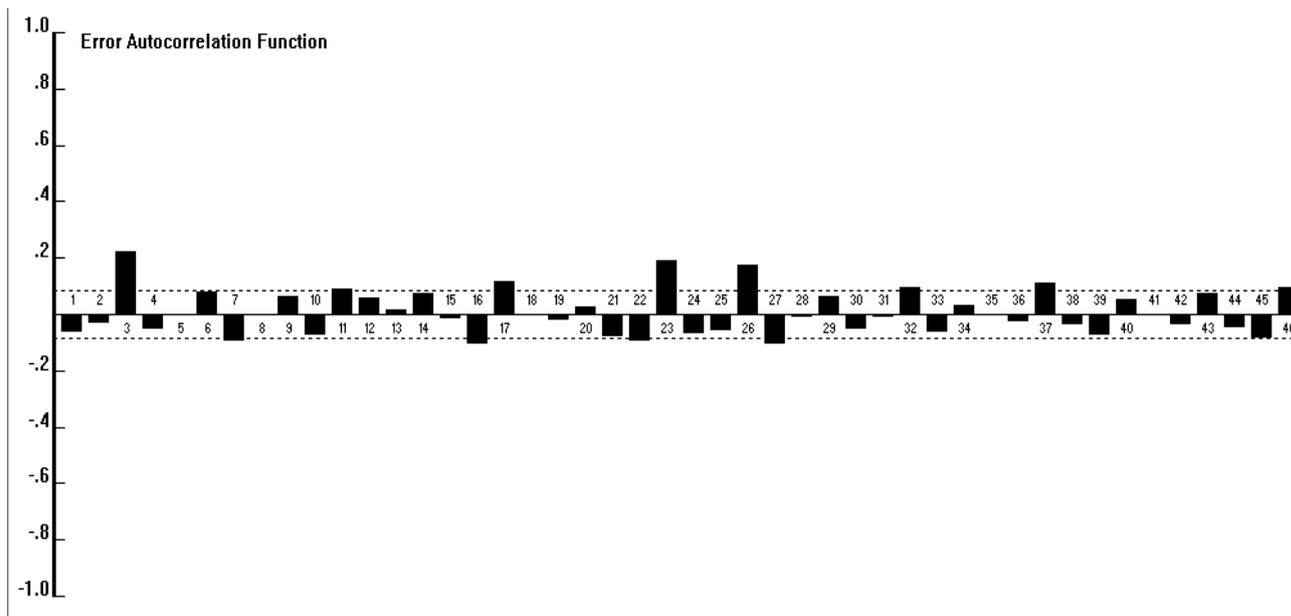
Para a verificação da fidedignidade do modelo¹⁸, foi realizado o teste de Ljung-Box e o gráfico ACF Erro. No teste de Ljung-Box foi testada a hipótese de que as "k" primeiras autocorrelações eram nulas, isto é:

$$H_0: \rho_1 = \rho_2 = \dots = \rho_k = 0$$

O p-valor encontrado foi igual a 1. No FPW, caso o p-valor seja menor que o nível de confiança (0,95), aceita-se a hipótese nula de erros descorrelatados (H_0). Apesar de se rejeitar a hipótese nula no teste de Ljung-Box, observando o gráfico da autocorrelação dos erros (figura 6), pode-se considerar que os resíduos são na verdade descorrelatados, com as correlações variando entre -0,1 e 0,2.

¹⁸ Ver Zanini (2023).

Figura 6 – Função de Autocorrelação dos Erros – Método de Amortecimento Exponencial



Fonte: Elaboração própria

Para a análise do desempenho preditivo do modelo, observou-se o R^2 ajustado, que representa o poder de explicação da variação do consumo do diesel, e o valor do MAPE, que indica o erro médio absoluto percentual equivalente às previsões um passo à frente. Neste trabalho, o modelo de Holt-Winters estimado explica 98,4% do consumo no período e erra, em média, 4,1% para cima ou para baixo para projetar o próximo período (mês).

Utilizando o método de Box & Jenkins (MBJ), o modelo estimado para o consumo do diesel foi um $ARIMA(1,1,1)*(1,0,2)$ com transformação de raiz quadrada. Na tabela 4 abaixo estão os valores calculados para os parâmetros. Observou-se que todos os parâmetros estimados são significantes ao nível de significância de 5%.

Tabela 4 – Parâmetros do Modelo de Box & Jenkins

Termo	Coefficiente	P-valor
ϕ	-0,1144	0,9582
θ	0,7246	1,0000
Φ	0,9998	1,0000
Θ_1	0,7887	1,0000
Θ_2	0,0844	0,9466

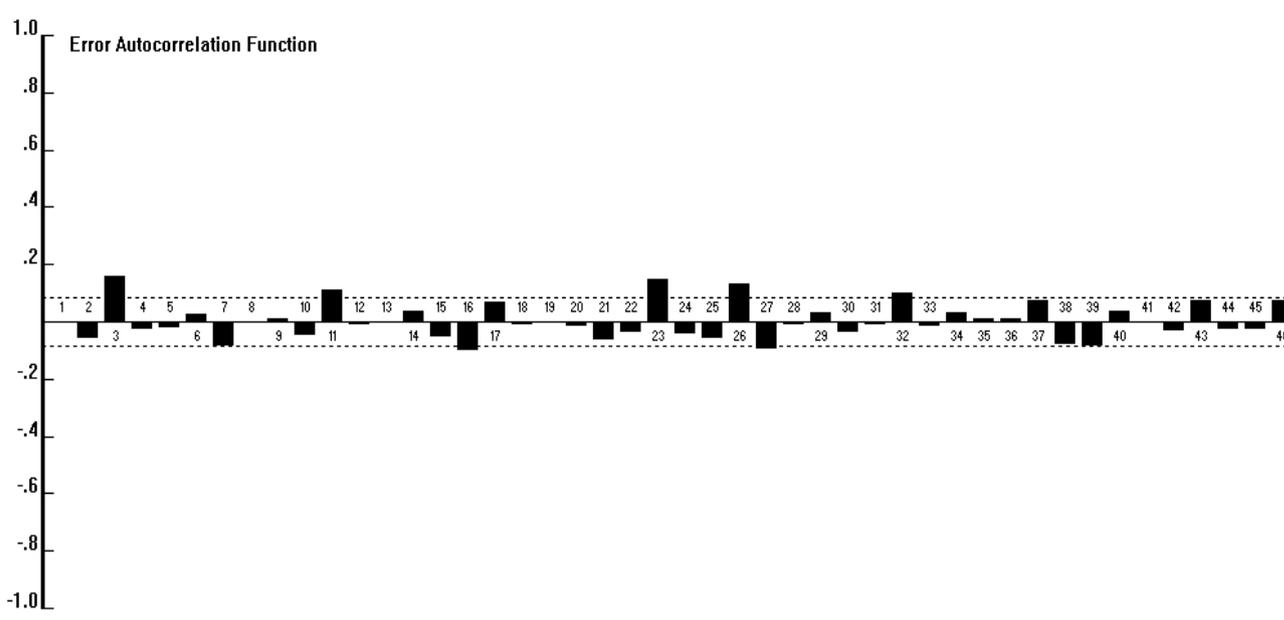
Fonte: Elaboração própria.

A equação 5.2.2 abaixo apresenta o modelo com os valores encontrados em cada um dos parâmetros:

$$(1 - 0,9998B^{12})(1 + 0,1144B)(1 - B) \sqrt{Z_t} = (1 - 0,7887B^{12} - 0,0844B^{24}) (1 - 0,7246B)a_t \quad (5.2.2)$$

Quanto ao teste de Ljung-Box, verificou-se que o p-valor encontrado foi igual a 1, neste caso, a hipótese nula é rejeitada. Porém, observando o gráfico da autocorrelação dos erros (figura 7), os resíduos são descorrelatados variando também entre -0,1 e 0,2.

Figura 7 – Função de Autocorrelação dos Erros – Método de Box & Jenkins



Fonte: Elaboração própria

O R^2 ajustado apresentou uma capacidade de explicação de 98,45% do consumo do diesel, e o MAPE indicou que, ao projetar o próximo período, o modelo errou em média 4,052% para cima ou para baixo.

5.3 Competição de Métodos

Para escolher o melhor método para gerar as previsões de consumo de óleo diesel, foi realizada uma análise recursiva fora da amostra, conhecido como *out-of-sample rolling evaluation* e uma análise dentro da amostra (*in-sample*). Foram retiradas da amostra os últimos 12 dados. A tabela 5 a seguir apresenta a comparação dos modelos.

Tabela 5 – Competição entre os Métodos

Dentro da Amostra (<i>In sample</i>)			Fora da Amostra (<i>Out of Sample</i>)			
Modelo	R ² adj(%)	MAPE (%)	MAPE (%) H=1 N=12	MAPE ACUM (%)	GMRAE H=1 N=12	GMRAE ACUM
MAE	98,32	4,11	2,70	2,60	0,75	0,31
MBJ	98,37	4,10	2,50	2,70	0,68	0,31

Fonte: Elaboração própria

Primeiramente, analisando cada método separadamente (análise na linha da tabela 05), vê-se que ambos têm bom poder de generalização, ou seja, prevêm melhor a realidade mais recente. Vê-se que ambos os métodos possuem um MAPE fora da amostra menor do que o MAPE dentro da amostra. No caso do MAE, 2,70% fora da amostra contra 4,11% dentro da amostra. Já no MBJ, tem-se 2,50% fora da amostra contra 4,10% dentro da amostra. Ambos inclusive apresentam o MAPE acumulado nas 78 previsões abaixo do MAPE dentro da amostra.

Num outro momento, comparando-se os dois modelos, analisando os dados dentro da amostra, vê-se que o método de Amortecimento Exponencial explicou 98,32% da variação do consumo do óleo diesel e obteve um erro médio para o próximo período de 4,11%. Já o método de Box & Jenkins obteve um desempenho melhor de 98,37% do poder de explicação e um erro médio de 4,10%. Vê-se, portanto, que o MBJ tem um desempenho levemente superior ao MAE dentro da amostra.

Examinando os resultados fora da amostra, observou-se que o MAPE na metodologia Box & Jenkins foi menor do que no Amortecimento Exponencial, 2,50% contra 2,70%, ou seja, o erro foi menor no MBJ. Porém, o MAPE acumulado do erro no Amortecimento Exponencial foi menor do que no Box & Jenkins. Já no GMRAE, percebe-se que ambos os modelos ganham do método ingênuo, com GMARAE para as previsões um passo à frente e acumulado menores que 1.

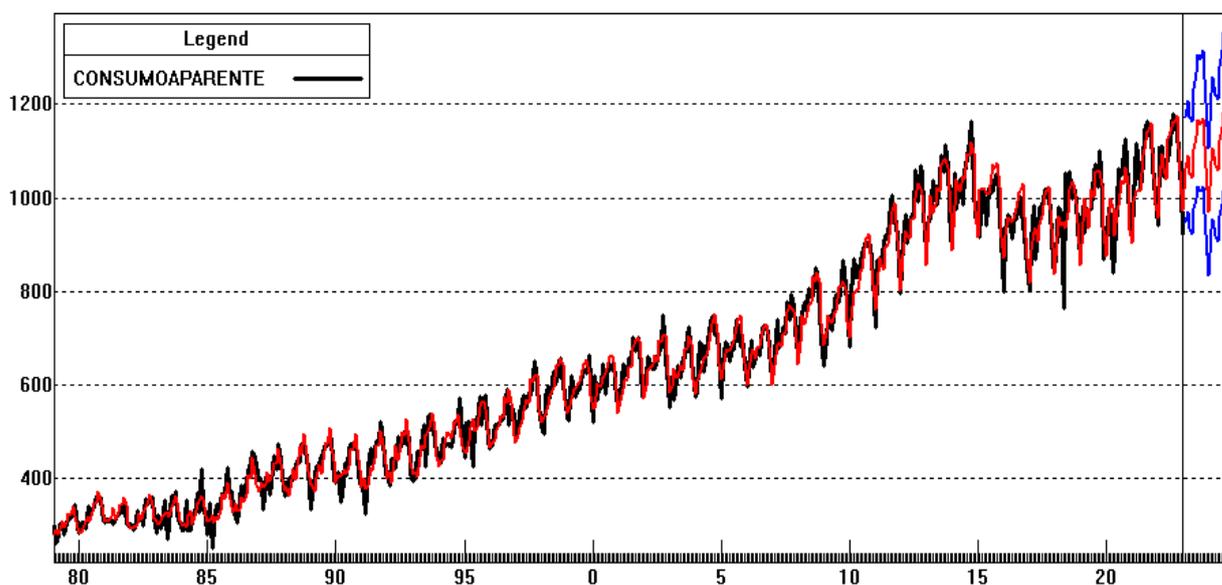
Em síntese, analisando as estatísticas dentro e fora da amostra, vê-se que ambos os modelos possuem bom poder de generalização (erram menos fora da amostra do que dentro da amostra), sendo que o MBJ possui desempenho melhor que o MAE tanto dentro quanto fora da amostra. Estes resultados indicam que o MBJ pode então ser considerado o método vencedor neste caso e será utilizado para previsão do consumo de óleo diesel no Brasil.

Como foi apresentado no capítulo 4 (Metodologia), realizada a análise recursiva fora da amostra e definido o método vencedor, os dados anteriormente retirados da amostra são novamente inseridos nesta e o modelo reestimado estando pronto para calcular as previsões no horizonte de previsão. Desta forma, o modelo que será utilizado para calcular as previsões do consumo de óleo diesel no Brasil é o estimado pela MBJ e apresentado na equação 5.2.2.

5.3 Previsão do Consumo de Óleo Diesel

A figura 8 abaixo exhibe a evolução do consumo de óleo diesel no Brasil, com os valores projetados pelo método vencedor (MBJ), além do intervalo de confiança de 95%. Importante atentar que uma das vantagens de se trabalhar com modelos estatísticos e matemáticos é que não se gera apenas um número, a previsão, mas uma previsão na forma probabilística, ou seja, calcula-se um intervalo de confiança que possui determinada probabilidade de conter o valor do consumo de óleo diesel.

Figura 8 – Consumo de óleo diesel (em milhares de barris) – período de 1979 a 2024. Valores reais, ajustado e previsões com intervalo de confiança de 95%



Fonte: Elaboração própria. Nota: ____ Valores reais ____ Valores projetados ____ Intervalo de Confiança de 95%

Na tabela 6 a seguir, estão as previsões para o período de fevereiro de 2023 a dezembro de 2024. Pode-se observar ainda os valores do intervalo de confiança de 95%.

Tabela 6 – Valores da Previsão do Consumo de Óleo Diesel (em milhares de barris) – Intervalo de Confiança de 95%

Mês	Limite Inferior	Previsão	Limite Superior
2023-02	948	1.061	1.173
2023-03	971	1.090	1.209
2023-04	935	1.054	1.172
2023-05	922	1.042	1.162
2023-06	975	1.103	1.232
2023-07	999	1.133	1.267
2023-08	1.025	1.165	1.305
2023-09	1.014	1.156	1.297

Fonte: Elaboração própria

Tabela 7 – Valores da Previsão do Consumo de Óleo Diesel (em milhares de barris)
– Intervalo de Confiança de 95% - Continuação

2023-10	1.024	1.169	1.315
2023-11	968	1.111	1.255
2023-12	878	1.016	1.154
2024-01	832	969	1.105
2024-02	931	1.078	1.225
2024-03	956	1.107	1.259
2024-04	920	1.071	1.222
2024-05	907	1.059	1.211
2024-06	962	1.121	1.280
2024-07	987	1.151	1.315
2024-08	1.015	1.184	1.352
2024-09	1.004	1.174	1.344
2024-10	1.015	1.188	1.361
2024-11	957	1.129	1.300
2024-12	866	1.033	1.200

Fonte: Elaboração própria

Na tabela 06, pode-se observar que foram feitas 23 previsões. Calculando as variações percentuais neste período, encontra-se um valor médio de 0,6%, ou seja, está sendo projetado um aumento médio mensal no consumo aparente. Analisando o período anterior equivalente, ou seja, os 23 dados anteriores (março de 2021 a janeiro de 2023), encontra-se uma variação do consumo médio mensal de -0,1%. Percebe-se, portanto, um crescimento no consumo do diesel, o que vai ao encontro das projeções da Empresa de Pesquisa Energética (EPE) (2023), que afirma que a demanda total de combustíveis deve seguir em alta em 2023 e 2024, sendo o consumo do diesel o mais expressivo. Este aumento se dará em grande parte pelo crescimento do agronegócio e da mobilidade.

6 CONCLUSÃO

O objetivo principal do trabalho foi estimar um modelo de consumo de óleo diesel no Brasil. Este objetivo se justifica, como evidenciado no trabalho monográfico, pela importância deste derivado do petróleo para a economia. Este estudo pode gerar informações que subsidiem as decisões dos agentes econômicos em relação à demanda e a oferta de consumo futuro, planejando assim melhor suas ações.

Foi apresentado como o mercado de combustível, mais especificamente, o mercado do diesel se comporta e a sua importância para os diversos setores e como isso impacta na economia do país. Algumas das áreas citadas no projeto foram: industrial, agrícola, marítima, ferroviária e para geração de energia. Portanto, é um combustível muito relevante, seus bens são gerados a partir de insumos de difícil substituição na matriz produtiva. Ademais, neste estudo também foi abordado como os derivados do petróleo são regularizados e fiscalizados.

As revisões de literatura foram expostas com a finalidade de mostrar algumas das diferentes metodologias utilizadas para realizar previsões nesta área de combustíveis.

Para atingir o objetivo proposto por este trabalho monográfico, foram estudadas e aplicadas duas diferentes técnicas explicadas: Método de Amortecimento Exponencial (MAE) e o Método de Box & Jenkins (MBJ), sendo ambos métodos univariados ou autoprojetivos.

Através da competição de métodos, utilizando a análise do desempenho preditivo *in-sample* e *out-of-sample rolling evaluation*, foi determinado que o modelo vencedor foi o de Box & Jenkins, apesar das duas metodologias apresentarem um desempenho muito parecido. Para o cálculo das previsões no horizonte até final de 2024, foi utilizado então um ARIMA(1,1,1)(1,0,2) com transformação de raiz quadrada.

Além disso, utilizando os valores estimados, foi possível analisar a variação percentual média do consumo no período projetado. Estima-se um crescimento médio mensal de 0,6% no consumo de óleo diesel no Brasil no período projetado, ou seja, até o final de 2024.

Foi possível evidenciar também que uma das vantagens de se trabalhar com métodos estatísticos de análise de séries de tempo, é que não se gera apenas a previsão, um número, mas sim um intervalo de confiança que possui determinada probabilidade (chance) de conter o valor da previsão. Em sendo assim, poder-se-ia,

por exemplo, elaborar cenários utilizando-se o limite inferior como um “cenário pessimista”, a previsão como um “cenário esperado” e o limite superior como um “cenário otimista”.

Como propostas futuras de trabalhos, poder-se-ia utilizar métodos que incorporem outras variáveis explicativas (como preço, exportação, importação, produção, vendas, entre outras) e que possam ter sua relação com o consumo de óleo diesel estimada, como, por exemplo, nos Modelos de Regressão Dinâmica. Em síntese, além das técnicas aplicadas nesta monografia, outros métodos também podem ser testados.

REFERÊNCIAS

ANP. Agência Nacional do Petróleo, Gás Natural e Biocombustíveis, **Anuário Estatístico Brasileiro do Petróleo, Gás Natural e Biocombustíveis 2022**. Disponível em: < <https://www.gov.br/anp>>. Rio de Janeiro: ANP, 2022.

Bastos, R. F. Avaliação de desempenho de modelos de séries temporais para previsões da produção de óleo diesel nacional. 2014. Disponível em: <<https://www.proceedings.blucher.com.br/article-details/avaliacao-de-desempenho-de-modelos-de-sries-temporais-para-previses-da-produo-de-leo-diesel-nacional-22711>>. Acesso em: 27 out.2022.

BOX, G. E. P., JENKINS, G. M.. Time Series Analysis, Forecasting and Control, San Francisco, Holden-Day, 1994.

BRAFMAN, Ilana. **Modelo econométrico para a projeção do consumo aparente de combustível no Brasil - Otto e Diesel**. Rio de Janeiro: IBMEC, 2009.

BRASIL, Câmara dos Deputados. **Projeto de Lei nº 9.478, de 6 de agosto de 1997**. Brasília: Câmara dos Deputados, 1997. Disponível em: <<https://www2.camara.leg.br/legin/fed/lei/1997/lei-9478-6-agosto-1997-365401-publicacaooriginal-1-pl.html>>. Acesso em: 16 out. 2022.

BRUNETTI, Franco. Motores de combustão interna – Vol 1. São Paulo: Edgar Blücher Ltda, 2012.

CAETANO, Lucas Pereira. **Mapeamento e Projeção do Consumo de Óleo Diesel no Brasil**. Orientadora: Rosemarie Bröker Bone. 2013. 77f. TCC (Graduação) - Curso de Engenharia de Produção da Escola Politécnica, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2013. Disponível em: <https://pantheon.ufrj.br/bitstream/11422/11892/1/monopoli10009174.pdf>. Acesso em: 27 abr. 2023.

CANELAS, ANDRÉ LUÍS DE SOUZA. Evolução da Importância Econômica da Indústria de Petróleo e Gás Natural no Brasil: Contribuição a Variáveis Macroeconômicas. Tese (Doutorado em Engenharia) - Universidade Federal do Rio de Janeiro. Rio de Janeiro, 2007.

CEPEA - Centro de Estudos Avançados em Economia Aplicada. PIB do Agronegócio Brasileiro, jan/jun 2022. Disponível em: <<https://www.cepea.esalq.usp.br/br/pib-do-agronegocio-brasileiro.aspx>>. Acesso em: 27 out.2022.

DELÁZARO, Walter; RODRIGUES, Marco Julio L; COSTILHES, Bertrand. A necessidade de substituição e racionalização do consumo de óleo diesel no Brasil. **Revista de Administração de Empresas**, v. 24, n. 1, p. 45-50, jun.2013. Disponível em: <<https://doi.org/10.1590/S0034-75901984000100006>>. Acesso em: 27 out.2022.

Diesel para Agronegócio. Centenário Diesel, g. Disponível em <<http://centenariodiesel.com.br/diesel-para-agronegocio/>>. Acesso em: 27 out.2022.

EPE [Empresa de Pesquisa Energética]. Perspectiva para o Mercado Brasileiro de Combustíveis no Curto Prazo. Abril de 2023. Disponível em <<https://www.epe.gov.br/sites-pt/publicacoes-dados-abertos/publicacoes/Paginas/Perspectivas-para-o-mercado-brasileiro-de-combustiveis-no-curto-prazo.aspx>>. Acesso em: 7 abr.2023

GOODRICH, R.L.. Applied Statistical Forecasting, Belmont, Business Forecast Systems, 1989.

IBGE, Volume de serviços cresce 3,1% em dezembro e fecha 2022 com alta de 8,3%. **Instituto Brasileiro de Geografia e Estatística**. [S. l.], 2023. Disponível em: <<https://agenciadenoticias.ibge.gov.br/agencia-sala-de-imprensa/2013-agencia-de-noticias/releases/36236-volume-de-servicos-cresce-3-1-em-dezembro-e-fecha-2022-com-alta-de-8-3>>. Acesso em: 31 jun. 2023.

INSTITUTO DE PESQUISA ECONÔMICA APLICADA – **Ipeadata**. Dados macroeconômicos e regionais. Disponível em: <<http://www.ipeadata.gov.br>>. Acesso em 25 out. 2022.

MELO, Isabela Esterminio. As Crises do Petróleo e seus Impactos sobre a Inflação do Brasil. 2008. 35f. Trabalho de Conclusão de Curso - Pontifícia Universidade Católica do Rio de Janeiro, [S. l.], 2008.

MONTGOMERY, D.C., JOHNSON, L.A.. **Forecasting and Time Series Analysis**, New York, McGraw-Hill Book Co., 1990.

MOREIRA, Ajax R.B. Modelos para a Projeção do Consumo Nacional e Regional de Óleo Diesel. Rio de Janeiro: Ipea, 1996. (Texto para Discussão, n. 443).

RODRIGUES, N. ; LOSEKANN, L. D. . Análise da demanda por óleo diesel no Brasil. In: 5 Encuentro Latinoamericano de Economía de la Energía, 2015, Medellín. 5 Encuentro Latinoamericano de Economía de la Energía, 2015.

SILVA, Juliana Rodrigues de Melo. Implicações Econômicas do Uso do Óleo Diesel no Setor de Transporte Rodoviário Brasileiro. 2010. 76 f. Trabalho de Conclusão de Curso - Universidade Federal do Rio de Janeiro, [S. l.], 2010.

SOUZA, R. C., CAMARGO, M. E., *Análise e Previsão de Séries Temporais: Os Modelos ARIMA*. SEDIGRAF, 1996.

SOUZA, Y. M. de; RODRIGUES, L.; FURLAN AMARAL, D. ANÁLISE ESPACIAL DA DEMANDA POR DIESEL NOS ESTADOS BRASILEIROS. **Revista de Economia e Agronegócio**, [S. l.], v. 19, n. 1, p. 1–26, 2021. DOI: 10.25070/rea.v19i1.11671. Disponível em: <https://periodicos.ufv.br/rea/article/view/11671>. Acesso em: 16 out. 2022.

UCHÔA, Frederico et al. Fuel Demand Elasticities in Brazil: A Panel Data Analysis with Instrumental Variables. *International Journal of Energy Economics and Policy*, v. 10, n. 2, p. 450-457, 2020.

ZANINI, A. Modelos de Previsão para Séries Temporais. Material Didático. Universidade Federal de Juiz de Fora, 2023.