



DISSERTAÇÃO DE MESTRADO PROFISSIONAL

**Vulnerabilidade e Eficiência no Setor Energético Brasileiro:
Uma Análise Estocástica com Modelos DEA e Forecast**

GABRIELA MAYUMI SAIKI

Orientador Prof. Dr. André Luiz Marques Serrano

Programa de Pós-Graduação Profissional em Engenharia Elétrica

DEPARTAMENTO DE ENGENHARIA ELÉTRICA

FACULDADE DE TECNOLOGIA

UNIVERSIDADE DE BRASÍLIA

UNIVERSIDADE DE BRASÍLIA
FACULDADE DE TECNOLOGIA
DEPARTAMENTO DE ENGENHARIA ELÉTRICA

**Vulnerabilidade e Eficiência no Setor Energético Brasileiro: Uma Análise Estocástica com Modelos
DEA e Forecast**

**Vulnerabilidade e Eficiência no Setor Energético Brasileiro: Uma Análise Estocástica com Modelos
DEA e Forecast**

Gabriela Mayumi Saiki

Orientador: Prof. Dr. André Luiz Marques Serrano, PPEE/UnB

PUBLICAÇÃO: PPEE.MP.088

BRASÍLIA-DF

UNIVERSIDADE DE BRASÍLIA
Faculdade de Tecnologia

DISSERTAÇÃO DE MESTRADO PROFISSIONAL

**Vulnerabilidade e Eficiência no Setor Energético Brasileiro:
Uma Análise Estocástica com Modelos DEA e Forecast**

GABRIELA MAYUMI SAIKI

Orientador Prof. Dr. André Luiz Marques Serrano

*Dissertação de Mestrado Profissional submetida ao Departamento de Engenharia
Elétrica como requisito parcial para obtenção
do grau de Mestre em Engenharia Elétrica*

Banca Examinadora

Prof. Dr. Andre Luiz Marques Serrano, PPEE/UnB _____
Presidente

Prof. Dr. Vinicius Pereira Gonçalves, PPEE FT/UnB _____
Examinador Interno

Dr. Fabiano Mezadre Pompermayer, Instituto de _____
Pesquisa Econômica Aplicada (IPEA)
Examinador externo

Prof. Dr. Geraldo Pereira Rocha Filho, PPEE _____
FT/UnB
Suplente

FICHA CATALOGRÁFICA

SAIKI, GABRIELA MAYUMI

Vulnerabilidade e Eficiência no Setor Energético Brasileiro: Uma Análise Estocástica com Modelos DEA e Forecast [Distrito Federal] 2024.

xvi, 68 p., 210 x 297 mm (ENE/FT/UnB, Mestre, Engenharia Elétrica, 2024).

Dissertação de Mestrado Profissional - Universidade de Brasília, Faculdade de Tecnologia.

Departamento de Engenharia Elétrica

1. Texto

2. Texto

3. Texto

4. Texto

I. ENE/FT/UnB

II. Título (série)

REFERÊNCIA BIBLIOGRÁFICA

SAIKI, G.M. (2024). *Vulnerabilidade e Eficiência no Setor Energético Brasileiro: Uma Análise Estocástica com Modelos DEA e Forecast*. Dissertação de Mestrado Profissional, Departamento de Engenharia Elétrica, Universidade de Brasília, Brasília, DF, 68 p.

CESSÃO DE DIREITOS

AUTOR: GABRIELA MAYUMI SAIKI

TÍTULO: Vulnerabilidade e Eficiência no Setor Energético Brasileiro: Uma Análise Estocástica com Modelos DEA e Forecast .

GRAU: Mestre em Engenharia Elétrica ANO: 2024

É concedida à Universidade de Brasília permissão para reproduzir cópias desta Dissertação de Mestrado e para emprestar ou vender tais cópias somente para propósitos acadêmicos e científicos. Do mesmo modo, a Universidade de Brasília tem permissão para divulgar este documento em biblioteca virtual, em formato que permita o acesso via redes de comunicação e a reprodução de cópias, desde que protegida a integridade do conteúdo dessas cópias e proibido o acesso a partes isoladas desse conteúdo. O autor reserva outros direitos de publicação e nenhuma parte deste documento pode ser reproduzida sem a autorização por escrito do autor.

GABRIELA MAYUMI SAIKI

Depto. de Engenharia Elétrica (ENE) - FT

Universidade de Brasília (UnB)

Campus Darcy Ribeiro

CEP 70919-970 - Brasília - DF - Brasil

“Seja você quem for, seja qual for a posição social que você tenha na vida, a mais alta ou a mais baixa, tenha sempre como meta muita força, muita determinação e sempre faça tudo com muito amor e com muita fé em Deus, que um dia você chega lá. De alguma maneira você chega lá.”

Ayrton Senna

AGRADECIMENTOS

Em primeiro lugar, gostaria de agradecer a Deus por mais essa etapa concluída em minha vida. Sua graça, amor e cuidado permitiram que eu chegasse até aqui. Há dias em que somente Ele acalma minha mente e meu coração. Somente Ele consegue extrair o que há de melhor em mim. É a razão do meu viver, minha força nos dias difíceis, meu refúgio e porto seguro.

Em segundo lugar, expresso meus sinceros agradecimentos ao Prof. Dr. André Luiz Marques Serrano, a quem tenho como orientador, mentor, conselheiro e amigo. Desde 2022, ele acreditou em mim e me deu todas as oportunidades possíveis para me tornar uma profissional, estudante e, acima de tudo, uma pessoa melhor. No fundo, nos somos o resultado das pessoas que apostam e acreditam em nós. Não sei o que o futuro me reserva, mas certamente o Professor André será um exemplo de dedicação e companheirismo que levarei comigo ao longo da vida.

Não poderia deixar de agradecer aos meus pais, Pedro Kunio Saiki e Vandinéia Sofia de Aquino Saiki, que, apesar de todas as dificuldades da vida, sempre me deram o melhor que podiam, ensinaram-me sobre honestidade e trabalho duro, e incentivaram-me a buscar um futuro melhor e diferente da realidade que tiveram. Da mesma forma, sou grata aos meus primos e irmãos de consideração, Suzy Sumie Saiki e Douglas Satochi da Silveira Saiki, que, além de serem companheiros de vida, me deram carona todos os dias para que eu pudesse concluir os estudos.

Por fim, registro meus sinceros agradecimentos ao grupo Projetum, em especial aos pesquisadores Gabriel Arquelau Pimenta Rodrigues, Guilherme Dantas Bispo, Guilherme Fay Vergara, Letícia Rezende Mosquéra, Matheus Noschang de Oliveira, Patrícia Helena dos Santos Martins e Raquel Valadares Borges, bem como à gerente do projeto, Patrícia Ferreira Missel, que foram meus parceiros de projeto durante a execução desta pesquisa.

RESUMO

Para alcançar o Objectivo de Desenvolvimento Sustentável 7 (ODS7), é crucial ter algoritmos e modelos que possam prever com precisão e melhorar a eficiência dos processos. Estas ferramentas fornecem contribuições valiosas para as estratégias de planeamento dos tomadores de decisão. Neste estudo, foi utilizado análise envoltória de dados (DEA) e algoritmos de bootstrap para analisar a eficiência energética do Brasil de 2004 a 2023. Foi comparado o desempenho dos modelos de média móvel integrada autorregressiva (ARIMA) e de média móvel integrada autorregressiva sazonal (SARIMA) para previsão a tendência das variáveis nos próximos vinte e quatro meses. Além disso, é aplicado o modelo de heterocedasticidade condicional autorregressiva generalizada (GARCH) para avaliação das previsões diárias de preços do ELET3 correlacionadas com as tendências dos preços da energia industrial e do PIB. O estudo traz conclusões importantes sobre o cenário energético brasileiro. Conforme os resultados encontrados, a taxa anual de consumo de energia deverá aumentar em média 2,1% até 2030, o que é acompanhado por uma tendência de crescimento do PIB. Utilizando as tecnologias existentes no país, é possível reduzir os custos com consumo de energia elétrica em uma média de 30,58%, mantendo o mesmo valor do PIB. Além disso, foi encontrada uma correlação negativa significativa de 33,91% entre a variável ELET3 e o preço médio da energia na indústria, sugerindo que, à medida que a ELET3 diminui, o preço tende a aumentar. Ademais, observou-se um aumento na volatilidade dos preços das ações durante períodos marcados por problemas climáticos severos, como secas e estiagens, bem como flutuações causadas pela COVID-19. Os achados revelaram uma tendência de queda nos preços das ações juntamente com o aumento da instabilidade do mercado, indicando a eficácia do modelo GARCH para os agentes interessados em compreender o setor do mercado de energia.

Palavras-chave: Energia, PIB, Análise Envoltória de Dados, Bootstrap, Forecast

ABSTRACT

To achieve Sustainable Development Goal 7 (SDG 7), it is crucial to have algorithms and models capable of accurately forecasting and improving process efficiency. These tools provide valuable insights for decision-makers' planning strategies. In this study, Data Envelopment Analysis (DEA) and bootstrap algorithms were used to assess Brazil's energy efficiency from 2004 to 2023. The performance of Autoregressive Integrated Moving Average (ARIMA) and Seasonal Autoregressive Integrated Moving Average (SARIMA) models was compared to forecast the trend of variables over the next twenty-four months. Additionally, the Generalized Autoregressive Conditional Heteroskedasticity (GARCH) model was applied to evaluate daily price forecasts for ELET3, correlated with industrial energy price trends and GDP. The study provides important conclusions about Brazil's energy landscape. According to the results, the annual energy consumption rate is expected to increase by an average of 2.1% until 2030, which is accompanied by an upward trend in GDP. Using existing technologies in the country, it is possible to reduce electricity consumption costs by an average of 30.58% while maintaining the same GDP level. Furthermore, a significant negative correlation of 33.91% was found between the ELET3 variable and the average industrial energy price, suggesting that as ELET3 prices fall, energy prices tend to rise. Moreover, an increase in stock price volatility was observed during periods marked by severe climate issues, such as droughts and dry spells, as well as fluctuations caused by the COVID-19 pandemic. The findings revealed a downward trend in stock prices along with increasing market instability, indicating the effectiveness of the GARCH model for stakeholders aiming to understand the energy market sector.

Keywords: Energy, GDP, Data Envelopment Analysis, Bootstrap, Forecast

SUMÁRIO

1	INTRODUÇÃO	1
1.1	OBJETIVOS	2
1.2	CONTRIBUIÇÕES DO TRABALHO	2
1.3	ESTRUTURA DA DISSERTAÇÃO	2
2	APLICAÇÃO DE ALGORITMOS BASEADOS EM MODELOS NÃO PARAMÉTRICOS E DE PRE- VISÃO PARA O DESENVOLVIMENTO SUSTENTÁVEL DO SETOR ENERGÉTICO NO BRASIL .	4
2.1	CONTEXTUALIZAÇÃO DO SETOR ENERGÉTICO BRASILEIRO	4
2.2	MODELOS DE PREVISÃO E TRABALHOS RELACIONADOS.....	6
2.2.1	MODELOS PREDITIVOS USANDO TÉCNICAS DE APRENDIZADO DE MÁQUINA	6
2.2.2	MODELOS ECONÔMICOS: ARIMA E SARIMA.....	7
2.2.3	COMPARATIVO ENTRE MODELOS DE PREVISÃO DE APRENDIZADO DE MÁQUINA E ECONÔMICOS CLÁSSICOS	8
2.3	CONSOLIDAÇÃO DO BANCO DE DADOS	10
2.4	ESTIMATIVA DA EFICIÊNCIA ENERGÉTICA	11
2.4.1	ANÁLISE ENVOLTÓRIA DE DADOS (DEA)	11
2.4.2	MÉTODO BOOTSTRAP PARA REMOÇÃO DE OUTLIERS	12
2.4.3	CONSTRUÇÃO DO INTERVALO DE CONFIANÇA COM <i>Bootstrap</i>	14
2.5	IDENTIFICAÇÃO DOS PARÂMETROS (P,D,Q).....	15
2.5.1	MODELOS AUTORREGRESSIVOS INTEGRADOS DE MÉDIAS MÓVEIS (ARIMA) E SAZONAIS (SARIMA)	16
2.5.2	TRANSFORMANDO VARIÁVEIS EM SÉRIES TEMPORAIS	17
2.5.3	TESTE DE SAZONALIDADE DE SÉRIES TEMPORAIS	19
2.5.4	TESTE DE PARÂMETROS COM AUTO.ARIMA	22
2.6	VERIFICAÇÃO DIAGNÓSTICA.....	24
2.7	RESULTADOS E DISCUSSÕES.....	24
2.7.1	ANÁLISE DE EFICIÊNCIA ENERGÉTICA	25
2.7.2	PREVISÃO SARIMA.....	26
2.8	DISCUSSÕES FINAIS SOBRE O ESTUDO DE CASO DE PREVISÃO DO SETOR ENERGÉTICO.....	29
3	APLICAÇÃO DA ANÁLISE GARCH PARA PREVISÃO DA VOLATILIDADE DOS PREÇOS DAS AÇÕES DA ELETROBRAS	31
3.1	CONTEXTUALIZAÇÃO DA VOLATILIDADE DOS PREÇOS DAS AÇÕES DE ENERGIA ...	31
3.2	MODELOS DE VOLATILIDADE E TRABALHOS RELACIONADOS	33
3.2.1	MODELOS TRADICIONAIS DE PREVISÃO ECONÔMICA	34
3.2.2	MODELOS DE PREVISÃO DE INTELIGÊNCIA ARTIFICIAL	35

3.2.3	MODELOS HÍBRIDOS DE PREVISÃO ECONOMÉTRICA E DE INTELIGÊNCIA ARTIFICIAL	35
3.2.4	DISCUSSÃO SOBRE OS PONTOS FORTES E AS LIMITAÇÕES DA AVALIAÇÃO DA VOLATILIDADE	36
3.3	MATERIAIS E MÉTODOS	38
3.3.1	MODELO AUTORREGRESSIVO GENERALIZADO DE HETEROCEDASTICIDADE CONDICIONAL (GARCH).....	38
3.4	PREPARAÇÃO DO BANCO DE DADOS	40
3.4.1	CONSOLIDAÇÃO DO BANCO DE DADOS	40
3.4.2	ANÁLISE DA PRESENÇA DE VOLATILIDADE NOS DADOS	41
3.5	TESTE DOS PARÂMETROS USANDO AUTO.ARIMA.....	43
3.6	VERIFICAÇÃO DOS RESÍDUOS DOS MODELOS.....	44
3.6.1	RESÍDUOS DO MODELO ARMA	44
3.6.2	RESÍDUOS DO MODELO GARCH	45
3.7	DISCUSSÃO E RESULTADOS	46
3.7.1	APLICAÇÃO DA PREVISÃO	46
3.7.2	CORRELAÇÃO COM OS PREÇOS DA ENERGIA INDUSTRIAL.....	49
3.8	DISCUSSÕES FINAIS SOBRE O ESTUDO DE CASO DE VOLATILIDADE DAS AÇÕES DE ENERGIA.....	51
4	CONCLUSÃO	53
4.1	PERSPECTIVAS PARA PESQUISAS FUTURAS	56
	REFERÊNCIAS BIBLIOGRÁFICAS	59

LISTA DE FIGURAS

2.1	Estrutura do algoritmo para aplicação da metodologia.....	9
2.2	<i>Boxplot</i> para o índice de eficiência dos dados com e sem <i>outliers</i>	13
2.3	Densidade do índice de eficiência dos dados com e sem <i>outliers</i>	13
2.4	<i>Boxplot</i> para o índice de eficiência dos dados sem e com correção.....	15
2.5	Série temporal de demanda de energia.....	17
2.6	Série temporal do PIB.....	18
2.7	Função de autocorrelação total da demanda de energia.....	18
2.8	Função de autocorrelação total do PIB.....	19
2.9	Decomposição da série temporal de demanda de energia.....	20
2.10	Decomposição da série temporal do PIB.....	20
2.11	FAC e FACP da demanda de energia.....	21
2.12	FAC e FACP do PIB.....	21
2.13	FAC e FACP da demanda de energia com uma diferença.....	22
2.14	FAC e FACP do PIB com uma diferença.....	22
2.15	Índice de eficiência de 2004 a 2023.....	26
2.16	Previsão da demanda de energia com SARIMA.....	28
2.17	Previsão do PIB com SARIMA.....	28
3.1	Estrutura da pesquisa desenvolvida.....	38
3.2	Volatilidade dos preços das ações da Eletrobras ao longo do tempo (diário).	41
3.3	Retorno Diário Eletrobras - ELET3.SA.....	42
3.4	Gráfico da função ACF do modelo.....	42
3.5	Gráfico da função PACF do modelo.....	43
3.6	Resíduo quadrado - ELET3.....	44
3.7	Gráfico de resíduos.....	44
3.8	Gráfico ACF dos resíduos.....	45
3.9	Gráfico PACF dos resíduos.....	45
3.10	Gráfico de probabilidade do modelo GARCH(1,1).....	45
3.11	Volatilidade condicional modelo GARCH(1, 1).....	46
3.12	Análise de resíduos.....	47
3.13	Previsões para 1.460 dias.....	48

LISTA DE TABELAS

2.1	Resumo estatístico da oferta e demanda de eletricidade e PIB. O intervalo de tempo considerado é entre 2004 e 2023.....	8
2.2	Resumo estatístico da oferta e demanda de eletricidade e PIB. O intervalo de tempo considerado é entre 2004 e 2023.....	11
2.3	Resumo estatístico dos dados de demanda e PIB sem <i>outliers</i> para o período entre 2004 e 2023.....	13
2.4	Resumo estatístico dos dados sem e com correção com intervalo de confiança de 95%.	15
2.5	Resumo estatístico para a série temporal de demanda de energia e PIB.	18
2.6	Valores de AIC, AICc e BIC para diferentes parâmetros usados no <code>auto.arima</code>	23
2.7	Valores de RMSE, MAE, MAPE e MASE para diferentes parâmetros usados no <code>auto.arima</code> . 23	
2.8	ME, RMSE, MAE, MAPE e MASE para diferentes parâmetros usados no ARIMA.	24
2.9	Teste de <i>Ljung-Box</i>	24
2.10	Escore de eficiência energética de 2004 a 2013.....	25
2.11	Escore de eficiência energética de 2014 a 2023.....	25
2.12	Previsão da demanda de energia (MWh)	27
2.13	Previsão do PIB (milhares de R\$)	27
3.1	Resumo estatístico dos preços da Eletrobras. Valores em BRL.	40
3.2	Resumo estatístico dos preços da Eletrobras com logaritmo e com uma diferenciação	42
3.3	Valores AIC, AICc e BIC dos parâmetros <code>auto.arima</code>	43
3.4	Valores ME, RMSE, MAE, MPE, MAPE e MASE para os parâmetros <code>auto.arima</code>	43
3.5	Valores ME, RMSE, MAE, MPE, MAPE e MASE para os modelos GARCH	47
3.6	Testes de resíduos quadrados de modelos GARCH.....	48
3.7	Resumo da previsão para os próximos 1.460 dias	49
3.8	Teste <i>Shapiro-Wilk</i> para Preço das Ações	50
3.9	Teste de Correlação de Spearman e Kendall	50
3.10	Intensidade da Correlação	50

1 INTRODUÇÃO

No Brasil, grande parte da matriz energética é composta por fontes renováveis, especialmente energia de origem hidráulica. Em 2022, a fonte hidráulica foi responsável pela produção de 427,1TWh, representando cerca de 61,9% da produção total do país de 690,1TWh, enquanto a fonte de energia solar fotovoltaica teve um valor de produção de 30,1TWh (4,4%), a energia solar térmica de 11,6TWh (1,7%) e a fonte de energia eólica de 81,6TWh (11,8%). De todos os consumidores finais, a indústria é a principal responsável pelo consumo de energia elétrica, que teve um consumo de 218,7TWh (31,7%) no ano de 2022. Se desconsideradas as perdas operacionais, a indústria é responsável pelo consumo de 37,3% de energia elétrica do país (EPE).

Isso mostra que tornar a matriz energética eficiente engloba fatores sustentáveis e fatores econômicos para tornar a indústria brasileira mais competitiva internacionalmente, no entanto, existe uma série de melhorias e mudanças necessárias para tornar essa afirmação em uma realidade palpável e aplicável ao País considerando a confiabilidade, segurança e capacidade dos processos de geração, transmissão e distribuição de energia elétrica. Uma dessas questões apontadas na literatura tem sido problemas com ataques tanto no domínio de geração, de distribuição e até o consumo de energia. Como o caso de clientes mal-intencionados, que invadem a rede elétrica para acessarem a rede de controladores inteligentes de geração de energia elétrica renováveis, a fim de adulterar os valores marcados e reivindicarem maiores valores de energia fornecida (Ismail et al., 2020).

Com crescimento de tecnologias de comunicação os ataques tornaram-se um problema que não deve ser ignorado, nas redes elétricas os ataques podem ocorrer nos sensores, medidores inteligentes, serviços auxiliares de manutenção de frequências, estabilidade de tensões, estabilidade de rede de fluxo de energia, além de ataques relacionados aos fluxos de informação gerando uma série de malefícios (Sahoo et al., 2021). Na etapa de distribuição da rede corrente, há os chamados ataques cibernéticos nas microrredes de corrente contínua (CC), onde por meio da injeção de dados falsos (FDIAs) na rede, tem a intenção de gerar uma desregulação dos valores referentes a tensão e a corrente (Habibi et al., 2021). Lin et al. (2024) aponta que os ataques FDIA têm sido muito estudados nos últimos anos devido às suas ameaças potenciais no gerenciamento de energia em tempo real.

A sociedade na sua totalidade tem caminhado para se tornar cada vez mais dependente de tecnologias de redes inteligentes (sistemas avançados interligados que aumentam a eficiência, produtividade, confiabilidade e flexibilidade dos mecanismos de gestão e distribuição de energia) o que por sua vez torna o sistema de energia mais fácil de gerir; no entanto, concomitantemente, mais suscetível a sofrer ataques, devido ao aumento da sua vulnerabilidade. Além disso, há desafios sendo introduzidos com o aumento da integração das redes convencionais de energia com as novas fontes de energia renovável, que produzem uma maior instabilidade nos sistemas, devido ao aumento de oscilações que podem gerar ruídos potenciais para perturbar a rede de energia (Aygul et al., 2024).

Dessa forma, torna-se essencial compreender o mercado energético brasileiro, bem como suas principais implicações. Propondo metodologias e métodos, que auxiliem no desenvolvimento de tecnologias

e alternativas, que permitam melhorias na infraestrutura de redes inteligentes, tornando o uso de recursos energéticos mais eficiente, assim como propor alterações de políticas para o devido planejamento das partes interessadas e trazer benefício para a indústria, que possui uma parcela considerável dos seus gastos advindo do custo com a energia elétrica (Yasmeen et al., 2022).

1.1 OBJETIVOS

O objetivo geral do estudo é propor metodologias para análise da eficiência energética brasileira e analisar o comportamento de tendências e volatilidade do mercado de energia para o futuro. Os objetivos específicos para o cumprimento do objetivo geral são descritos da seguinte forma:

- Avaliar a eficiência energética brasileira, por meio da demanda de energia em comparação com o resultado do produto Interno Bruto (PIB) do país, por meio da Análise Envoltória de Dados (DEA) juntamente com métodos computacionais bootstrap.
- Aplicação dos modelos Autorregressivo Integrado de Médias Móveis (ARIMA) e Autorregressivos Integrados de Médias Móveis com Sazonalidade (SARIMA).
- Aplicação dos modelos Heterocedasticidade Condicional Autorregressiva (ARCH) e Heterocedasticidade Condicional Autorregressiva Generalizada (GARCH) para previsão da volatilidade de preços do mercado de energia.
- Correlação do comportamento do índice da bolsa da Eletrobras (ELET3) com os preços anuais de energia da indústria e a previsão do PIB para previsão dos preços diários de energia elétrica.

1.2 CONTRIBUIÇÕES DO TRABALHO

O presente estudo contribui para o crescimento da literatura com a demonstração de métodos e metodologias que possam servir de base para análise do setor energético brasileiro. De acordo com Arnob et al. (2023a) existem lacunas na literatura frente aos parâmetros e ajustes para os modelos de previsão, por haver inconsistências quanto a padronização do consumo de energia, clima e economia de cada país, onde a maioria dos estudos são conduzidos na China, Irã, Turquia e Índia, tornando o Brasil um dos países na fronteira de singularidade, que demonstra a necessidade de estudos voltados para área, assim como a apresentação de ajustes e parâmetros condizentes a realidade brasileira.

1.3 ESTRUTURA DA DISSERTAÇÃO

O estudo está organizado em três seções, sendo que a seção 2 apresenta a metodologia e a aplicação de um estudo de caso do método não paramétrico da Análise Envoltória de Dados (DEA) para avaliação da eficiência energética da matriz brasileira e os modelos Autorregressivo Integrado de Médias Móveis

(ARIMA) e Autorregressivos Integrados de Médias Móveis com Sazonalidade (SARIMA) para previsão do comportamento de energia nos próximos 2 anos. A seção 3 apresenta a metodologia e a aplicação de um segundo estudo de caso utilizando os modelos Heterocedasticidade Condicional Autorregressiva (ARCH) e Heterocedasticidade Condicional Autorregressiva Generalizada (GARCH) para previsão da volatilidade dos preços de energia da matriz energética brasileira, correlacionando os preços anuais da energia elétrica da indústria, a previsão do Produto Interno Bruto (PIB) e o índice de eletricidade (IEE B3). A seção 4 finaliza o estudo com as principais conclusões da pesquisa.

2 APLICAÇÃO DE ALGORITMOS BASEADOS EM MODELOS NÃO PARAMÉTRICOS E DE PREVISÃO PARA O DESENVOLVIMENTO SUSTENTÁVEL DO SETOR ENERGÉTICO NO BRASIL

2.1 CONTEXTUALIZAÇÃO DO SETOR ENERGÉTICO BRASILEIRO

A integração de fontes de energia renováveis é uma tendência global nos sistemas de distribuição de energia. Permite-lhes i) enfrentar os desafios energéticos não resolvidos colocados pelas centrais elétricas tradicionais centralizadas, ii) reduzir as emissões globais de CO₂ e iii) aumentar o fornecimento a longo prazo de energia sustentável (Ang et al., 2022). Os recursos energéticos distribuídos (DERs), especialmente as energias eólica e solar fotovoltaica, desempenham um papel cada vez mais importante no setor energético (Weschenfelder et al., 2020).

É crucial compreender a procura de energia elétrica para poder alocar e gerir eficientemente os recursos energéticos dentro de um sistema com capacidade de produção limitada. Esta compreensão também pode ajudar a alcançar os objetivos de desenvolvimento sustentável, particularmente o Objetivo de Desenvolvimento Sustentável 7 (ODS 7) e as suas metas relacionadas. É importante notar que a energia representa um custo econômico significativo para um país e deve ser tida em conta na medição do Produto Interno Bruto (PIB). Isto destaca a importância de uma gestão eficiente neste setor (Caldeira et al., 2016; Yasmeen et al., 2022).

Os Objetivos de Desenvolvimento Sustentável abrangem uma ampla gama de metas e diretrizes, que incluem o seguinte: garantir o acesso universal a serviços energéticos modernos, viáveis e acessíveis (7.1), aumentar a participação das energias renováveis na matriz energética global (7.2), duplicar a taxa global de melhoria na eficiência energética (7.3), fortalecimento das medidas de cooperação internacional para pesquisa e tecnologia de energia limpa (7.4) e melhoria da infraestrutura para serviços de energia sustentável nos países em desenvolvimento (7.5) (Shyu, 2021). Além disso, mostra a ação do Brasil para atingir essas metas. A compreensão da produção de energia, conforme mostrado nas diversas matrizes que compõem a rede elétrica nacional do Brasil, bem como a capacidade de prever a demanda futura de energia, são características críticas para os serviços públicos, formuladores de políticas e partes interessadas, pois lhes permitem tomar decisões informadas sobre os objetivos de desenvolvimento sustentável (Caldeira et al., 2016; Alizadeh et al., 2020).

As tendências positivas nas áreas socioeconômicas e no tratamento dos fatores climáticos, juntamente com os avanços tecnológicos, influenciaram os padrões de fornecimento e consumo de energia durante longos períodos e atingiram assim o objetivo do ODS 7, que é alargar o acesso a serviços energéticos modernos (7.1) e melhorar a eficiência energética (7.3). A sazonalidade de curto prazo, incluindo variações nas condições meteorológicas e nos ciclos de produção industrial, também conduz a flutuações na procura de energia. Estes são fatores cruciais para alcançar as metas do ODS 7, particularmente na integração de

fontes de energia renováveis. (7.2) e desenvolvimento de infraestrutura (7.5) (Yasmeen et al., 2022). Deve ser enfatizado que as flutuações econômicas e as mudanças políticas sublinham a necessidade de mecanismos robustos e capacidade de previsão para apoiar medidas de cooperação internacional em investigação e tecnologia de energia limpa (7.4) (Yasmeen et al., 2022). Ao selecionar técnicas de previsão apropriadas, serviços públicos e serviços privados, os investigadores podem contribuir para os esforços globais para melhorar o acesso à energia limpa e criar uma infraestrutura energética sustentável (7.a, 7.b) (Yasmeen et al., 2022).

O uso de técnicas de previsão tem atraído interesse significativo na literatura e tem sido explorado em diversos contextos. Estes incluem a sua aplicabilidade prática, eficácia em diferentes setores industriais, adequação para ajudar organizações de vários tamanhos, utilização para horizontes temporais e graus de precisão (Ahmad and Chen, 2020). Desta forma, este estudo realiza uma análise comparativa de modelos ARIMA sazonais para energia no Brasil, buscando *insights* sobre seus respectivos pontos fortes e fracos na captura da complexa dinâmica da relação entre a demanda de energia e o PIB no Brasil. Além disso, o estudo pretende oferecer recomendações para melhorar a previsão da procura, o que contribuirá para uma alocação e gestão mais eficiente dos recursos no setor energético, promovendo o progresso em direção aos objetivos de desenvolvimento sustentável.

Além disso, este estudo procura fazer recomendações para melhorar a previsão da procura de energia, melhorando a gestão eficiente dos recursos no setor energético. Isto significa que o progresso rumo aos ODS pode ser acelerado e beneficiar diretamente a economia. Usando dados mensais de demanda de eletricidade e PIB do Brasil entre janeiro de 2004 e dezembro de 2023, os métodos de previsão estatística do ETS (erro, tendência e sazonalidade) foram comparados com a modelagem ARIMA usando simulações computacionais no software R. O objetivo foi avaliar o desempenho dos modelos de previsão e sua relação com o PIB brasileiro, alinhados aos objetivos do ODS 7. Visa aprimorar as práticas de previsão de demanda de energia, considerando a contribuição das fontes hidráulica, eólica e solar.

A novidade deste trabalho é ilustrada pelo fato de não terem sido encontrados na literatura estudos semelhantes que adotem uma abordagem com métodos computacionais *bootstrap-DEA*, assim como propõe utilizar os modelos ARIMA e SARIMA para previsão de tendências para avaliar o setor energético brasileiro. Apesar desses modelos gerarem dois produtos diferentes, a ideia principal para utilização dos dois modelos é fornecer um panorama geral do cenário energético brasileiro. Assim, as metodologias de pesquisa são empregadas para avaliar este setor, auxiliar na formulação de políticas públicas e fornecer uma base para a criação de novas tecnologias adequadas. Os dados em séries temporais muitas vezes acumulam volatilidade. Porém, este estudo não busca explorar a volatilidade dos dados, como os modelos de Heterocedasticidade Condicional Autorregressiva (ARCH) e Heterocedasticidade Condicional Autorregressiva Generalizada (GARCH). A análise centrar-se-á, em vez disso, na determinação da presença ou ausência de fatores sazonais, uma vez que os dados são organizados mensalmente e não diariamente.

A estrutura do capítulo está organizada da seguinte forma. A Seção 2.2 apresenta uma revisão da literatura sobre modelos de previsão de séries temporais e sua aplicabilidade no setor de energia. Esta seção reúne os estudos mais recentes sobre o tema e inclui uma análise crítica dos achados. Na sequência, a Seção 3.3 fornecerá uma visão geral dos materiais e métodos utilizados neste estudo, detalhando o algoritmo empregado e explicando a estrutura das Seções 2.3, 2.4, 2.5 e 2.6. Essas seções oferecem uma

explicação detalhada dos modelos matemáticos e procedimentos aplicados, incluindo a implementação do algoritmo *bootstrap-DEA* e o uso de previsões ARIMA e SARIMA para analisar o setor de energia no Brasil. A Seção 2.7 discute os principais achados e resultados da aplicação do *bootstrap-DEA*. Os cálculos do índice de eficiência e seus respectivos intervalos de confiança de 95% também são apresentados. Além disso, é discutida a aplicação do modelo SARIMA para prever a série temporal de demanda de energia e produto interno bruto (PIB) no Brasil. Finalmente, a Seção 2.8 conclui o capítulo resumindo as principais contribuições da pesquisa.

2.2 MODELOS DE PREVISÃO E TRABALHOS RELACIONADOS

A necessidade de desenvolver um plano energético sustentável e a ênfase colocada na previsão da demanda de energia ressaltam a importância de gerenciar essa demanda de maneira eficaz, alinhada aos objetivos de longo prazo das metas de sustentabilidade ambiental, social e econômica. Ao integrar a previsão da demanda de energia no planejamento energético sustentável, as sociedades podem avançar em direção a um futuro que seja tanto sustentável quanto equitativo—combinando conhecimento sobre produtividade, sustentabilidade e indústria (Bispo et al., 2024; Serrano et al., 2024).

Portanto, prever a demanda de energia é fundamental para o planejamento e a gestão de energia, permitindo que as partes interessadas tomem decisões informadas sobre alocação de recursos, desenvolvimento de infraestrutura, formulação de políticas e seu impacto no PIB. Os métodos de previsão da demanda de energia podem ser amplamente categorizados em abordagens qualitativas e quantitativas. Conforme documentado por Arnob et al. (2023b), os modelos de séries temporais são amplamente utilizados para previsão da demanda de energia porque capturam os padrões temporais e as tendências inerentes aos dados de consumo de energia.

2.2.1 Modelos Preditivos Usando Técnicas de Aprendizado de Máquina

Vários estudos compararam o desempenho de modelos de séries temporais em diferentes contextos e para diversas aplicações. Uma revisão abrangente dos métodos de previsão, incluindo abordagens clássicas e modernas, foi realizada (Rafayal et al., 2022). O estudo mostrou que os modelos de aprendizado profundo mais genéricos, como LSTM e GRU, tiveram melhor desempenho do que modelos mais simples, como *DeepAR* e *DeepState*. No entanto, esses ganhos não foram significativos, e o resultado é pior ao se observar cenários mais prolongados. De forma semelhante, Riady and Apriani (2023) usou uma rede neural autorregressiva não linear (NAR) para prever a demanda de energia da próxima década com base em um conjunto de dados públicos sobre consumo global de energia; entretanto, o modelo que apresentou melhor desempenho (*FB prophet*) é difícil de aplicar (Serrano et al., 2024).

A previsão da demanda de energia tem sido uma área significativa de pesquisa, com vários métodos propostos para previsões precisas. Redes neurais, como redes neurais autorregressivas não lineares, podem lidar com problemas estatísticos, empíricos e teóricos de forma eficaz, conforme apresentado por Al-Haija et al. (2023); no entanto, de forma semelhante ao estudo de Riady and Apriani (2023), onde o melhor modelo de previsão, que neste caso é o NAR, também apresenta limitações em termos de aplicação devido

à complexidade do modelo, resultando em vários erros. Redes neurais convolucionais (CNNs) e campos aleatórios condicionais (CRFs) têm sido usados para prever o consumo de energia, alcançando alta precisão, de acordo com Thangavel and Govindaraj (2022), embora altamente precisos em um cenário real onde há uma grande quantidade de dados, uma quantidade considerável de GPUs será necessária para executar os modelos. A previsão de preços de eletricidade também é relevante neste contexto (Gundu and Simon, 2021).

Os países em desenvolvimento enfrentam desafios para alcançar o crescimento econômico sustentável com baixas emissões de carbono. Ainda assim, para Arnob et al. (2023b), redes neurais artificiais (ANNs) hibridizadas com técnicas meta-heurísticas são superiores na previsão de carga, embora naturalmente mais complexas. Modelos de previsão probabilística, como *DeepAR* e *deep state space*, têm melhor desempenho para horizontes de previsão mais longos, conforme conduzido por Rafayal et al. (2022).

Técnicas de aprendizado de máquina (ML) são amplamente utilizadas, especialmente para previsões de eletricidade de curto prazo, enquanto modelos baseados em engenharia cobrem horizontes temporais longos e o consumo de eletrodomésticos; de acordo com Verwiebe et al. (2021), questões importantes levantadas pelo estudo apontam para a necessidade de se ter modelos de demanda de energia industrial, assim como o fato de que diferentes modelos têm desempenhos diferentes quando aplicados a outros cenários. Na geração de energia, modelos de previsão também têm sido usados para prever parâmetros de fornecimento e produção de eletricidade, como a velocidade do vento (Sengar and Liu, 2020; Liu et al., 2018). No entanto, essa abordagem é interessante para ser aplicada em cenários que consideram a energia eólica; ela não ajuda a avaliar uma matriz energética com predominância de geração hidrelétrica e solar, como no Brasil (Serrano et al., 2024).

2.2.2 Modelos Econométricos: ARIMA e SARIMA

Dessa forma, entende-se que os estudos mais recentes que tratam da previsão da demanda de energia têm se concentrado no uso de técnicas de aprendizado de máquina. Apesar de apresentarem bons resultados em comparação com métodos mais convencionais, essas técnicas têm enfrentado dificuldades devido à complexidade para aplicação prática. Portanto, modelos tradicionais encontrados na literatura, como ARIMA e SARIMA, são mais adequados para países como o Brasil, pois são simples de aplicar e interpretar.

Por outro lado, o modelo ARIMA baseia-se no ajuste dos valores observados com o objetivo de reduzir a diferença entre os valores produzidos pelo modelo e os valores observados para próximo de zero. Para a construção de modelos ARIMA, *Box-Jenkins* sugeriu as seguintes etapas iterativas; segundo Bayer, o modelo ARIMA (p, d, q) consiste na combinação de um modelo autorregressivo (AR) de ordem p com uma série temporal diferenciada d vezes (número de diferenças necessárias para tornar a série estacionária) e um modelo de média móvel (MA) de ordem q . Assim, ele tem a seguinte forma: ϵ é o ruído branco, $f(B)$ e $q(B)$ são o polinômio autorregressivo e o polinômio de média móvel, respectivamente. O modelo ARIMA sazonal, ou SARIMA, segue este formato de apresentação: SARIMA $(p, d, q)(P, D, Q)$, onde os parâmetros (P, D, Q) são os equivalentes sazonais de (p, d, q) . A parte sazonal é representada por três parâmetros adicionais semelhantes, sendo assim denominado SARIMA. O modelo SARIMA é definido por uma parte simples de parâmetros (p, d, q) e uma parte sazonal composta por parâmetros (P, D, Q)

(Gujarati and Porter, 2011).

2.2.3 Comparativo entre Modelos de Previsão de Aprendizado de Máquina e Económétricos Clássicos

Numerosos artigos na literatura aprofundam-se nos aspectos conceituais e técnicos dos modelos de previsão temporal, como ARIMA e SARIMA (Kontopoulou et al., 2023; Yamak et al., 2020; Siami-Namini et al., 2019; Box et al., 2015; Rundo et al., 2019; Siami-Namini et al., 2018; Spyrou et al., 2022; ArunKumar et al., 2022; Zhou et al., 2020; Priyadarshini et al., 2022; Liu et al., 2021; Dimri et al., 2020; Malki et al., 2021; ArunKumar et al., 2021; Shadab et al., 2020).

Muitos dos estudos incluíram comparações entre modelos tradicionais e modelos de aprendizado profundo. Algumas pesquisas indicaram a superioridade dos modelos estatísticos convencionais sobre algoritmos de inteligência artificial (IA) (Yamak et al., 2020; Spyrou et al., 2022; Priyadarshini et al., 2022; Liu et al., 2021; Dimri et al., 2020), enquanto outros demonstraram o melhor desempenho de modelos de previsão baseados em IA (Kontopoulou et al., 2023; Siami-Namini et al., 2019; Rundo et al., 2019; Siami-Namini et al., 2018).

Alguns estudos não mostraram diferenças significativas entre modelos tradicionais e de IA (ArunKumar et al., 2022; Zhou et al., 2020). A Seção 3.3 discute os modelos matemáticos que envolvem esses testes estatísticos e os modelos de previsão usados para as previsões com os modelos ARIMA e SARIMA. Esta seção descreve a metodologia empregada nesta pesquisa, detalhando as ferramentas estatísticas e o conjunto de dados utilizados. A tabela 2.1 faz um resumo de algum dos trabalhos relacionados, para melhor contribuição da pesquisa.

Tabela 2.1: Resumo estatístico da oferta e demanda de eletricidade e PIB. O intervalo de tempo considerado é entre 2004 e 2023.

Referência	ARIMA/SARIMA	DEA	Bootstrap-DEA
Kontopoulou et al. (2023)	✓	×	×
Yamak et al. (2020)	✓	×	×
Siami-Namini et al. (2019)	✓	×	×
Box et al. (2015)	✓	×	×
Rundo et al. (2019)	✓	×	×
Siami-Namini et al. (2018)	✓	×	×
Zhou et al. (2020)	✓	×	×
Priyadarshini et al. (2022)	✓	×	×
Liu et al. (2021)	✓	×	×
Dimri et al. (2020)	✓	×	×
Malki et al. (2021)	✓	×	×
ArunKumar et al. (2021)	✓	×	×
Shadab et al. (2020)	✓	×	×
Este trabalho	✓	✓	✓

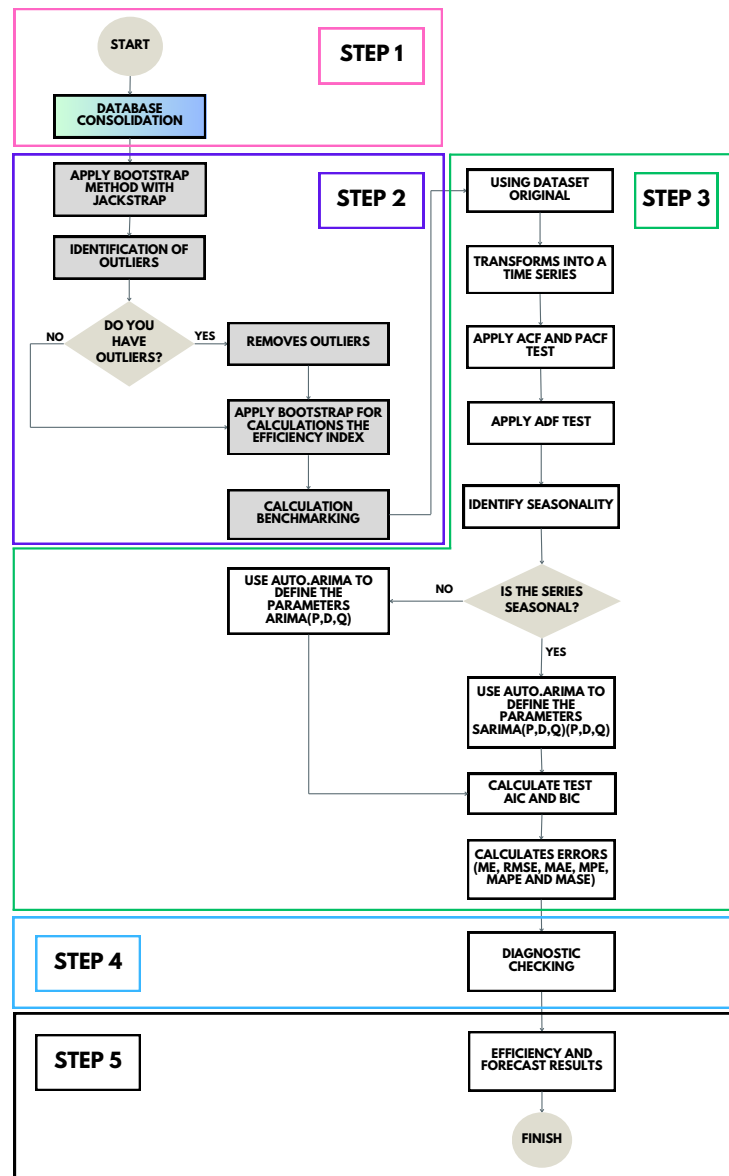


Figura 2.1: Estrutura do algoritmo para aplicação da metodologia.

A estrutura do algoritmo utilizado para implementar a metodologia de pesquisa é ilustrada na Figura 2.1.

A primeira etapa (STEP 1) do algoritmo envolve a consolidação do banco de dados, que inclui o pré-processamento e a estruturação dos dados. Os dados são originários de duas fontes principais: dados de demanda de eletricidade da Empresa de Pesquisa Energética (EPE) (, EPE) e dados do PIB do Instituto de Pesquisa Econômica Aplicada (IPEA). O conjunto de dados abrange o período de 2004 a 2023, mensalmente. A consolidação do banco de dados é descrita na Seção 2.3.

A segunda etapa (STEP 2) do algoritmo tem como objetivo medir a eficiência energética do Brasil de 2004 a 2023. A análise envoltória de dados (DEA) é uma metodologia utilizada para medir a eficiência energética. Para aplicar a DEA, uma abordagem *Jackknife* é utilizada (Stošić and de Sousa, 2003) para remover *outliers* e, em seguida, métodos computacionais de *bootstrap* são aplicados para construir

intervalos de confiança para os índices de eficiência (Simar and Wilson, 1998; Rosano-Peña et al., 2022; Marques Serrano et al., 2024; Saiki et al., 2024). A metodologia e os procedimentos utilizados para aplicar o modelo DEA são detalhados na Seção 2.4.

A terceira etapa (STEP 3) do algoritmo aplica as duas primeiras etapas do método Box-Jenkins: identificação dos parâmetros (p, d, q) usando a função de autocorrelação e a função de autocorrelação parcial, utilizando correlogramas, função de autocorrelação (ACF) e função de autocorrelação parcial (PACF), teste de *Dickey-Fuller* aumentado (ADF), teste de *Kwiatkowski-Phillips-Schmidt-Shin* (KPSS) e teste de *Phillips-Perron* para verificar a sazonalidade do modelo, e estimativa dos parâmetros selecionados, garantindo que os parâmetros escolhidos, como o número de termos autorregressivos (p), número de diferenciais (d) e número de termos de média móvel (q), se ajustem aos dados originais, utilizando resíduos e funções ACF e PACF. Outra opção é utilizar `auto.arima` para determinar esses parâmetros. A Seção 2.5 descreve a metodologia ARIMA e SARIMA, incluindo os procedimentos para identificação dos parâmetros (p, d, q) e estimativa dos erros desses parâmetros.

A quarta etapa (STEP 4) do algoritmo aplica a terceira etapa do método *Box-Jenkins*: verificação e diagnóstico, analisando se as previsões são adequadas e precisas, e verificando o desempenho do modelo por meio da análise dos resíduos estimados. Nesta etapa, se for identificado que os parâmetros não são adequados para a previsão, a terceira etapa (STEP 3) deve ser repetida até que a quarta etapa (STEP 4) seja aceita, para que os parâmetros definidos possam ser utilizados para prever séries temporais (Jamil, 2020). Os procedimentos adotados para verificar o diagnóstico são descritos na Seção 2.6.

A quinta etapa (STEP 5) apresenta os resultados e a discussão dos índices de eficiência energética e o melhor modelo para prever os dados. Os resultados são apresentados na Seção 2.7.

2.3 CONSOLIDAÇÃO DO BANCO DE DADOS

Esta seção detalha a etapa 1 do algoritmo mostrado na Figura 2.1.

O banco de dados da pesquisa é composto por 240 linhas com duas colunas representando as variáveis (demanda e PIB) para cada mês, de 2004 a 2023. Para a análise de eficiência, considera-se o valor de x_1 , que representa a demanda de energia como entrada, e y_1 , que representa o produto interno bruto como saída. Da mesma forma, a demanda de energia e o produto interno bruto também são considerados na análise de previsão. Para avaliar se as variáveis estão sujeitas a comportamento sazonal durante as previsões, os valores devem ser mensais, por isso foram escolhidos valores mensais. O banco de dados estruturado e o código utilizado está disponível no *figshare*¹. A Tabela 2.2 mostra os resumos estatísticos das variáveis do estudo, incluindo o mínimo, 1º quartil, mediana, média, 3º quartil e máximo para demanda e PIB.

¹doi.org/10.6084/m9.figshare.25714932.v2

Tabela 2.2: Resumo estatístico da oferta e demanda de eletricidade e PIB. O intervalo de tempo considerado é entre 2004 e 2023.

	Demanda (GWh)	PIB (milhões de R\$)
Mín	26.508	142.861
1º Qu.	32.234	267.691
Mediana	37.866	459.337
Média	36.622	457.192
3º Qu.	40.269	582.831
Máx	46.407	954.063

O software R foi utilizado para estruturação do banco de dados, remoção de *outliers*, cálculo de eficiência e previsão.

2.4 ESTIMATIVA DA EFICIÊNCIA ENERGÉTICA

Esta seção explica a metodologia DEA, os procedimentos para remoção de *outliers* e o cálculo dos intervalos de confiança por *bootstrap* para o índice de eficiência energética. Portanto, esta seção apresenta a metodologia e os procedimentos relacionados à etapa 2, conforme mostrado na Figura 2.1.

2.4.1 Análise Envoltória de Dados (DEA)

A Análise Envoltória de Dados (DEA) é um modelo matemático baseado em econometria e pesquisa operacional. O modelo tem origem nos conceitos de produtividade introduzidos por Farrell (1957), onde produtividade é definida como a relação entre produtos e insumos de um processo produtivo específico. Medidas e métricas anteriores eram excessivamente específicas e impraticáveis, especialmente ao lidar com várias variáveis. Com base neste conceito de produtividade, Charnes et al. (1978) propuseram o modelo de eficiência relativa, com o objetivo de utilizar de forma ótima tanto os produtos quanto os insumos, considerando os retornos constantes de escala das unidades de tomada de decisão (DMU_i) por meio de problemas de programação matemática (Chen et al., 2021). As DMU_i referem-se às unidades sendo comparadas, como hospitais, escolas, organizações, empresas e países, entre outros (Bogetoft and Otto, 2010).

O modelo DEA é uma abordagem não paramétrica, o que significa que não exige conhecimento específico da natureza do problema. Essa flexibilidade inerente o torna mais aplicável e robusto do que modelos paramétricos, como a análise de fronteira estocástica (SFA), simplificando sua aplicação em vários cenários. Essa vantagem é particularmente evidente no setor empresarial, onde os modelos não paramétricos superam os paramétricos. Como resultado, o modelo DEA tem sido amplamente aplicado na literatura em diversos campos, incluindo industrial, energético, econômico, mineração, meio ambiente, indústria da água, políticas públicas e agronegócio (Matsumoto and Chen, 2021; Yu et al., 2021; Liu et al., 2019; Gökgöz and Yalçın, 2023; Sala-Garrido et al., 2023; Antunes et al., 2023; Chen et al., 2021; Pokushko

2.4.2 Método Bootstrap para Remoção de Outliers

Outliers são pontos de dados considerados anormais dentro de um conjunto de dados, pois diferem significativamente das outras observações. Vários fatores podem causar esse comportamento. Os principais fatores identificados como causadores de *outliers* nos dados são i) erros no banco de dados que ocorrem durante a coleta e organização dos dados, ii) dados precisos e altamente atípicos, e iii) dados reais e deficientes. Dependendo da sensibilidade do modelo a *outliers*, esses fatores podem causar distorções significativas e enviesar os resultados, tornando sua remoção necessária (Bogetoft and Otto, 2010).

A presença de *outliers* no modelo DEA cria uma situação problemática ao lidar com a fronteira de um conjunto de unidades produtivas eficientes, pois medidas e valores determinísticos são sensíveis a erros. Portanto, a remoção de *outliers* é necessária para evitar enviesamento na análise e produzir resultados mais robustos para o modelo (Rosano-Peña et al., 2022).

Uma abordagem é usar a técnica de alavancagem *Jackknife* (Stošić and de Sousa, 2003). Esta técnica avalia o impacto da remoção de cada ponto de dados do conjunto. Nesta aplicação, observa-se o efeito dos dados removidos sobre os valores de eficiência de outras DMU_i em combinação com a técnica de reamostragem *bootstrap* para remover *outliers*.

O procedimento é o seguinte: I) Calcular as pontuações de eficiência de todas as DMU_i pelo modelo clássico (Charnes et al., 1978), gerando um conjunto de eficiências dado por $\{\theta_i \mid i = 1, \dots, n\}$; II) Selecionar aleatoriamente um subconjunto K com $(K = 1, \dots, k)$ correspondente a 10% da amostra original de DMU_i , resultando em um conjunto de DMU_K ; III) Calcular as pontuações de eficiência de todas as DMU_K selecionadas $\{\theta_K \mid K = 1, \dots, k\}$ por meio de reamostragem *bootstrap* B vezes, onde B assume valores de $(B = 1, \dots, b)$; IV) Avaliar o impacto usando uma medida estatística para analisar se houve mudanças significativas nas pontuações de eficiência por meio da alavancagem de cada DMU_K selecionada em B , armazenando as informações de alavancagem em l_{KS} ; V) Repetir os passos ii, iii e iv S vezes com $(S = 1, \dots, s)$; VI) Calcular a alavancagem local; VII) Calcular a alavancagem global por meio do desvio padrão das medidas de eficiência antes e depois de remover os dados. Para mais detalhes sobre como calcular a alavancagem, consulte a referência fornecida (Stošić and de Sousa, 2003). O uso da alavancagem minimiza a probabilidade de selecionar *outliers* para reamostragem aleatória. Essa probabilidade é determinada usando a função de *Heaviside* (Stosic, 2005). Portanto, qualquer DMU com um valor de alavancagem significativamente maior que a alavancagem global será eliminada (Stošić and de Sousa, 2003).

Dois métodos foram usados para detectar anomalias nos dados. O primeiro método, o teste de *Kolmogorov-Smirnov* (K-S), identificou 2,08% dos dados como anômalos. O segundo método, o critério de *Heaviside*, detectou 5,42% de *outliers*. Como o critério de *Heaviside* foi mais eficaz, prosseguimos com este método e removemos os *outliers* do conjunto de dados original, resultando em uma estrutura de dados consolidada com 227 observações. O resumo estatístico dos dados de entrada e saída do modelo sem *outliers* pode ser encontrado na Tabela 2.3.

Tabela 2.3: Resumo estatístico dos dados de demanda e PIB sem *outliers* para o período entre 2004 e 2023.

	Demanda (GWh)	PIB (milhões de R\$)
Mín	27.657	156.954
1º Qu.	32.480	273.200
Mediana	37.867	458.517
Média	36.722	452.982
3º Qu	40.078	573.219
Máx	46.407	950.791

Ao comparar os valores nas Tabelas 2.2 e 2.3, nota-se uma diferença nos valores médios de demanda e PIB. Enquanto os valores de demanda são próximos aos valores originais (36.622 para 36.722), há uma diferença mais perceptível no produto interno bruto (PIB) (457.192 para 452.982). Para entender melhor o possível impacto negativo dessas observações, a medida de eficiência de cada mês foi calculada antes e depois do processo de remoção de *outliers*. Essas informações podem ser encontradas nas Figuras 2.2 e 2.3.

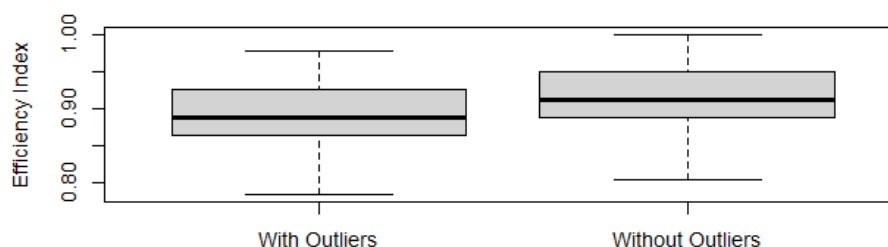


Figura 2.2: *Boxplot* para o índice de eficiência dos dados com e sem *outliers*.

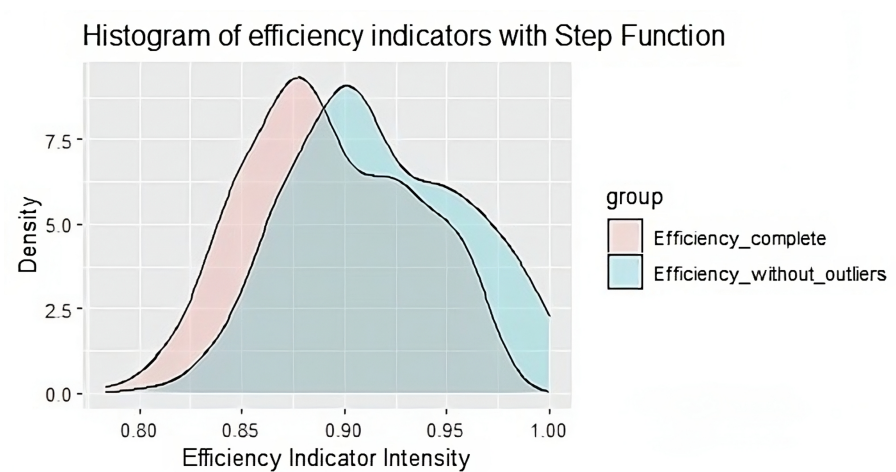


Figura 2.3: Densidade do índice de eficiência dos dados com e sem *outliers*.

Como visto na Figura 2.2, que mostra a mudança no valor mediano, e na Figura 2.3, que ilustra a distribuição de densidade dos dados, ao limpar o banco de dados, os valores se deslocam para valores mais altos. O conjunto completo de dados tinha um valor mediano de 0.888, enquanto os valores sem *outliers*

possuem uma mediana próxima de 0.912, representando uma diferença de 0.024. Foi realizado um teste de *Wilcoxon* para determinar se essa diferença é estatisticamente significativa. O valor $V = 25.878$ e o p-valor $< 2.2e^{-16}$ indicam que essa diferença é substancial e deve ser considerada relevante. Os dados foram então parametrizados. Isso ocorre porque o R.

2.4.3 Construção do Intervalo de Confiança com *Bootstrap*

Devido à ausência de ferramentas de inferência estatística para modelos clássicos não paramétricos, podemos utilizar métodos baseados em *bootstrap* para calcular a eficiência DEA, incluindo inferências estatísticas, testes de hipóteses e intervalos de confiança. Este modelo é um modelo DEA estocástico. Ele introduz o conceito de *bootstrap* e, posteriormente, métodos descritivos para identificar dados influentes para cálculos não paramétricos (Wilson, 1995; Simar and Wilson, 1998). Esses métodos permitem o uso de inferências estatísticas sem comprometer a natureza não paramétrica do problema (Stosic, 2005).

Os autores Simar and Wilson (1998) aplicaram o *bootstrap* para estimar intervalos de confiança das medidas de eficiência fornecidas pela DEA. O *bootstrap* simula uma amostra aplicando o estimador original e fazendo com que os resultados da simulação repliquem o original através de um processo de reamostragem, repetido W vezes (geralmente $W = 2000$). Este processo pode ser descrito em quatro etapas: I) Para cada observação $(x_i, y_i), i = 1, \dots, n$, calcular a correspondente pontuação de eficiência DEA $\hat{\theta}_i$ usando programação linear; II) Sorteie um conjunto de dados da amostra original aleatoriamente usando *bootstrap*, gerando uma amostra aleatória de tamanho P , onde P é do mesmo tamanho que a amostra original. Para esta amostra, obtenha $\hat{\theta}_r^*, r = 1, \dots, p$; III) A partir desta amostra aleatória $\hat{\theta}_r^*$, construa $\{x_r^* = [x_1^*, \dots, x_p^*], y_r^* = [y_1^*, \dots, y_p^*]\}$, onde para $\hat{\theta}_i$ com orientação de entrada, tem-se $x_r^* = \left(\frac{\hat{\theta}_i}{\hat{\theta}_r^*}\right) x_i$ e para $\hat{\theta}_i$ com orientação de saída, tem-se $y_r^* = \left(\frac{\hat{\theta}_i}{\hat{\theta}_r^*}\right) y_i$, com $r = 1, \dots, p$; IV) Calcule o *bootstrap* estimado $\hat{\theta}_{(r,w)}^*$ resolvendo o problema linear com as restrições DEA usando $\hat{\theta}_{(r,w)}^*$, conforme indicado na Equação 2.1 (Rosano-Peña et al., 2022).

$$\hat{\theta}_{(r,w)}^* = \min \left\{ \hat{\theta} \mid \sum_{r=1}^p \lambda_r x_r^* \leq \hat{\theta} x_o^*, \sum_{r=1}^p \lambda_r y_r^* \geq y_o^*, \sum_{r=1}^p \lambda_r = 1, \lambda_r \geq 0, \forall r \right\} \quad (2.1)$$

Onde x_o^* é a entrada virtual e y_o^* é a saída virtual. λ_r representa os pesos. Em seguida, repita os passos ii, iii e iv W vezes para obter o resultado para cada observação $r = 1, \dots, p$, resultando em um conjunto de estimativas *bootstrap* $\{\hat{\theta}_{r,w}^*, w = 1, \dots, W\}$. Com o número de repetições do bootstrap W . Obtenha o resultado com um intervalo de confiança de 95% para cada observação a partir do conjunto de estimativas (Rosano-Peña et al., 2022). Todos esses valores são calculados usando uma função fornecida pelo pacote *FEAR* no R, com a função *boot.sw98* (Wilson, 2008).

Após aplicar o método computacional *bootstrap-DEA*, os valores ajustados foram comparados para analisar o retorno de escala dos dados. O resultado estimado é $S = 1.010459$, e o valor crítico é $C_\alpha = 0.9350187$. Como resultado, não são observadas diferenças significativas nas fronteiras tecnológicas (T), indicando um comportamento de escala consistente. Isso está alinhado com a teoria econômica, pois a análise examina o Brasil em diferentes períodos. Em teoria, o retorno da relação entrada-saída deve permanecer constante, uma vez que o tamanho do Brasil não muda ao longo do tempo, prevenindo, assim, variações de custo devido ao tamanho. Como resultado, pode-se assumir um retorno constante à escala.

Assim, os valores fornecidos pelo modelo CRS usando o método computacional *bootstrap* com 2000 replicações foram utilizados. Um gráfico *boxplot* é plotado na Figura 2.4 para visualizar os valores com e sem correção.

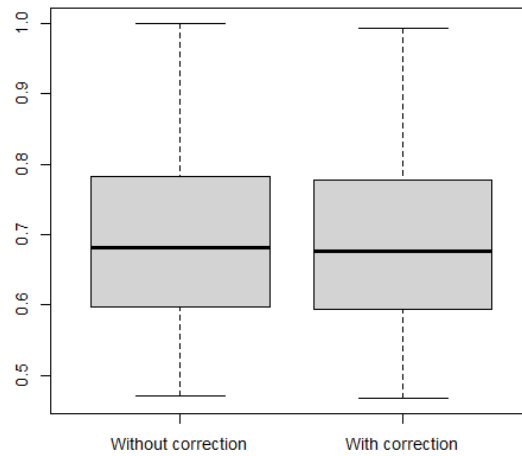


Figura 2.4: Boxplot para o índice de eficiência dos dados sem e com correção.

Visualmente, na Figura 2.4, pequenas mudanças nas pontuações de eficiência podem ser observadas. Portanto, decidiu-se avaliar os resultados usando o teste de *Wilcoxon* para verificar se as diferenças nas pontuações com e sem correção são significativas. A hipótese nula (H_0) afirmava que não havia diferenças significativas nas medianas. O teste retornou um valor de $p < 2.2e^{-16}$, portanto H_0 é rejeitada. O resumo estatístico dos dados é apresentado na Tabela 2.4.

Tabela 2.4: Resumo estatístico dos dados sem e com correção com intervalo de confiança de 95%.

	Mín.	1º Quartil	Mediana	Média	3º Quartil	Máx.
Sem correção	0.4703	0.5971	0.6813	0.6987	0.7823	1.0000
Com correção	0.4672	0.5933	0.6770	0.6942	0.7773	0.9936
Máx. 95%	0.4702	0.5970	0.6813	0.6986	0.7822	0.9999
Mín. 95%	0.4593	0.5832	0.6655	0.6825	0.7641	0.9768

A Tabela 2.4 mostra que as pontuações de eficiência sem correção resultaram em valores de eficiência mais altos (0.6813 e 0.6770). Portanto, a diferença entre os modelos sem e com correção é significativa e deve ser considerada.

2.5 IDENTIFICAÇÃO DOS PARÂMETROS (P,D,Q)

Esta seção explica o funcionamento dos modelos ARIMA e SARIMA e os procedimentos utilizados para realizar a etapa inicial da metodologia de *Box-Jenkins* para identificar os parâmetros (p, d, q). Esta seção refere-se ao passo 3 do algoritmo apresentado na Figura 2.1.

2.5.1 Modelos Autorregressivos Integrados de Médias Móveis (ARIMA) e Sazonais (SARIMA)

Uma série temporal é determinada por uma variável (y_t) com valores quantitativos ao longo de um período (t). Portanto, a análise de séries temporais visa explicar comportamentos futuros com base nos comportamentos passados da variável. É um procedimento lógico baseado em fatos históricos, e é essencial entender como os processos estocásticos funcionam, uma vez que as variáveis estão sujeitas a variáveis aleatórias, que podem surgir de pontos, eventos ou fenômenos observados ao longo do tempo. A previsão de séries temporais envolve várias áreas do conhecimento (multidisciplinares), e sua implementação não é complexa, principalmente porque requer apenas a variável em análise com seus dados históricos (Alsharif et al., 2019). Para isso, existem os modelos autorregressivos (AR), média móvel (MA), autorregressivo de média móvel (ARMA) e autorregressivo integrado de média móvel (ARIMA), que serão descritos detalhadamente na Seção 2.5.1.

O modelo autorregressivo baseia-se no conceito de que os valores (y_t) são explicados pelos (m) valores anteriores dados por ($y_{t-1}, y_{t-2}, y_{t-3}, \dots, y_{t-m}$). Assim, para um modelo autorregressivo de ordem (p) $AR(p)$, ele será dado pela Equação 2.2 (Gujarati and Porter, 2011).

$$y_t = \alpha_1 y_{t-1} + \alpha_2 y_{t-2} + \dots + \alpha_p y_{t-p} + e_t \quad (2.2)$$

Onde μ é a média da variável (y), α_t é a proporção dada pelo período ($t - 1$), e (e_t) representa o erro aleatório dos valores que são não correlacionados com a média (μ), mas causam mudanças devido a razões desconhecidas e incontroláveis nos valores de ($y_t - \mu$), sendo distúrbios aleatórios do período (t). Portanto, os valores envolvidos nos modelos anteriores consideram seus valores atuais e anteriores, demonstrando que o modelo $AR(p)$ é explicado por seus valores e independente de seus regressores. O modelo de média móvel considera os valores do termo de erro (e_t) em termos do presente e do passado, bem como uma constante (c) para explicar (y_t). Portanto, para um processo de média móvel de ordem q $MA(q)$, pode ser expresso pela Equação 2.3 de forma geral (Gujarati and Porter, 2011).

$$y_t = c + \beta_0 e_t + \beta_1 e_{t-1} + \beta_2 e_{t-2} + \dots + \beta_q e_{t-q} \quad (2.3)$$

Portanto, o modelo de média móvel será uma previsão baseada na combinação linear do ruído branco. Isso é viável porque os modelos AR e MA assumem que são baseados em uma série de sistemas lineares com erro aleatório, uma média de zero, um desvio padrão constante e nenhuma autocorrelação. Os modelos autorregressivos e de média móvel, que resultam no modelo ARMA, são os termos autorregressivos e os termos de média móvel dados por $ARMA(p,q)$. Para simplicidade, o modelo $ARMA(1,1)$ é apresentado na equação 2.4 (Gujarati and Porter, 2011).

$$y_t = w + \alpha_1 y_{t-1} + \dots + \alpha_p y_{t-p} + e_t + \beta_0 e_t + \dots + \beta_q e_{t-q} \quad (2.4)$$

Onde w representa um termo constante.

Os modelos AR, MA e ARMA descritos assumem que a série temporal envolvida na análise é (fracamente) estacionária, o que significa uma série com média, variância e autocovariância constantes, não

variando ao longo do tempo, onde as medições dadas pela variável (y_t) terão variações com amplitudes mais ou menos constantes, e seus valores tendem a retornar à sua média. No entanto, muitas séries temporais são consideradas não estacionárias e não exibem esse comportamento contínuo. Uma maneira de lidar com esses casos é por diferenciação (d) da série temporal de acordo com a ordem de diferenciação, geralmente com um valor de (d) sendo 1 ou 2 (Alsharif et al., 2019). Assim, uma série temporal não estacionária de ordem ($d = 1$) deve ser dividida uma vez para se tornar estacionária. Portanto, uma série temporal ARIMA (1,2,1), diferenciada duas vezes ($d = 2$), pode ser analisada pelo modelo ARMA (1,0,1), conforme já apresentado pela equação 2.4 do modelo ARMA (Gujarati and Porter, 2011).

Nos casos em que há um comportamento sazonal da variável (y_t), o modelo autorregressivo integrado de média móvel sazonal, também conhecido como SARIMA, é utilizado. O modelo SARIMA pode ser representado pela Equação 2.5.

$$y_t = c + \sum_{n=1}^p \alpha_n y_{t-n} + \sum_{n=1}^q \beta_n e_{t-n} + e_{t-1} \sum_{n=1}^P \alpha_n y_{t-sn} + \sum_{n=1}^Q \beta_n e_{t-sn} + e_t \quad (2.5)$$

Consiste na inclusão de termos autorregressivos sazonais (P), diferenciação sazonal (D), média móvel sazonal (Q) e o componente sazonal (S).

2.5.2 Transformando Variáveis em Séries Temporais

Para a análise de dados sobre demanda e PIB, os dados são transformados em uma série temporal de 2004 a 2023 usando o logaritmo das variáveis. Essas séries históricas de dados podem ser vistas nas Figuras 2.5 e 2.6, respectivamente. Seus dados respectivos estão resumidos na Tabela 2.5.

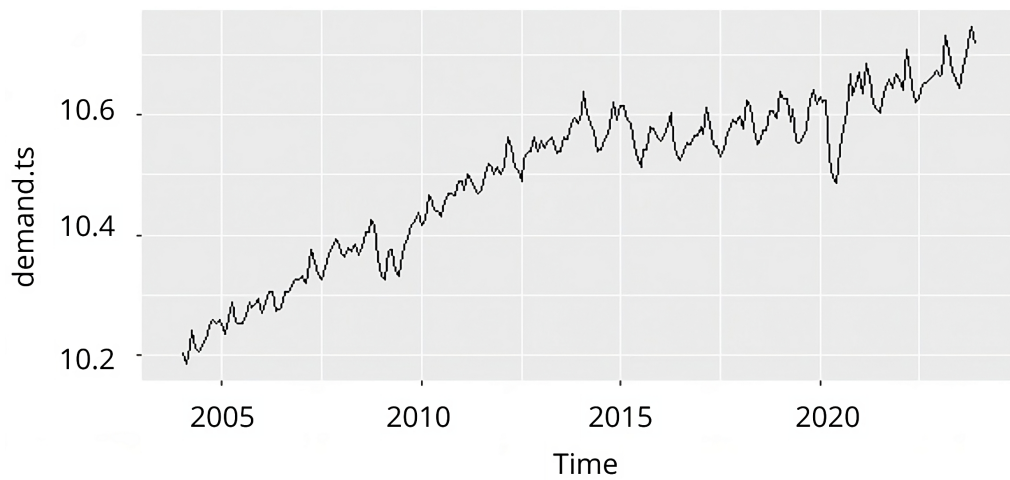


Figura 2.5: Série temporal de demanda de energia

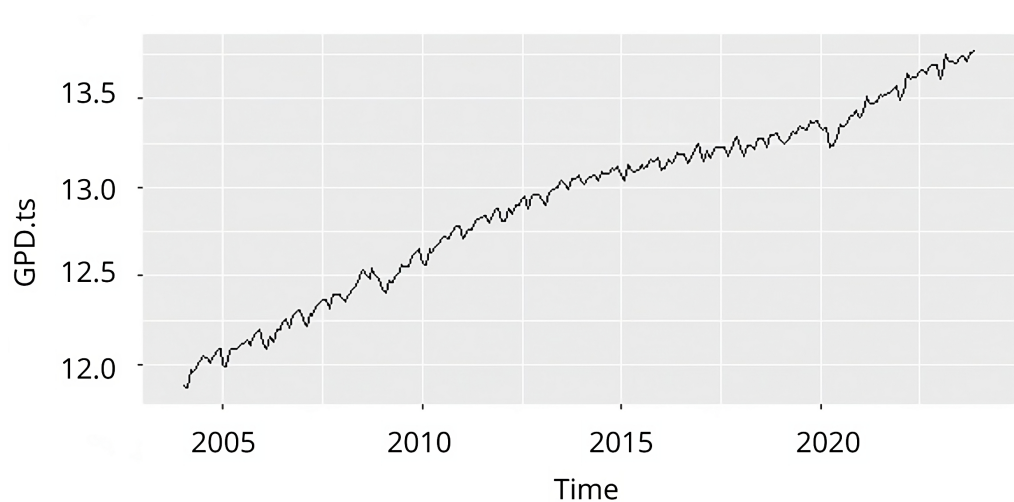


Figura 2.6: Série temporal do PIB

Tabela 2.5: Resumo estatístico para a série temporal de demanda de energia e PIB.

	Mín.	1° Quartil	Mediana	Média	3° Quartil	Máx.
Demanda	10.19	10.38	10.54	10.50	10.60	10.75
PIB	11.87	12.50	13.04	12.92	13.28	13.77

As Figuras 2.5 e 2.6 não demonstram visualmente um comportamento estacionário, um pré-requisito para o uso dos modelos AR, MA ou ARIMA. Esses modelos são baseados no tratamento de dados que têm uma média e variância constantes ao longo do tempo. Portanto, a tendência dos dados em uma análise preliminar tende a ser não estacionária, e a volatilidade dos dados sugere um comportamento sazonal, indicando que o modelo SARIMA pode ser apropriado para analisar previsões de dados. No entanto, essas conclusões devem ser testadas estatisticamente para obter resultados robustos e projeções correspondentes à realidade. A função de autocorrelação é usada para identificar séries temporais não estacionárias e observar a presença de raízes unitárias e tendências nos dados. A função de autocorrelação pode ser visualizada nas Figuras 2.7 e 2.8.

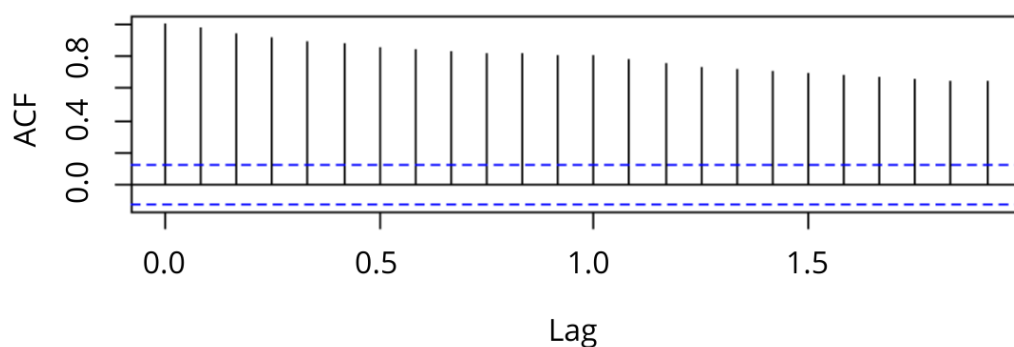


Figura 2.7: Função de autocorrelação total da demanda de energia.

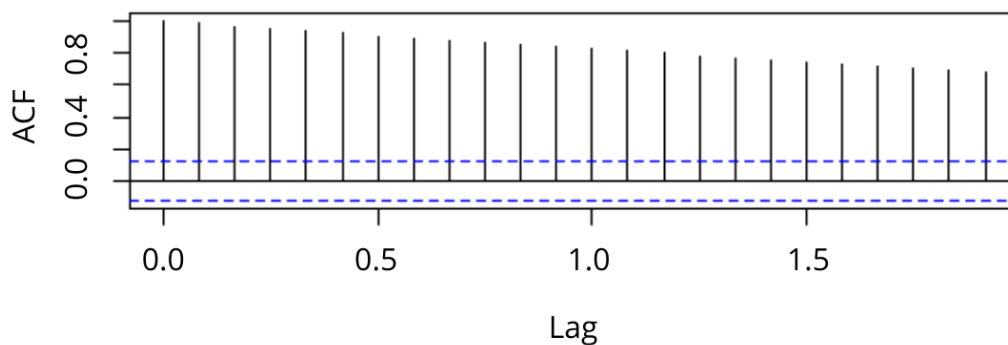


Figura 2.8: Função de autocorrelação total do PIB.

A partir dos gráficos de ACF, tanto a demanda quanto o PIB exibem tendências nos dados e a presença de raízes unitárias, não havendo evidência de ergodicidade nos modelos. No teste ACF, se os lags decaem muito lentamente, isso indica não estacionariedade nos dados do modelo. Para que um modelo seja considerado estacionário, ele deve apresentar média, variância e autocorrelação constantes, mostrando assim que não há mudanças ao longo do tempo. No entanto, os valores podem flutuar acima e abaixo, mas convergir para a mesma média. Um dos testes de estacionariedade é o teste de *Dickey-Fuller* aumentado (ADF), onde a hipótese nula (H_0) indica a presença de uma raiz unitária, sugerindo que o modelo não é estacionário. O teste ADF para demanda resultou em um valor p de 0.2599, o que sugere a aceitação de H_0 , indicando que o mercado não é estacionário. O teste ADF para o PIB retornou um valor p de 0.7107, o que sugere que a série temporal do PIB também não é estacionária.

2.5.3 Teste de Sazonalidade de Séries Temporais

Para visualizar a série sazonal, a série é decomposta para identificar padrões por meio de componentes de tendência (análise do gráfico da série), sazonalidade (análise do gráfico do índice sazonal), ciclos (análise do gráfico do índice sazonal e irregularidades) e resíduos (análise do gráfico das irregularidades). Isso avança o estudo das séries temporais, uma vez que isola um termo específico determinado pela série. A decomposição das séries pode ser vista nas Figuras 2.9 e 2.10.

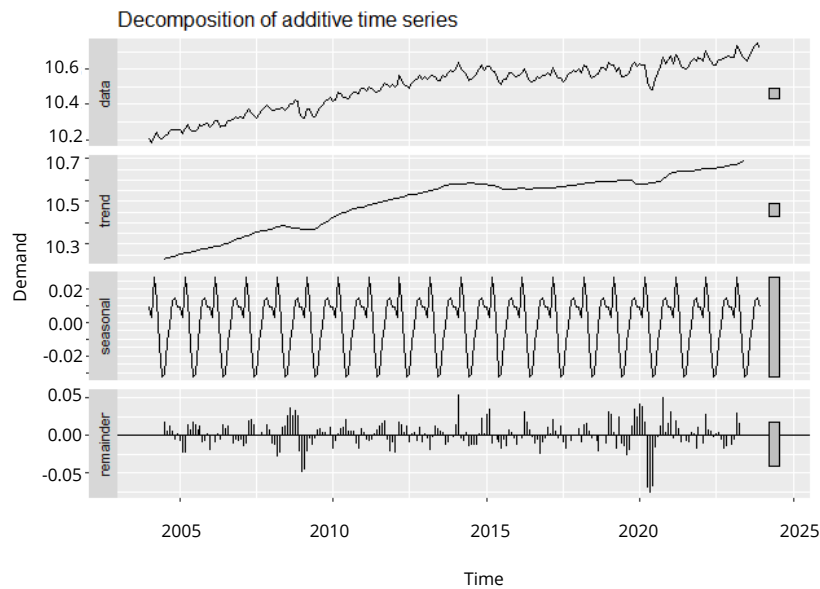


Figura 2.9: Decomposição da série temporal de demanda de energia

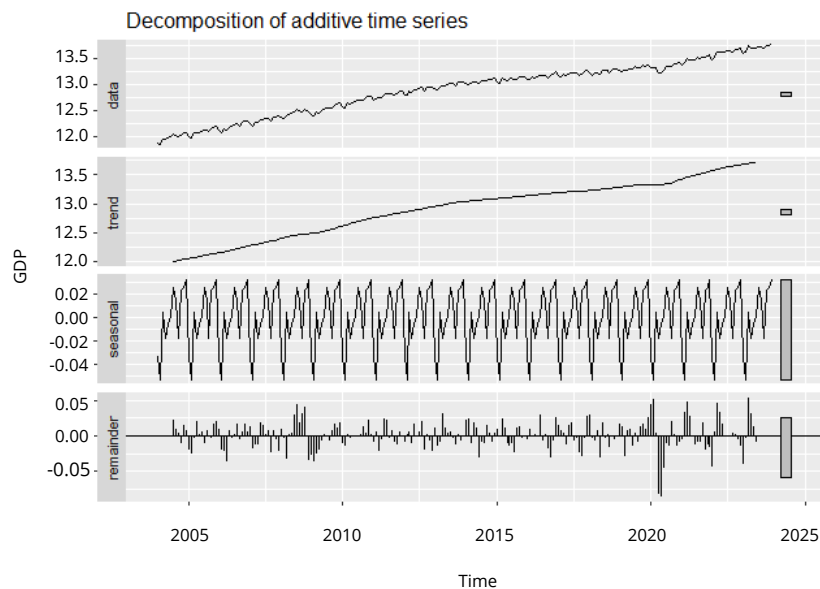


Figura 2.10: Decomposição da série temporal do PIB

Os dados de demanda de energia e PIB mostram uma tendência de crescimento ao longo do tempo, exigindo uma diferença para alcançar a estacionariedade. Observa-se um comportamento de variação sazonal causado por fatores externos com relação ao índice sazonal. Para entender se essa variação sazonal é significativa e deve ser considerada, examinou-se o intervalo de variação do componente sazonal. Em ambos os casos, é significativo, indicando que o componente sazonal está presente na série e é substancial o suficiente para ser considerado na previsão. Não há ciclos irregulares com alternância sistemática ou longos períodos de valores baixos e altos. Em vez disso, há regularidade na distribuição dos dados. Assim, para confirmar as análises feitas usando os gráficos de decomposição, foram criados gráficos de autocorrelação e autocorrelação parcial das séries, como mostrado nas Figuras 2.11 e 2.12.

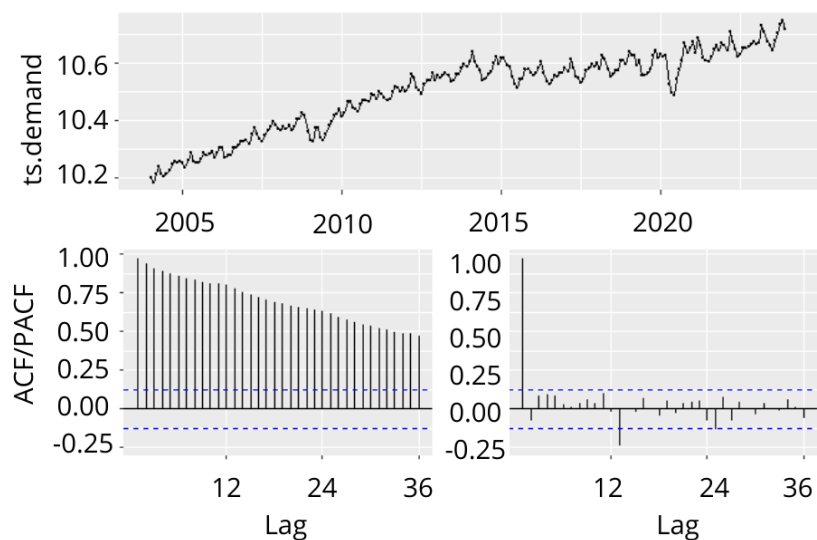


Figura 2.11: FAC e FACP da demanda de energia

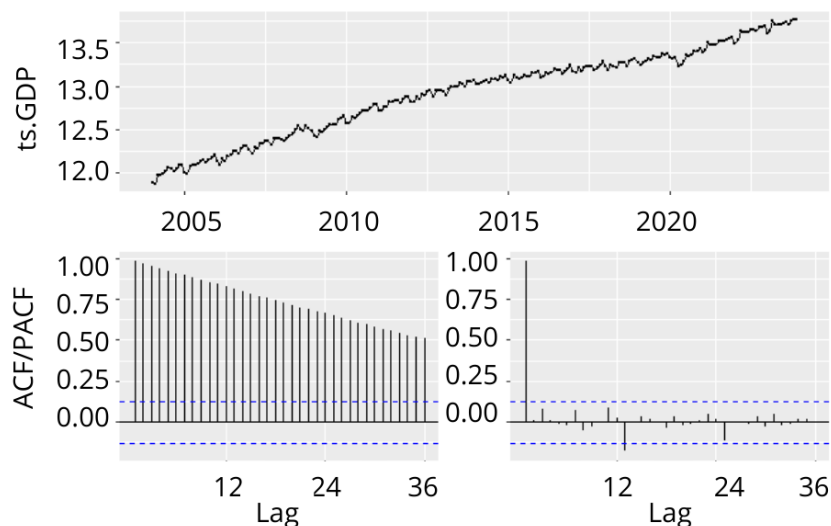


Figura 2.12: FAC e FACP do PIB

O gráfico de ACF também mostra tendências de demanda e PIB ao longo do tempo, e o gráfico de PACF demonstra a ausência de sazonalidade nos dados, já que a maioria dos valores de autocorrelação parcial varia dentro do intervalo representado pela linha azul pontilhada. No entanto, como a série exibe uma tendência e não estacionariedade, foram usadas séries logarítmicas e diferenciadas para reexaminar as funções ACF e PACF. A decomposição da série indicou uma possível sazonalidade, que pode ser visualizada nas Figuras 2.13 e 2.14.

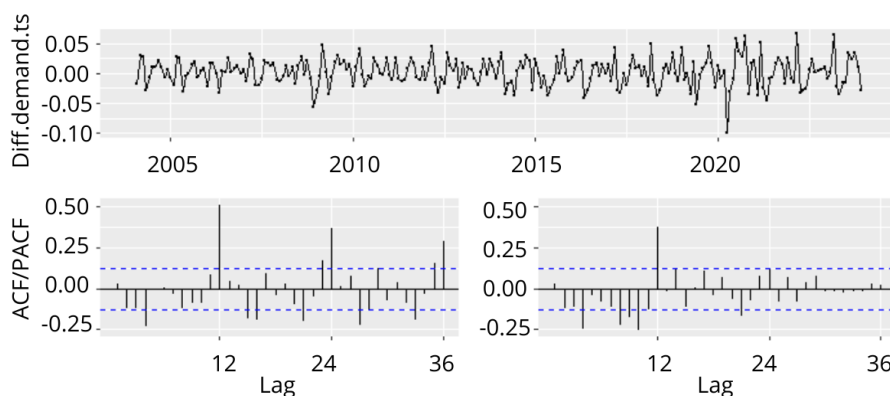


Figura 2.13: FAC e FACP da demanda de energia com uma diferença

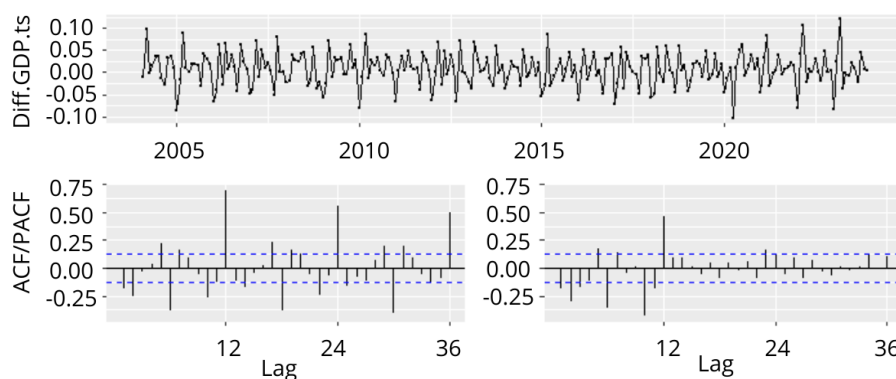


Figura 2.14: FAC e FACP do PIB com uma diferença

Assim, ao remover as tendências dos dados por meio da primeira diferenciação, é possível identificar um comportamento sazonal dado pelas funções ACF e PACF, já que há vários períodos onde os valores de autocorrelação excedem os intervalos dados pela linha azul pontilhada. Extrapolações de intensidade significativa são observadas a cada 12 *lags* tanto na demanda quanto no PIB, com extrapolações de intensidade moderada a cada quatro *lags* na demanda e a cada seis *lags* no PIB. Portanto, o modelo SARIMA é mais adequado para aplicar previsões do que o modelo ARIMA, pois leva em consideração os fatores sazonais presentes nessas séries. Uma alternativa é definir esses parâmetros usando a função `auto.arima`, que identifica o melhor modelo ajustando os parâmetros e usando os índices AICc para avaliar o desempenho do modelo. Esse processo dinâmico envolve várias dependências e pode ajudar na seleção dos parâmetros.

2.5.4 Teste de Parâmetros com Auto.arima

A metodologia de *Box-Jenkins* definirá os parâmetros (p, d, q) . Essa metodologia envolve o ajuste de modelos autorregressivos integrados de média móvel, $ARIMA(p, d, q)$, a um conjunto de dados. Para a construção do modelo, um algoritmo é estruturado no qual a escolha da estrutura do modelo é baseada nos próprios dados.

A função `auto.arima` do R realiza a maior parte desse processo testando diferentes combinações de parâmetros $(p, 1, q)$ e comparando as métricas AICc para determinar o melhor modelo. O algoritmo

`auto.arima` é uma técnica automatizada para identificar o modelo ARIMA mais adequado para uma série temporal. Consiste em três componentes principais. O componente autorregressivo (AR) usa valores passados da série temporal para prever valores futuros, o componente integrado (I) representa o número de diferenças necessárias para tornar a série temporal estacionária removendo tendências, e o componente de média móvel (MA) usa erros de previsão de valores passados para corrigir previsões futuras.

O modelo `auto.arima` oferece várias vantagens, incluindo a automação, que reduz a necessidade de intervenção manual na seleção dos parâmetros do modelo e aumenta a eficiência, economizando tempo e esforço ao automatizar o processo de busca do modelo. Além disso, o modelo proporciona flexibilidade para lidar com dados temporais complexos, incluindo tendências e sazonalidade. No entanto, o algoritmo `auto.arima` tem algumas limitações, como a presença padrão de uma abordagem *stepwise* para melhorar o potencial do modelo, podendo desativar a função *stepwise* para explorar um espaço de modelo mais amplo. Além disso, há a opção de desativar a função de aproximação para melhorar a precisão do modelo, embora isso possa desacelerar o processo de ajuste.

Os modelos definidos como os melhores parâmetros pelo `auto.arima` para demanda foram SARIMA (1,1,1)(0,1,2) e, para o modelo ajustado, SARIMA (1,1,1)(1,1,1). Para o PIB, os parâmetros foram SARIMA (1,1,3)(0,1,2) e, para o modelo ajustado, SARIMA (0,1,2)(2,1,1). Os valores dos testes AIC, AICc e BIC, bem como os parâmetros de erro, podem ser vistos nas Tabelas 2.6 e 2.7.

Tabela 2.6: Valores de AIC, AICc e BIC para diferentes parâmetros usados no `auto.arima`.

Variável	SARIMA	AIC	AICc	BIC
Demanda	(1,1,1)(0,1,2)	-1164.68	-1164.41	-1147.56
	(1,1,1)(1,1,1)	-1164.68	-1164.41	-1147.56
PIB	(1,1,3)(0,1,2)	-1079.2	-1078.69	-1055.23
	(0,1,2)(2,1,1)	-1078.71	-1078.33	-1058.16

Tabela 2.7: Valores de RMSE, MAE, MAPE e MASE para diferentes parâmetros usados no `auto.arima`.

Variável	SARIMA	RMSE	MAE	MAPE	MASE	ACF1
Demanda	(1,1,1)(0,1,2)	0.0174	0.0132	0.1253	0.3928	0.0041
	(1,1,1)(1,1,1)	0.0173	0.0132	0.1250	0.3917	0.0062
PIB	(1,1,3)(0,1,2)	0.0206	0.0154	0.1192	0.1686	-0.0011
	(0,1,2)(2,1,1)	0.0208	0.0157	0.1209	0.1710	-0.0076

A Tabela 2.6 apresenta os valores dos testes AIC, AICc e BIC fornecidos pelos modelos `auto.arima`, com os menores valores descrevendo o melhor modelo. No entanto, os modelos padrão e ajustados para demanda de energia mostraram os mesmos valores, e para os modelos do PIB, o primeiro teve um valor de BIC mais baixo, enquanto o modelo ajustado teve um AIC mais baixo. Seguindo o princípio da parcimônia, foi escolhido o modelo mais simples, portanto, prosseguindo com o modelo não ajustado. Isso é confirmado ao observar os valores dados por RMSE, MAE, MAPE e MASE, que são medidas que avaliam erros. Quanto mais próximo o valor de zero, melhor é o modelo. Portanto, o SARIMA(1, 1, 1) (0, 1, 2) apresentou um desempenho melhor em comparação ao SARIMA(1, 1, 3)(0, 1, 2) para demanda, e ambas as variáveis

tiveram um desempenho melhor com o modelo ajustado. O ajuste dos parâmetros para aplicar o modelo ARIMA mostra uma queda no rendimento. Os valores de RMSE, MAE, MAPE e MASE também são apresentados na Tabela 2.8.

Tabela 2.8: ME, RMSE, MAE, MAPE e MASE para diferentes parâmetros usados no ARIMA.

Variável	ARIMA	RMSE	MAE	MAPE	MASE	ACF1
Demanda	(1,1,1)	0.0231	0.0184	0.1749	0.5478	-0.0404
PIB	(1,1,3)	0.0359	0.0287	0.2235	0.3136	-0.1028
	(0,1,2)	0.0372	0.0302	0.2346	0.3292	-0.1340

Nas Tabelas 2.8, são observados os parâmetros do modelo ARIMA e os modelos ajustados para o PIB.

Comparando o modelo SARIMA (Tabela 2.7) com o modelo ARIMA (Tabela 2.8), observa-se que os modelos SARIMA apresentam um desempenho melhor. Portanto, analisaremos os resíduos do modelo SARIMA para determinar se eles exibem características de ruído branco, o que é essencial para a aplicação das previsões.

2.6 VERIFICAÇÃO DIAGNÓSTICA

Esta seção apresenta os testes estatísticos aplicados para analisar os resíduos. Esta seção refere-se ao passo 4 do algoritmo, que pode ser visto na Figura 2.1.

O teste de *Ljung-Box* foi aplicado aos modelos para diagnosticar os parâmetros usados. Os resultados apresentados mostram que os modelos não falham. Os valores do teste qui-quadrado indicam que não há autocorrelação nos resíduos até o atraso 1. Os valores do p-valor indicam que a hipótese nula não pode ser rejeitada e que não há autocorrelação nos resíduos, mostrando que os resíduos se comportam como ruído branco (Tabela 2.9).

Tabela 2.9: Teste de *Ljung-Box*

SARIMA	Qui-quadrado	df	p_{valor}
(1,1,1)(0,1,2)	0.0040	1	0.9494
(1,1,3)(0,1,2)	0.0003	1	0.9869

Assim, aceita-se que os resíduos se comportam como ruído branco.

2.7 RESULTADOS E DISCUSSÕES

Esta seção apresenta os resultados encontrados pelo estudo, incluindo o índice de eficiência energética e a previsão gerada pelo modelo SARIMA para os próximos anos. Esta seção corresponde ao passo 5 do algoritmo visualizado na Figura 2.1, e apresenta a consolidação dos resultados encontrados nos passos 2,

2.7.1 Análise de eficiência energética

Os escores de eficiência para o conjunto de dados estão mostrados na Tabelas 2.10 e 2.11. Espaços em branco indicam *outliers* para o mês.

Tabela 2.10: Escores de eficiência energética de 2004 a 2013

Data	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013
JAN	-	0.697	0.728	0.573	0.553	0.733	0.583	0.529	0.573	0.556
FEB	-	0.837	0.594	0.581	0.549	0.734	0.548	0.556	0.543	0.547
MAR	0.735	0.591	0.556	0.467	0.582	0.608	0.520	0.552	0.519	0.583
MAY	-	0.771	0.794	0.560	0.571	0.766	0.584	0.614	0.608	0.598
JUN	-	0.840	0.769	0.636	0.658	0.830	0.605	0.650	0.621	0.640
JUL	-	0.865	0.805	0.691	0.656	0.777	0.657	0.637	0.689	0.663
AUG	0.994	0.765	0.674	0.611	0.555	0.664	0.599	0.590	0.624	0.607
SEP	0.744	0.618	0.628	0.516	0.548	0.619	0.563	0.530	0.554	0.589
OCT	0.703	0.695	0.656	0.544	0.539	0.613	0.592	0.572	0.597	0.593
NOV	0.798	0.700	0.628	0.504	0.531	0.606	0.625	0.613	0.560	0.570
DEC	0.757	0.680	0.639	0.532	0.674	0.590	0.570	0.596	0.599	0.597

Tabela 2.11: Escores de eficiência energética de 2014 a 2023

Data	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023
JAN	0.543	0.549	0.677	0.684	0.693	0.654	0.718	0.688	0.795	0.889
FEB	0.489	0.533	0.662	0.692	0.698	0.660	0.725	0.786	0.868	0.908
MAR	0.552	0.629	0.666	0.655	0.652	0.676	0.725	0.760	0.831	0.890
APR	0.588	0.614	0.616	0.679	0.669	0.764	0.858	0.778	0.878	0.899
MAY	0.617	0.677	0.701	0.787	0.720	0.741	0.963	0.872	0.947	0.976
JUN	0.661	0.737	0.805	0.802	0.836	0.848	-	0.897	-	-
JUL	0.693	0.795	0.819	0.853	0.820	0.901	0.940	0.951	-	-
AUG	0.649	0.711	0.793	0.816	0.786	0.863	0.839	0.903	0.975	-
SEP	0.631	0.718	0.712	0.706	0.747	0.821	0.790	0.852	0.948	0.912
OCT	0.602	0.677	0.745	0.720	0.728	0.762	0.704	0.831	0.986	0.887
NOV	0.564	0.672	0.739	0.719	0.730	0.719	0.775	0.889	0.971	0.873
DEC	0.623	0.713	0.776	0.764	0.770	0.782	0.766	0.849	0.951	-

A DMU com o maior escore de eficiência foi em agosto de 2004, enquanto o menor escore de eficiência foi observado em abril de 2006. Em média, o Brasil pode reduzir os gastos com energia em 30,58% mantendo o mesmo valor de PIB, considerando a fronteira produtiva dos anos de 2004 e 2023. O valor médio de 0.6942 para o índice de eficiência é relativamente baixo, indicando que na maioria dos meses,

o Brasil opera muito abaixo do que pode ser considerado eficiente. A DEA destaca a necessidade de mudanças substanciais para minimizar o uso de energia. Também sugere que tais reduções são alcançáveis, já que houve períodos em que os valores de eficiência energética estavam próximos de 1. A tendência dos índices de eficiência de 2004 a 2023 é apresentada na Figura 2.15.

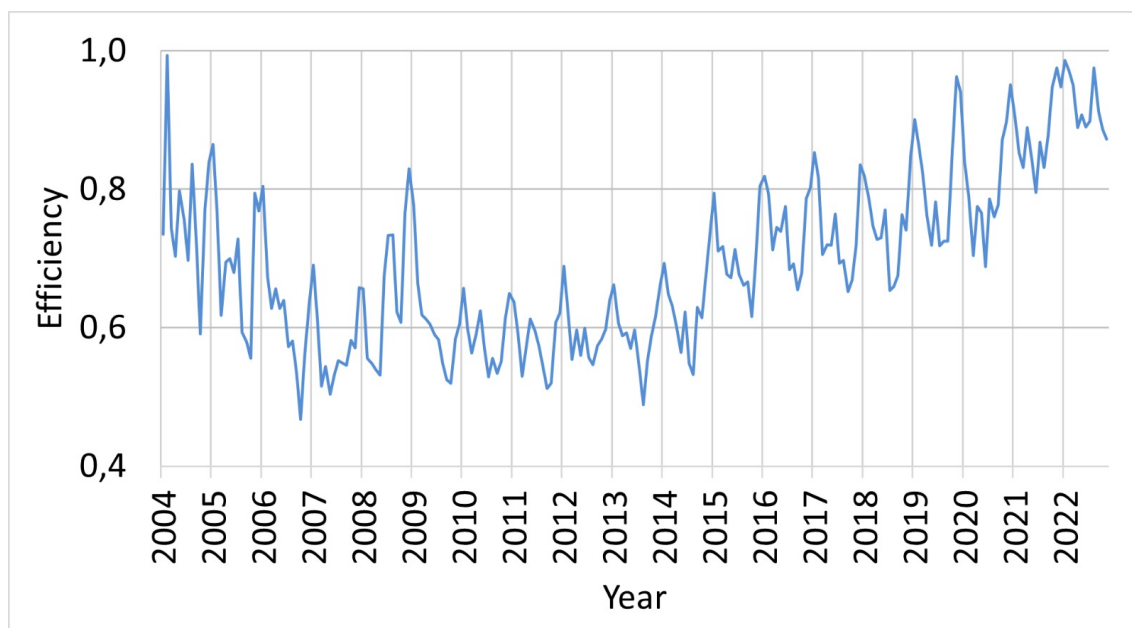


Figura 2.15: Índice de eficiência de 2004 a 2023.

Apesar de o Brasil apresentar um índice de eficiência ineficiente em média, a Figura 2.15 mostra uma queda inicialmente seguida por uma recuperação ao longo do tempo, indicando que por meio dos avanços tecnológicos, as operações estão se tornando mais eficientes. No entanto, há uma variação significativa de mês a mês, demonstrando que há períodos em que as operações se tornam eficientes e períodos em que não se tornam.

2.7.2 Previsão SARIMA

Com os parâmetros (p, d, q) definidos, foram realizadas previsões com o modelo SARIMA. Os modelos foram projetados para um período de 72 meses, tendo como referência o mês de dezembro de 2023. Os resultados com intervalos de confiança de 95% do modelo SARIMA(1,1,1)(0,1,2) para a demanda e do modelo SARIMA(1,1,3)(0,1,2) para o PIB são apresentados, respectivamente, nas Tabelas 2.12 e 2.13.

Tabela 2.12: Previsão da demanda de energia (MWh)

Mês	2024	2025	2026	2027	2028	2029	2030
Jan	45223,43	46442,76	47464,77	48527,26	49613,77	50724,61	51860,32
Feb	44885,43	46019,91	47037,97	48090,97	49167,71	50268,56	51394,06
Mar	46702,57	47694,36	48753,33	49844,79	50960,80	52101,80	53268,34
Apr	45577,83	46511,62	47546,96	48611,44	49699,84	50812,61	51950,29
May	44254,19	45233,10	46241,76	47277,05	48335,57	49417,79	50524,24
Jun	43446,27	44366,31	45356,86	46372,36	47410,62	48472,13	49557,41
Jul	43354,12	44422,76	45415,42	46432,24	47471,84	48534,72	49621,40
Aug	44459,96	45466,72	46483,31	47524,05	48588,10	49675,97	50788,20
Sep	45265,66	46200,20	47233,62	48291,16	49372,38	50477,82	51608,00
Oct	46496,98	47352,26	48411,76	49495,68	50603,87	51736,88	52895,26
Nov	46648,16	47442,90	48504,64	49590,64	50700,96	51836,14	52996,74
Dec	46200,65	47209,60	48266,27	49346,94	50451,80	51581,41	52736,30

Tabela 2.13: Previsão do PIB (milhares de R\$)

Mês	2024	2025	2026	2027	2028	2029	2030
Jan	894130,7	976779,9	1065062,3	1160704,3	1264826,2	1378243,9	1501813,5
Feb	898594,8	975965,6	1063101,4	1158260,6	1262037,6	1375153,7	1498425,1
Mar	972345,1	1045097,5	1139144,1	1241412,9	1352764,2	1474063,1	1606221,8
Apr	949898,1	1024401,4	1115918,2	1215828,3	1324772,7	1443515,7	1572916,8
May	956952,5	1035981,0	1129154,4	1230504,7	1340868,8	1461097,3	1592092,1
Jun	966061,6	1050497,1	1144394,5	1246874,4	1358608,9	1480388,1	1613096,0
Jul	991411,8	1080793,6	1177950,5	1283661,4	1398785,3	1524203,6	1660854,9
Aug	993631,4	1079827,0	1176389,2	1281751,8	1396619,0	1521808,2	1658230,4
Sep	972186,6	1057416,6	1152433,2	1255838,2	1368460,2	1491156,9	1624844,3
Oct	1016733,6	1106198,2	1205156,1	1313110,6	1430794,5	1559049,6	1698811,3
Nov	1024990,2	1115216,2	1215391,6	1324431,4	1443198,9	1572594,3	1713581,9
Dec	1038563,0	1132211,8	1233529,5	1344038,9	1464500,1	1595778,8	1738834,1

Para melhorar a clareza, os dados das séries temporais foram representados graficamente nas Figuras 2.16 e 2.17. Nos gráficos, a região azul indica a previsão feita pelo modelo SARIMA, sendo a linha mais escura a projeção e a área ao redor o intervalo de confiança do modelo.

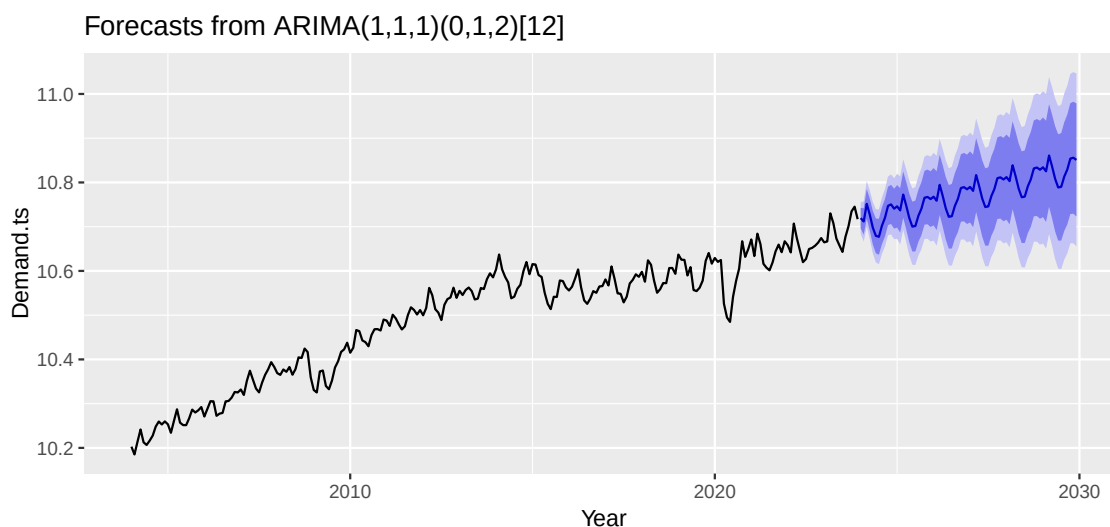


Figura 2.16: Previsão da demanda de energia com SARIMA

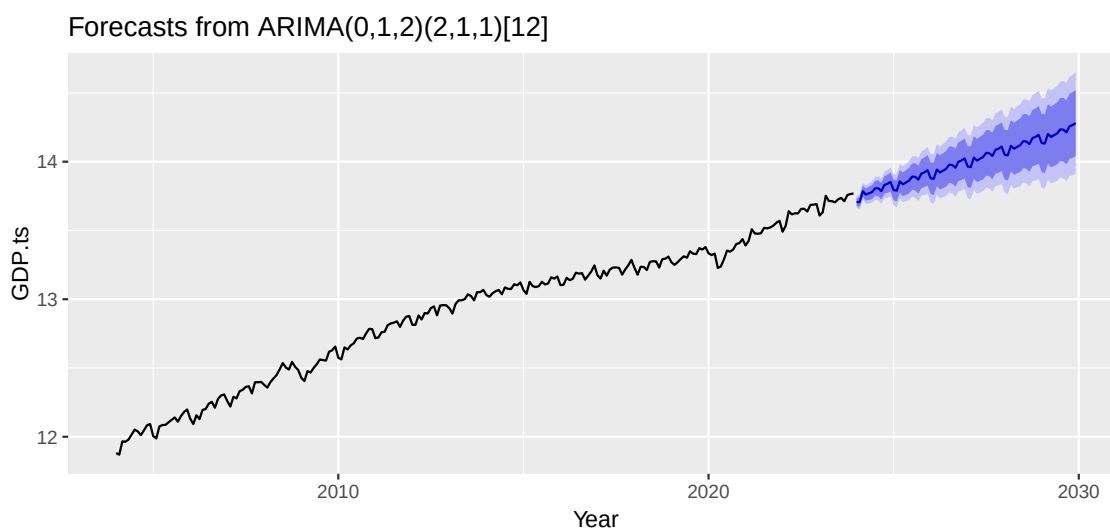


Figura 2.17: Previsão do PIB com SARIMA

Os dados de previsão indicam uma tendência de crescimento consistente para os próximos dois anos. A estimativa é de um crescimento anual médio da demanda de 2,1% entre 2023 e 2030. Esses aumentos estão abaixo da média prevista pela Agência Internacional de Energia (IEA), que projeta um crescimento anual médio de 2,5% até 2026 (International Energy Agency, 2023), mas estão em consonância com as previsões da Empresa de Pesquisa Energética (EPE), coordenadas pelo Ministério de Minas e Energia (MME) (Ministério de Minas e Energia, 2024). Dessa forma, o modelo SARIMA demonstra alta confiabilidade e grande potencial de aplicação, uma vez que permite projeções simples e de fácil replicação para diferentes cenários. A elevação no consumo de energia é atribuída ao aumento da atividade econômica e ao uso residencial de energia, impulsionado pelas altas temperaturas que o Brasil tem enfrentado nos últimos anos. O modelo SARIMA prevê que esse aumento será mais gradual em comparação à matriz energética de 2023. Segundo as projeções, o PIB também segue uma tendência de alta. Em 2024, deve encerrar o ano em torno de 11,7 trilhões de reais, podendo alcançar 19,5 trilhões em 2030.

2.8 DISCUSSÕES FINAIS SOBRE O ESTUDO DE CASO DE PREVISÃO DO SETOR ENERGÉTICO

Este estudo utilizou análise envoltória de dados (DEA) para avaliar a eficiência energética brasileira e aplicou modelos ARIMA/SARIMA para previsão das tendências de variáveis de demanda e PIB, usando dados mensais de 2004 a 2023, por meio de simulações computacionais usando o software R. Contribui com informações e recomendações para melhorar a previsão da demanda de energia, melhorar a gestão eficiente dos recursos no setor energético, avaliar o desempenho dos modelos de previsão e sua relação com o PIB brasileiro, e melhorar as práticas de previsão da demanda de energia, considerando a contribuição dos recursos hidráulicos, eólicos e fontes solares. Avançar no cumprimento do ODS 7 para o desenvolvimento econômico e sustentável da matriz energética brasileira.

O estudo atingiu seus objetivos: i) Utilizou o modelo não paramétrico juntamente com algoritmos dos métodos computacionais *bootstrap-DEA* para analisar o setor energético do país. Considerou a variável demanda como insumo e o PIB como produto da matriz energética. O estudo constatou um comportamento ineficiente da matriz energética ao longo de um período considerável de tempo, indicando uma margem significativa de melhoria para tornar o país mais eficiente. ii) O estudo proporcionou uma compreensão mais aprofundada da demanda de eletricidade, utilizando modelos ARIMA e SARIMA para antecipar o comportamento futuro da matriz energética através de tendências. Esta informação pode ser utilizada pelos decisores do país. iii) O estudo propôs uma metodologia de avaliação da eficiência energética brasileira, que pode auxiliar no desenvolvimento de políticas públicas e ações para garantir o cumprimento do ODS7 pelo país, tornando mais difundido o conhecimento das estratégias energéticas.

O procedimento adotado, que envolve DEA combinado com métodos computacionais como o *bootstrap*, tem se mostrado eficaz na medição da eficiência energética e na avaliação das fronteiras das oportunidades produtivas do país. Com base nas variáveis de demanda energética e PIB, as conclusões de sua eficácia mostram que o Brasil tem se tornado cada vez mais eficiente energeticamente ao longo do tempo devido aos seus avanços tecnológicos. No entanto, ainda há espaço significativo para melhorias para alcançar a eficiência real, uma vez que a redução potencial da margem de lucro nos custos de consumo de energia é de aproximadamente 30,58%. Conforme mostrado na Seção 2.7.1, há períodos de altos e baixos ao longo dos meses em consideração, sugerindo que fatores sazonais impactam a eficiência energética.

As séries temporais tanto da procura como do PIB apresentam um padrão de comportamento semelhante, evidenciado pelos testes de estacionariedade realizados, que confirmaram que ambos necessitam de diferenciação de 1 para atingir a estacionariedade. Além disso, ambas as séries apresentaram um padrão sazonal com valores diferentes ao longo do tempo. Devido a esta sazonalidade, o SARIMA produziu melhores resultados do que o modelo ARIMA para a procura de energia e previsão do PIB. Isto foi observado como resultado das métricas de erro e pelo fato de o SARIMA ter regiões mais estreitas (dentro de intervalos de confiança de 80% e 95%) do que os modelos ARIMA.

As conclusões retratam uma perspectiva mais pessimista do que a que tem sido veiculada nos canais de notícias brasileiros sobre fatores econômicos e consumo de energia, o que sugere o risco de uma desaceleração mais significativa no crescimento econômico do país. Embora o Brasil tenha melhorado, em média, sua eficiência ao longo dos anos, houve uma variação considerável de mês para mês, o que destaca

a necessidade de melhorias adicionais. Ainda há espaço significativo para melhorias, com a introdução de novas tecnologias que podem otimizar a produção para os custos exatos ou manter o valor da produtividade e, ao mesmo tempo, reduzir significativamente os custos de energia. Assim, uma alternativa é investir em tecnologias energéticas sustentáveis para aumentar substancialmente a quantidade de energia renovável na matriz energética do país. O Brasil se tornará cada vez mais eficiente e competitivo se isso puder reduzir os custos de eletricidade para as residências e a indústria. Isto apenas sublinha a importância dos objetivos da agenda 2030 para o setor industrial, particularmente o ODS 7, que procura garantir o acesso a energia fiável, sustentável, renovável e acessível para todos.

3 APLICAÇÃO DA ANÁLISE GARCH PARA PREVISÃO DA VOLATILIDADE DOS PREÇOS DAS AÇÕES DA ELETROBRAS

3.1 CONTEXTUALIZAÇÃO DA VOLATILIDADE DOS PREÇOS DAS AÇÕES DE ENERGIA

A integração de fontes de energia renováveis é uma tendência mundial nos sistemas de produção e distribuição porque podem responder aos desafios energéticos não resolvidos pelas tradicionais centrais elétricas centralizadas, reduzindo as emissões globais e aumentando a oferta de energia sustentável a longo prazo (Ang et al., 2022). Os recursos energéticos distribuídos (DERs), especialmente a energia eólica e a energia solar fotovoltaica, desempenham papéis cada vez mais importantes nas estruturas energéticas (Weschenfelder et al., 2020).

Contudo, o conhecimento dos preços dos ativos das empresas de geração e distribuição de energia elétrica negociados no mercado financeiro é essencial para a alocação de recursos no setor energético num sistema com produção limitada. Além disso, contribui para desencadear ou promover objetivos de desenvolvimento sustentável, especialmente o Objetivo de Desenvolvimento Sustentável 7 (ODS 7) e as suas metas. Vale ressaltar que a energia representa um dos principais custos da indústria para a economia, ou seja, é um fator que influencia o Produto Interno Bruto (PIB), destacando ainda a relevância de garantir uma gestão eficiente neste setor (Yasmeen et al., 2022).

Desta forma, a compreensão dos Objetivos de Desenvolvimento Sustentável 7 abrange vários objetivos, incluindo garantir o acesso universal a serviços energéticos modernos, acessíveis e viáveis (7.1), aumentar a quota de energias renováveis na matriz energética global (7.2), duplicar a taxa global de melhoria na eficiência energética (7.3), fortalecendo a cooperação internacional para pesquisa e tecnologia de energia limpa (7.4) e promovendo a expansão da infraestrutura para serviços de energia sustentável nos países em desenvolvimento (7.5) (Shyu, 2021; Caldeira et al., 2016) .

Compreender o preço das ações da Companhia Brasileira de Energia Elétrica (Eletrobras), bem como antecipar o comportamento futuro da energia e seus preços, são elementos fundamentais para que empresas públicas, formuladores de políticas e *stakeholders* tomem decisões informadas e alinhadas aos objetivos relacionados ao desenvolvimento sustentável e investimentos futuros na geração e distribuição deste insumo (Alizadeh et al., 2020).

A compreensão das tendências, tais como os factores socioeconômicos e os avanços tecnológicos, influenciam os padrões dos preços das ações durante longos períodos, pelo que é necessário aumentar o acesso a serviços energéticos modernos (7.1) e melhorar a eficiência energética (7.3). As sazonalidades de curto prazo, incluindo variações nas condições meteorológicas e nos ciclos de produção industrial, também têm impacto nas flutuações na procura de energia. Estas são considerações cruciais para alcançar as metas do ODS 7, particularmente sobre a integração das energias renováveis (7.2) e o desenvolvimento

de infraestruturas (7.5). Sublinham a relevância de mecanismos robustos de geração de conhecimento e de previsão para apoiar a cooperação internacional em investigação e tecnologia de energia limpa (7.4) (Yasmeen et al., 2022).

Ao selecionar técnicas de previsão apropriadas, os serviços públicos, os serviços privados e os investigadores podem contribuir para os esforços globais para melhorar o acesso à energia limpa e promover infraestruturas energéticas sustentáveis (7.A, 7.B) (Yasmeen et al., 2022), daí a importância de compreender o futuro preços e tendências na produção e distribuição de eletricidade para apoiar a formulação de estratégias para futuros investimentos e desenvolvimento da indústria e aplicação do PIB.

As técnicas de previsão têm gerado interesse significativo na literatura, pois são utilizadas em diversos contextos, com sua aplicabilidade prática, eficácia em diferentes setores industriais, adequação a diferentes tamanhos de organizações, horizontes de tempo e níveis de precisão. Tendo isto em mente, o estudo procura oferecer recomendações para melhorar a análise da volatilidade no setor de produção de eletricidade, com vista a uma análise mais eficiente dos recursos que impactam diretamente a economia. Foram considerados dados diários sobre os preços das ações da principal empresa de geração de energia elétrica do Brasil, a Eletrobras, usando dados diários entre janeiro de 2007 e maio de 2024. Os métodos de previsão estatística mais comumente usados para modelar a volatilidade são modelos de heterocedasticidade condicional.

Introduzido por Engle (1982), o modelo de heterocedasticidade condicional autorregressiva (ARCH), inicialmente para modelar a volatilidade dos preços das ações da Eletrobras (ELET3). Observando que os modelos ARCH ajustados eram muito complexos, ele generalizou o modelo ARCH, criando o modelo de heterocedasticidade condicional autorregressiva generalizada (GARCH) mais parcimonioso. A ideia básica desta classe de modelos é que os retornos não são autocorrelacionados serialmente, mas a volatilidade depende dos retornos passados de uma função quadrática.

Contudo, uma dificuldade na utilização destes modelos é encontrar uma distribuição de probabilidade apropriada para os resíduos. As séries de retorno geralmente apresentam caudas pesadas, com excesso de curtose, o que afasta suas distribuições da distribuição normal padrão (Engle and Patton, 2001). É, portanto, importante que a especificação do modelo incorpore esta característica. Por esta razão, uma alternativa amplamente utilizada é propor uma distribuição *Student t* para os resíduos, ou alguma distribuição particular para modelos GARCH, como a distribuição de erro generalizada (GED) (Nelson, 1991), ou a distribuição assimétrica *t* (Fernández and Steel, 1998). Outro problema reside na avaliação da previsão.

Em geral, a relevância prática de um modelo em previsão é feita comparando os resultados obtidos com as realizações do processo (Reis et al., 2013). Assim, a volatilidade não é diretamente observável, tornando irrelevante a abordagem clássica de avaliação de previsões. Os avanços computacionais e a recente maior disponibilidade de dados de alta frequência (HFD) tornaram possível o desenvolvimento desta medida (Reis et al., 2013). No entanto, os HFD raramente estão disponíveis gratuitamente ou durante longos períodos. Além disso, estão sujeitos a erros de digitação ou falta de sincronização. A vantagem de usar modelos GARCH é que eles modelam a volatilidade com base nos retornos diários.

A falta de dados é uma realidade em muitas áreas (Kong et al., 2022). Essa limitação, falta de clareza e disponibilidade de dados é um problema no caso dos preços diários de energia elétrica no Brasil, tornando um desafio planejar estratégias que possam analisar tendências para o país reduzir custos e desenvolver de forma sustentável a matriz energética brasileira.

As contribuições deste trabalho são significativas no mercado financeiro e nos setores energéticos brasileiro e mundial. Uma contribuição é o avanço das discussões sobre a análise de risco do setor energético no mercado financeiro e a avaliação de padrões e comportamento das ações da Eletrobras. Com essas informações, a abordagem também pode apoiar o processo de tomada de decisão e a elaboração de políticas públicas no setor energético, direcionando investimentos e contribuindo para um futuro energético sustentável.

Além disso, este estudo pretende contribuir para uma melhor compreensão dos preços de energia, considerando que as bases de dados diárias são inexistentes, e quando estão disponíveis em meses ou anos, não há padronização de valores ou períodos históricos curtos, como no caso do IBGE ¹, ou ANEEL ², devido à falta de clareza sobre como os preços são definidos, pois são influenciados por diversas variáveis exógenas (tarifas, impostos, demanda, geração, capacidade instalada, condições climáticas, entre outros), é difícil pensar em soluções que envolvam o mercado de energia, por isso estruturar uma metodologia que permita previsões de preços de energia é de suma importância para a indústria.

A novidade deste trabalho reside no fato de não terem sido encontrados outros estudos que utilizem uma abordagem GARCH para analisar as tendências do ELET3 para medir as tendências diárias dos preços de energia elétrica na indústria brasileira. Assim, uma nova metodologia de pesquisa é empregada para avaliar este setor, auxiliar na formulação de políticas públicas e fornecer uma base para a criação de novas tecnologias adequadas.

O capítulo está estruturado como segue. A seção 3.2 revisa a literatura sobre previsão de séries temporais, os modelos utilizados (econometria tradicional, inteligência artificial e híbrido) e sua aplicabilidade no setor energético. A seção 3.3 apresenta o material e a metodologia utilizada, descrevendo a modelagem matemática de ARIMA e GARCH. A seção 3.7 contém os resultados encontrados e discussões sobre a análise de padrões e testes estatísticos nas séries temporais referentes aos preços das ações da Eletrobras negociadas na B3, principal bolsa de valores do Brasil. A seção compara a precisão e o desempenho temporal dos diferentes modelos, apresentando valores de previsão para até dez anos. A seção 3.8 conclui o capítulo com os principais resultados e contribuições do estudo.

3.2 MODELOS DE VOLATILIDADE E TRABALHOS RELACIONADOS

Compreender o comportamento de uma empresa de geração de energia elétrica ao longo do tempo ajuda a verificar a importância de traçar um plano de geração e distribuição desse insumo, destacando a necessidade de alinhamento de longo prazo das metas de sustentabilidade ecológica, social, de mercado e econômica. Ao combinar a previsão energética com o planejamento energético sustentável, as sociedades podem progredir em direção a um futuro que seja simultaneamente sustentável e equitativo. Assim, compreender e prever o comportamento dos preços das empresas geradoras de energia é fundamental para o planejamento estratégico energético, permitindo que as partes interessadas tomem decisões informadas relativamente à atribuição de recursos, ao desenvolvimento de infraestruturas, à formulação de políticas e

¹<https://sidra.ibge.gov.br/tabela/7060>

²<https://www.ibge.gov.br/estatisticas/economicas/precos-e-custos/9256-amplio.html?=&t=resultados>

[índice-nacional-de-precos-ao-consumidor-](https://www.aneel.gov.br/indicadores/indice-nacional-de-precos-ao-consumidor)

até ao impacto na economia e no Produto Interno Bruto. Os métodos de previsão de energia podem ser amplamente categorizados em abordagens qualitativas e quantitativas.

Conforme documentado por Arnob et al. (2023b), dentro dos métodos quantitativos, os modelos de séries temporais são amplamente utilizados para a previsão da procura de energia porque capturam os padrões temporais e tendências inerentes aos dados de consumo de energia. Vários estudos compararam o desempenho de modelos de séries temporais em diferentes contextos e para diversas aplicações. Rafayal et al. (2022) conduziu uma revisão abrangente dos métodos de previsão, incluindo abordagens clássicas e modernas. Da mesma forma, Riady and Apriani (2023) usou uma rede neural autorregressiva (RNA) não linear para prever a demanda e a variação do preço da energia na próxima década com base em um conjunto de dados publicamente disponível para o consumo global de energia.

Várias aplicações recentes e estudos de caso na literatura têm utilizado modelos de previsão para prever os preços dos participantes no mercado de energia. Tem havido vários estudos comparando vários modelos econométricos tradicionais de previsão com modelos de Inteligência Artificial (IA), seja para prever o preço da eletricidade ou de outras fontes de energia, como o petróleo, ou o custo da produção de carbono a partir destes processos (Markovics and Mayer, 2022; Mardani et al., 2020; Ning et al., 2022).

As pesquisas encontradas na literatura podem ser divididas em três vertentes. i) estudos que utilizam modelos econométricos tradicionais aplicados ou comparados com modelos de IA, caso em que os modelos econométricos apresentam melhor desempenho (Narajewski and Ziel, 2020), ii) estudos de previsão utilizando Inteligência Artificial aplicada ou comparada com modelos econométricos tradicionais, que nesse caso, os modelos de IA apresentam melhor desempenho (Spiliotis et al., 2022), e iii) estudos que utilizam métodos econométricos tradicionais juntamente com modelos de IA para melhorar o desempenho dos modelos de previsão (Makridakis et al., 2023).

3.2.1 Modelos tradicionais de previsão econométrica

Existe um estudo de aplicação para previsão de preços de eletricidade para o dia seguinte nos mercados alemão e finlandês, usando uma abordagem com *Extreme Learning Machine* e *Bootstrap* para construir intervalos de confiança para previsões de preços (Loizidis et al., 2024). A análise dos modelos de previsão do preço do petróleo, utilizando o modelo GARCH de duas fases para melhorar o desempenho da previsão, demonstra que é um modelo eficaz para prever preços neste tipo de mercado (Zhu et al., 2021).

Previsão de preços de energia no mercado colombiano comparando os modelos ARIMA e GARCH, mostrando que este último tem melhor desempenho nas previsões de curto prazo (Ramírez et al., 2021). Há também um estudo que utiliza dados de preços de ações de gás da *State Gas Company* para prever os preços das Empresas Nacionais de Gás da Indonésia, usando o modelo GARCH de 2019 a 2020, para ajustar o melhor modelo para prever 186 dias antes e 176 dias após a Covid-19, obtendo o melhor modelo como sendo AR(1) - GARCH(1, 1) com erro inferior a 0,0015 (Suripto and Supriyanto, 2021).

A previsão dos preços da energia na Colômbia também utiliza modelos econométricos; neste estudo, ARIMA e GARCH foram utilizados para prever o mercado de energia elétrica, obtendo os parâmetros ARMA (14, 10) e GARCH (1, 1) ideais para desenhar estratégias de investimento no setor (Arango et al., 2020). A previsão dos preços das opções de carbono, utilizando uma combinação de modelos GARCH

e o movimento browniano fracionário da China (FBM), neste estudo, a abordagem combinada revelou-se satisfatória (Liu and Huang, 2021).

Previsão dos preços da eletricidade fora de ponta e de ponta no cenário germânico, estimando a média e os resíduos, calculando os mínimos quadrados ordinários (MQO) destas séries temporais (Muniain and Ziel, 2020). Aplicando modelos de previsão ao preço das ações de entidades de energia no mercado indonésio *Indika Energy Tbk*, os resultados apresentados indicaram que os modelos tradicionais, dados pelos parâmetros $AR(4)$ -GARCH(1, 1), obteve resultados satisfatórios (Azhar et al., 2020).

3.2.2 Modelos de previsão de Inteligência Artificial

Aplicabilidade na previsão de preços diários utilizando modelos de regressão e séries temporais para o período de 2007 a 2024, utilizando métodos convencionais (ARIMA, GARCH e OLS) em comparação com modelos de Redes Neurais; neste estudo, as redes neurais obtiveram melhores resultados que os modelos econométricos tradicionais (Xu et al., 2023). Prever o preço futuro das emissões de carbono para o dia seguinte usando o modelo de aprendizado de máquina *Random Forest* com alguns ajustes adicionais, obtendo melhores resultados que os modelos ARIMA, GARCH ou híbridos (Pawłowski, 2021).

3.2.3 Modelos híbridos de previsão econométrica e de inteligência artificial

Análise da previsão dos preços da energia em Portugal entre 2016 e 2021 utilizando o modelo *Markov-Switching GARCH Models* para eliminar os efeitos da volatilidade dos preços da eletricidade (Entezari and Fuinhas, 2024). Aplicação à previsão diária de preços de eletricidade no mercado da Nova Zelândia, utilizando o modelo GARCH e SV, o estudo faz uma comparação com outros modelos de aprendizado de máquina, incluindo LSTM, GRU, XGBoost, LEAR e um DNN de quatro camadas, demonstrando que o GARCH e o modelo SV supera os modelos de IA para previsão de preços neste estudo de caso específico (Kapoor and Wichitaksorn, 2023).

Previsão do preço do mercado de eletricidade eólica *Nord Pool* com base em técnicas de aprendizado de máquina, a saber, *Support Vector Regression (SVR)*, *Generalized Autoregressive Conditional Heteroscedastic (GARCH)*, *Autoregressive Integrated Moving Average (ARIMA)*, *Multiple Linear Regression (MLR)* e um híbrido de SVR-GARCH (Vijaya Chandrakala and Rajan, 2022). Uma abordagem para a previsão do preço do carbono usando uma adaptação do modelo GARCH, chamada VMD-GARCH/LSTM-LSTM, combinando o algoritmo de decomposição de modo variacional (VMD), a rede de memória longa de curto prazo (LSTM) e o modelo GARCH no estudo de caso, os resultados mostraram que os modelos econométricos tradicionais em conjunto com a Inteligência Artificial tiveram melhor desempenho (Huang et al., 2021).

A previsão do comportamento dos preços de energia tem sido uma área importante de pesquisa, com vários métodos propostos para uma previsão precisa. Entre eles estão as redes neurais autorregressivas não lineares, que podem lidar com problemas estatísticos, empíricos e teóricos de forma eficaz, conforme apresentado por (Al-Haija et al., 2023). As redes neurais convolucionais (CNNs) e os campos aleatórios condicionais (CRFs) também foram usados para prever o consumo de energia, alcançando alta precisão,

de acordo com (Thangavel and Govindaraj, 2022). A previsão do preço das ações no setor de eletricidade também é relevante, dada a possível mitigação de risco dos contratos futuros (Gundu and Simon, 2021).

Apesar da ampla aplicabilidade dos modelos de previsão GARCH, associados a vários outros modelos e, em especial, aos modelos de IA para previsão de preços de energia, não foi encontrado nenhum estudo relevante que utilize preços de ações aplicados com modelos Heteroscedásticos Condicionais Autorregressivos Generalizados no Brasil para fins de comparação com os resultados deste estudo. Dessa forma, o regime de níveis de preços e volatilidade ilustra a necessidade de mitigar riscos de forma eficiente, sendo os contratos futuros e de opções brasileiros e internacionais classificados como mecanismos potenciais. Assim, a correta previsão da volatilidade é essencial, pois permite a elaboração de um conjunto de estratégias, capturando momentos de grande incerteza no mercado. Diante dessas perspectivas, este estudo tem como objetivo modelar a volatilidade da cotação diária das ações da Eletrobras - ELET3.SA utilizando modelos da família GARCH.

3.2.4 Discussão sobre os pontos fortes e as limitações da avaliação da volatilidade

Uma série temporal representa valores quantitativos que seguem uma ordem progressiva no tempo. É essencial entender os processos estocásticos para estudar as séries temporais, um conjunto de variáveis aleatórias geralmente usadas para estudar a evolução dos fenômenos observados ao longo do tempo. A variável é observada em pontos de tempo discretos, geralmente no centro, e a análise desse comportamento temporal envolve a descrição do processo ou fenômeno que gera a sequência, explicando assim o passado. Um procedimento lógico para fazer previsões é usar esses dados históricos.

Usando dados financeiros históricos, temos modelos de volatilidade, que são especialmente adequados para modelar o comportamento de séries financeiras que têm pouca correlação nas observações dos retornos, mas se correlacionam com os quadrados dos retornos. As séries temporais financeiras têm características específicas, que correspondem aos fatos estilizados dos retornos financeiros: imprevisibilidade na média condicional e média incondicional próxima de zero, correlação pequena ou ausente na média; agrupamento de volatilidade, alternância de períodos de alta e baixa volatilidade que são agrupados e intercalados ao longo do tempo; assimetria, maior volatilidade quando os preços dos ativos estão caindo do que quando os preços estão subindo; a densidade ajustada aos retornos é simétrica e com excesso de curtose, maior probabilidade de valores extremos em comparação com a distribuição normal Tsay (2005); Enders (2004)). Os modelos de volatilidade descrevem essas características dos retornos financeiros.

Inicialmente, a pesquisa mostra a presença do modelo ARIMA, que se baseia no ajuste dos valores observados para reduzir a diferença entre os valores produzidos no modelo e os valores observados para próximo de zero. Para construir modelos ARIMA, *Box-Jenkins* (Gujarati and Porter, 2011) sugeriu as seguintes etapas iterativas; de acordo com *Bayer*, o modelo ARIMA (p, d, q) consiste na combinação de um modelo autorregressivo (AR) de ordem "p" em uma série diferenciada "d" vezes (número de diferenças necessárias para tornar a série estacionária) e um modelo de média móvel (MA) de ordem "q". Assim, ele tem a seguinte forma: é ruído branco, $f(B)$ e $q(B)$ são o polinômio autorregressivo e o polinômio de média móvel, respectivamente. O modelo ARIMA sazonal é apresentado da seguinte forma: $ARIMA(p, d, q)(P, D, Q)$, em que os parâmetros (P, D, Q) são os equivalentes sazonais de (p, d, q) .

Outro modelo pesquisado refere-se aos modelos tradicionais de volatilidade da família GARCH, cuja abordagem de volatilidade determinística tem uma configuração que descreve a volatilidade com parâmetros fixos. Essa classe inclui os modelos da família GARCH que, em geral, apresentam alta persistência e baixa capacidade de previsão. Nos modelos de volatilidade estocástica, a equação que descreve a volatilidade é estocástica. Essa classe contém modelos de volatilidade estocástica e modelos de volatilidade com mudança e regime (Hamilton and Susmel, 1994).

Os modelos de volatilidade são particularmente adequados para modelar o comportamento de séries financeiras que têm pouca correlação nas observações dos retornos, mas a correlação nos quadrados dos retornos. As séries temporais financeiras têm características que outras séries temporais não apresentam, correspondendo aos fatos estilizados dos retornos financeiros: imprevisibilidade na média condicional e média incondicional próxima de zero, correlação pequena ou ausente na média; agrupamento de volatilidade, alternância de períodos de alta e baixa volatilidade que são agrupados e intercalados no tempo; assimetria, maior volatilidade quando os preços dos ativos estão caindo do que quando os preços estão subindo; a densidade ajustada aos retornos é simétrica e com excesso de curtose, maior probabilidade de valores extremos em comparação com a distribuição normal (Gujarati and Porter, 2011).

Estudos mostram que os modelos GARCH têm alta persistência quando há uma quebra estrutural na variância incondicional ou condicional. Se considerarmos a mudança no nível de volatilidade, controlando com uma variável *dummy* na equação de volatilidade do modelo GARCH, podemos reduzir a persistência. Assim, embora o modelo descreva as características dos retornos dos ativos, suas principais limitações na presença de quebras estruturais são a alta persistência e o baixo desempenho preditivo (Hamilton and Susmel, 1994). Nesse caso, um choque nos retornos tem um efeito duradouro sobre a volatilidade. Isso leva a previsões limitadas, especialmente para períodos de análise após a mudança estrutural, uma vez que a persistência estimada resulta em uma estimativa de volatilidade que é geralmente maior do que a real. Os modelos da família GARCH geralmente apresentam uma persistência muito alta para a volatilidade, especialmente os modelos GARCH. A persistência estimada é geralmente próxima à unidade, sugerindo os modelos GARCH integrados ao IGARCH, que restringem a soma dos coeficientes ARCH e GARCH a ser igual à unidade, o que implica que a variância incondicional é infinita.

Finalmente, a dificuldade crítica com a modelagem é que a volatilidade do mercado não é diretamente observável - ao contrário dos preços de mercado, ela é uma variável latente. A volatilidade deve ser inferida por meio da observação da movimentação dos preços de mercado. Se os preços flutuam muito, sabe-se que a volatilidade é alta, mas não se pode determinar com precisão o quão alta é, e uma das razões é que não se pode dizer se um grande choque nos preços é transitório ou permanente. A natureza latente da volatilidade significa que ela deve ser prevista por um modelo estatístico, o que, inevitavelmente, envolve a elaboração de suposições sólidas. Assim, os modelos estatísticos que melhor preveem a volatilidade da série de retornos diários dos preços dos produtos analisados são os modelos da família ARCH, que são compostos por ARCH e GARCH (Gujarati and Porter, 2011).

3.3 MATERIAIS E MÉTODOS

Esta seção descreve a metodologia utilizada para aplicar o modelo de Heterocedasticidade Condicional Autoregressiva Generalizada (GARCH) na análise da volatilidade dos preços diários das ações da ELET3. A análise deste estudo é conduzida inteiramente utilizando o software R. A Figura 3.1 apresenta um fluxograma que resume a metodologia adotada pela pesquisa.

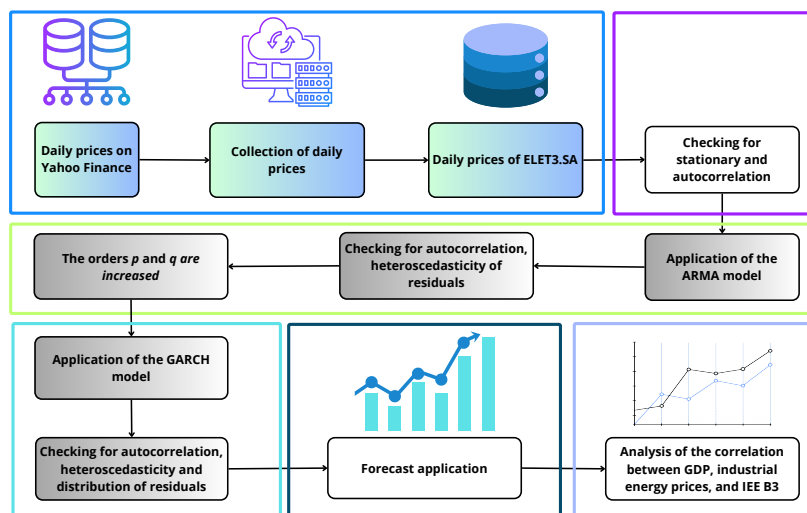


Figura 3.1: Estrutura da pesquisa desenvolvida

3.3.1 Modelo Autorregressivo Generalizado de Heterocedasticidade Condicional (GARCH)

Em se tratando de preços de ações de empresas geradoras de energia elétrica, há um comportamento volátil (períodos de grandes oscilações e depois um comportamento mais estável, seguido novamente por períodos de volatilidade significativa e assim por diante), pois as variáveis estão sujeitas a comportamentos influenciados por fatores econômicos, como o Balanço Energético de 2023 (, EPE), aumentos de ações, impostos, entre outros fatores, que podem causar variação no preço dos ativos de energia elétrica. Assim, uma das primeiras abordagens para o cálculo da volatilidade foi o modelo *Black-Scholes*, criado para auxiliar na precificação de opções europeias. Embora relevante, por ser considerado o ponto de partida para a mensuração da volatilidade, esse modelo apresenta uma desvantagem significativa na avaliação da variância condicional constante ao longo do tempo.

Considerando a variância condicional, sugere-se que os preços sigam uma distribuição normal. Entretanto, a literatura reconhece que os preços raramente são distribuídos de forma normal (Gujarati and Porter, 2011).

Assim, a instabilidade econômica não permite que a volatilidade seja considerada constante ao longo do tempo na maioria das situações. Por esse motivo, uma das abordagens mais comuns atualmente é calcular a volatilidade a partir de retornos diários usando modelos GARCH, propostos por Bollerslev (1986). Esses modelos consideram a volatilidade uma medida não constante ao longo do tempo e permitem flexibilidade na escolha de distribuições de probabilidade que melhor representem o conjunto de dados estudado. Sendo

uma série de retornos, a média condicional e a variância condicional do processo são, respectivamente, dadas pela equação 3.1, sendo que o modelo mais simples, de ordem (p,q) , é fornecido por GARCH(1, 1) (Gujarati and Porter, 2011):

$$\sigma_t^2 = \alpha_0 + \alpha_1 \mu_{t-1}^2 + \alpha_2 \sigma_{t-1}^2 \quad (3.1)$$

Onde (α) representa o parâmetro GARCH, (μ_t) representa o resíduo estimado da regressão (erro quadrático do período t) e (σ_t^2) representa a variância condicional do período t . Os parâmetros do modelo GARCH(p,q) são, respectivamente, os termos de defasagem necessários do termo de erro quadrático (p) e os termos de variação condicional referentes às defasagens (q). Como o modelo GARCH é uma variação do modelo ARCH, essa relação pode ser dada, em que GARCH(p,q) é o mesmo que o modelo Heteroscedástico Condicional Autorregressivo ARCH($p+q$), que é representado pela equação 3.2 (Gujarati and Porter, 2011).

$$\sigma_t^2 = \alpha_0 + \alpha_1 \mu_{t-1}^2 + \dots + \alpha_m \sigma_{t-m}^2 \quad (3.2)$$

Como você pode ver, a equação 3.2 é semelhante à equação 3.1. A diferença está no fato de que o modelo GARCH considera suas variâncias além dos dados passados. Portanto, o modelo GARCH seria o modelo ARCH com parâmetros adicionais (Ampountolas, 2021). Portanto, em termos gerais, GARCH(p, q) será dado por $\mu_t = \sigma^t \varepsilon$, reescrevendo assim o modelo de acordo com a equação 3.3.

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^m \alpha_i \mu_{t-i}^2 + \sum_{j=1}^s \beta_j \sigma_{t-j}^2 \quad (3.3)$$

Onde $\alpha_0 > 0$, $\alpha_i \geq 0$, $\beta_j \geq 0$ e $\sum_{i=1}^{max(m,s)} (\alpha_i + \beta_i) < 1$. Assim, um valor v_t é considerado, que será dado pela relação $v_t = \alpha_t^2 - \sigma_t^2$, de modo que $\sigma_t^2 = \alpha_t^2 - v_t$, observando os períodos passados $(t-1)$, pode ser reescrito com $\sigma_{t-i}^2 = \alpha_{t-i}^2 - v_{t-i}$, dado que $(i = 1, \dots, n)$, de modo que a equação 3.3 pode ser reescrita da seguinte forma, dada pela equação 3.4 (Yıldırım and Bekun, 2023).

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^{max(m,s)} (\alpha_i + \beta_i) \mu_{t-i}^2 + v_t + \sum_{j=1}^s \beta_j v_{t-j}^2 \quad (3.4)$$

Onde $\alpha_i + \beta_i$ representam o coeficiente de associados aos choques passados μ_{t-i}^2 , o β_j o coeficiente de ponderação dos choques passados não explicados v_{t-j}^2 e v_t sendo os choque imprevisíveis. Dessa forma, o modelo GARCH dado pela equação 3.4 será o mesmo que o modelo ARMA ao quadrado.

Para definir os parâmetros, a função *auto.arima()* identifica o melhor modelo, ajustando os parâmetros e usando os índices AIC, AICC e BIC para indicar o melhor modelo. Isso pode ajudar no processo de seleção de parâmetros, considerando que esse é um processo dinâmico com várias dependências. Além disso, todos os testes estatísticos devem ser realizados para verificar se o GARCH é ideal para aplicação nesse cenário, pois há casos em que o ARIMA teria um desempenho satisfatório (Chen et al., 2011).

3.4 PREPARAÇÃO DO BANCO DE DADOS

3.4.1 Consolidação do Banco de Dados

A volatilidade da ELET3, por ser a maior empresa do setor elétrico da América Latina, foi considerada ideal para medir a volatilidade dos preços da energia. Os dados das ações foram obtidos no *Yahoo Finance*³. Os dados utilizados correspondem aos preços diários de fechamento do mercado de ações em BRL. Assim, sempre que o algoritmo for executado, os dados mais recentes serão capturados, referentes ao dia anterior. Todos os dados e códigos utilizados nesta pesquisa estão disponíveis no *Figshare*⁴. A empresa possui três ativos listados na bolsa: ELET, ELET6 e ELET3. Dentre esses, a ELET3 foi escolhida por representar ações ordinárias, concedendo aos investidores direito a voto e participação direta nas decisões da empresa. Isso significa que a ELET3 está mais envolvida nas decisões estratégicas e políticas da companhia. A pesquisa tem como objetivo utilizar esses ativos para avaliar seu impacto sobre o preço da matriz energética brasileira. Foram utilizados dados do comportamento dos preços das ações da ELET3 negociadas na B3 no momento de seu fechamento. Assim, a Tabela 3.1 apresenta o resumo estatístico das variáveis estudadas; pode-se observar que os valores mínimos, 1º quartil, mediana, média, 3º quartil e máximos referem-se à data de fechamento da ELET3.SA.

Tabela 3.1: Resumo estatístico dos preços da Eletrobras. Valores em BRL.

	ELET3.SA	ÍNDICE
Mínimo	4.408	2007-01-02
1º Quartil	14.255	2011-05-09
Mediana	23.451	2015-09-08
Média	23.392	2015-09-03
3º Quartil	23.451	2020-01-02
Máximo	51.570	2024-04-29
Período	2007/2024	

Com base na Figura 3.2, fica evidente que os dados não apresentam comportamento estacionário. Isso é importante, pois os modelos ARIMA e GARCH são utilizados para dados com média e variância constantes ao longo do tempo. Os dados exibem uma tendência clara, com uma queda significativa a partir de 2010, seguida por uma tendência de alta a partir de 2016.

³<https://finance.yahoo.com>. Acesso em: 30 de abril de 2024.

⁴<https://doi.org/10.6084/m9.figshare.25726695>



Figura 3.2: Volatilidade dos preços das ações da Eletrobras ao longo do tempo (diário).

Portanto, uma análise preliminar sugere que os dados são não estacionários, resultando em volatilidade e variância não constante (heterocedasticidade do modelo). No entanto, essas conclusões devem ser rigorosamente testadas estatisticamente para garantir resultados e previsões robustas e condizentes com a realidade.

3.4.2 Análise da presença de volatilidade nos dados

A volatilidade pode ser alta em certos momentos e baixa em outros. Isso é caracterizado por mudanças contínuas, portanto, saltos não são consistentes. Além disso, a volatilidade tende a flutuar dentro de um intervalo fixo (indicando que a volatilidade é estacionária). Ela reage de forma diferente a um aumento substancial de preço e a uma diminuição igualmente significativa, sendo que a última tem um impacto maior. Com base no gráfico da Figura 3.2 e no valor p do teste *Dickey-Fuller* aumentado (ADF) de 0,7148, pode-se concluir que o modelo não é estacionário porque o valor p é mais significativo que 0,05. Isso indica que a hipótese nula é aceita e o modelo possui uma raiz unitária. Para que a série seja considerada estacionária, ela deve ter média, variância e autocorrelação constantes, para que não mude ao longo do tempo. Ela pode variar para valores mais altos ou mais baixos ao longo do tempo, mas convergirá para a mesma média.

Apesar de ser o ponto de partida, as características de uma série de preços não são suficientemente úteis para trabalhar com a estimativa de volatilidade. Uma transformação deve ser realizada para estabilizar e converter a série em uma estacionária. A abordagem mais comum é realizar uma transformação logarítmica, pois reduz a escala da série, tornando-a mais suave e eliminando a primeira diferença da série. Certas suposições são feitas ao calcular o retorno contínuo diário: o retorno médio é zero, a variação do retorno muda ao longo do tempo com base nos retornos contínuos e a volatilidade é medida com base em seu desvio padrão.

Uma alternativa para transformar a série não estacionária em uma série estacionária é aplicar um comando em R chamado *ndiffs* para auxiliar nesse processo de diferenciação. Os preços da ELET3.SA mostraram $ndiffs = 1$. Portanto, apenas uma diferenciação foi feita para esta série temporal diária. Os novos valores da série ELET3.SA podem ser vistos no resumo estatístico na Tabela 3.2.

Tabela 3.2: Resumo estatístico dos preços da Eletrobras com logaritmo e com uma diferenciação

	ELET3.SA	INDEX
Min	-0.3333303	2007-01-02
1° Qu.	-0.0157149	2011-05-09
Mediana	0.0000000	2015-09-08
Média	0.0000922	2015-09-03
3°Qu	0.0148975	2020-01-02
Máx	0.4007593	2024-04-29
Período	2007/2024	

Com base no gráfico abaixo, os retornos contínuos e o teste *Dickey-Fuller* aumentado (ADF) com um valor p inferior a 0.01 indicam que a hipótese nula (H_0) é rejeitada. Isso sugere que a série não possui uma raiz unitária, confirmando a presença de estacionariedade. O retorno diário, calculado como a diferença entre o logaritmo natural do preço atual e o logaritmo natural do preço anterior, pode ser observado na Figura (veja a Figura 3.3). A série de retornos indica variância estabilizada, o que sugere que os retornos são estacionários com uma média de zero.

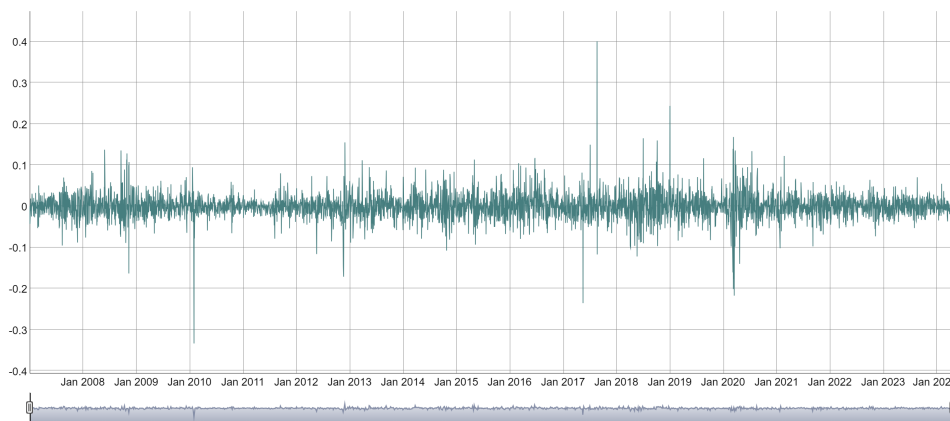


Figura 3.3: Retorno Diário Eletrobras - ELET3.SA

Os gráficos da função ACF do modelo revelam tendências e uma raiz unitária, e não há evidência de ergodicidade do modelo. De acordo com o teste ACF, uma diminuição gradual nos *lags* indica a não estacionariedade dos dados do modelo. A Figura 3.4 mostra o gráfico da função ACF.

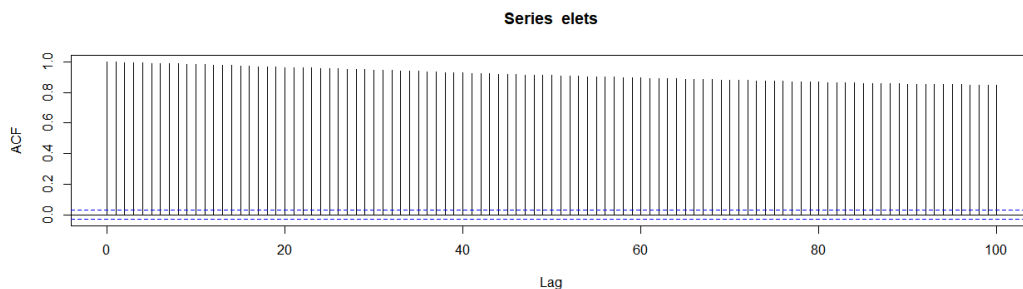


Figura 3.4: Gráfico da função ACF do modelo

O gráfico a seguir mostra o diagrama de autocorrelação parcial, que possibilita observar sazonalidade ou tendência através desses gráficos, onde gráficos sazonais mostram padrões de barras que variam entre valores positivos e negativos. Finalmente, gráficos com tendências mostrarão uma alta correlação, que diminui ao longo do tempo. O gráfico da função PACF pode ser visto na Figura 3.5.

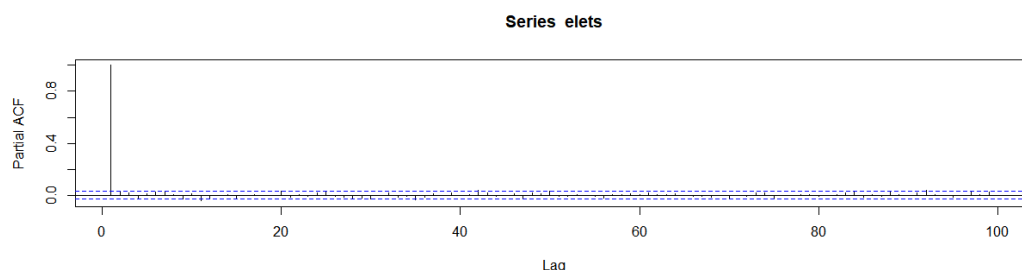


Figura 3.5: Gráfico da função PACF do modelo

Com base nas figuras 3.4 e 3.5, e no teste de estacionariedade ADF. Há evidências de autocorrelação na série. Com base na Figura 3.3, fica evidente que os dados são voláteis. Os resultados indicam a necessidade de aplicar um ARMA antes do modelo GARCH.

3.5 TESTE DOS PARÂMETROS USANDO AUTO.ARIMA

O *auto.arima* do R permite que esses parâmetros sejam determinados automaticamente. Ele funciona de forma semelhante à metodologia *Box-Jenkins*, onde os parâmetros são redefinidos até que o melhor modelo seja encontrado, de acordo com os testes AIC, AICc e BIC (quanto menor o módulo, melhor o modelo) para determinar o modelo de melhor ajuste para prever os dados. A vantagem de usar *auto.arima* é que ele testa todas as combinações (p, q) possíveis e, em seguida, compara todos os modelos para encontrar o melhor, o que difere do modo manual, que tenta estimar os melhores parâmetros e verificar se eles são satisfatórios para explicar as previsões, de modo que o melhor modelo nem sempre será necessariamente escolhido.

O modelo definido com o melhor parâmetro pelo comando *auto.arima* foi ARIMA(1, 1, 2). Os valores dos testes AIC, AICc e BIC fornecidos pelo modelo, bem como os valores de erro ME, RMSE, MAE, MPE, MAPE e MASE fornecidos pelo modelo, podem ser encontrados nas Tabelas 3.3 e 3.4.

Tabela 3.3: Valores AIC, AICc e BIC dos parâmetros *auto.arima*

ARIMA	AIC	AICc	BIC
(1,1,2)	9544.15	9544.161	9569.61

Tabela 3.4: Valores ME, RMSE, MAE, MPE, MAPE e MASE para os parâmetros *auto.arima*.

ARIMA	ME (BRL)	RMSE (BRL)	MAE (BRL)	MPE (%)	MAPE (%)	MASE
(1,1,2)	0.0031	0.7337	0.4633	-0.0403	2.0966	0.9996

Na Tabela 3.3, os valores AIC, AIC_c e BIC indicam baixas penalidades para o ajuste e complexidade do modelo. A Tabela 3.4 mostra o erro médio (ME), que está próximo de zero, indicando que o modelo não possui viés significativo e os erros positivos e negativos estão equilibrados. O erro absoluto médio (MAE) também é baixo ao desconsiderar a direção do erro (positivo ou negativo). Embora o erro da raiz quadrada média (RMSE) seja relativamente alto, penalizando grandes erros, é razoável considerando a volatilidade nas séries temporais financeiras. Os erros percentuais (MPE, MAPE e MASE) indicam que o modelo ARIMA(1,1,2), em média, tem um valor ligeiramente inferior aos valores reais (-0,0403%), com os valores previstos próximos dos valores reais (2,0966%), o que mostra previsões consistentes.

3.6 VERIFICAÇÃO DOS RESÍDUOS DOS MODELOS

3.6.1 Resíduos do modelo ARMA

O resíduo quadrado foi gerado, o que definiu o modelo - ARMA(1,2), dado pelo *auto.arima* como sendo o modelo mais representativo, conforme mostrado na Figura 3.6, que contém o resíduo quadrado dos preços ELET3.SA.

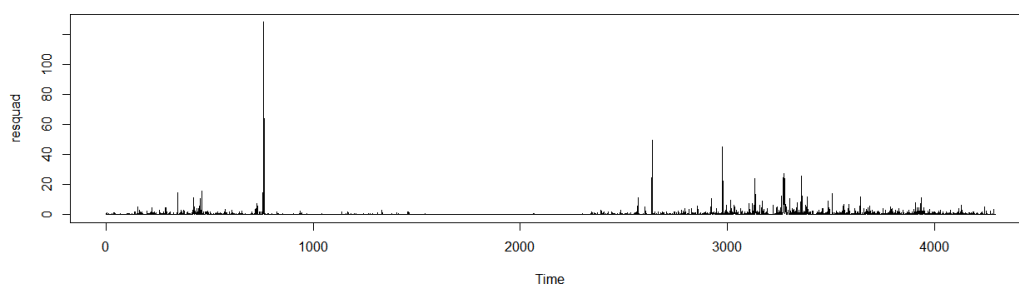


Figura 3.6: Resíduo quadrado - ELET3

Com base nas inferências analisadas, não há evidência de ruído branco, caracterizado por uma média nula, desvio padrão constante e nenhuma autocorrelação entre os *lags*. Os resíduos do modelo de regressão gerado pelo gráfico de resíduos podem ser vistos na Figura 3.7.

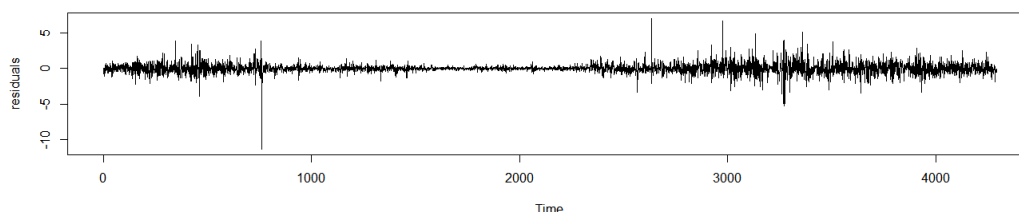


Figura 3.7: Gráfico de resíduos

Outra forma de verificar a heterocedasticidade condicional é visualizar os gráficos FAC e FACP da série de resíduos quadrados do modelo ARMA, observando que há uma dependência entre as variâncias. Essa

correlação é persistente mesmo em *lags* altos. Portanto, a análise visual está de acordo com os resultados esperados. Os gráficos com as funções ACF e PACF dos resíduos podem ser vistos nas Figuras 3.8 e 3.9.

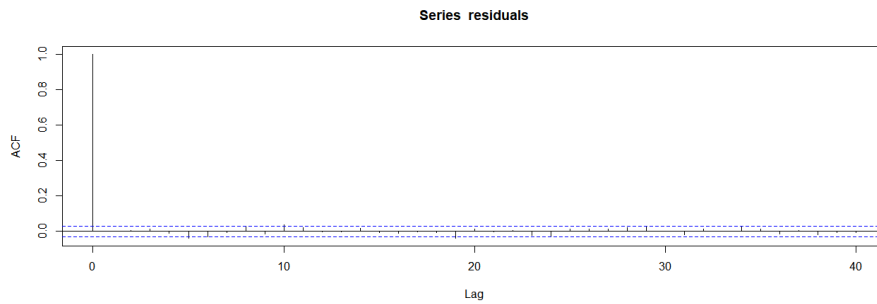


Figura 3.8: Gráfico ACF dos resíduos

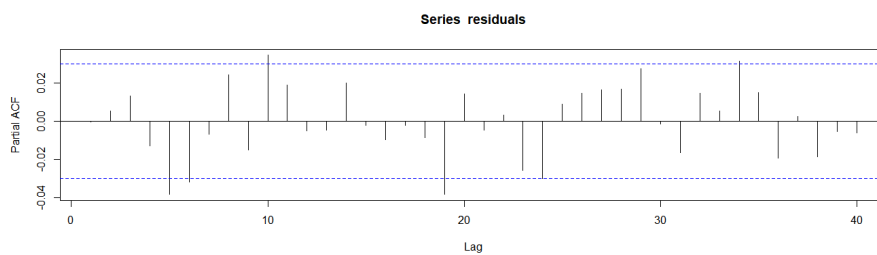


Figura 3.9: Gráfico PACF dos resíduos

3.6.2 Resíduos do modelo GARCH

Continuando com o diagnóstico dos resíduos, a figura a seguir mostra o gráfico com o envelope simulado de 95% do modelo GARCH(1, 1) seguindo a distribuição normal e o modelo GARCH(1, 1) seguindo a distribuição *t* de *Student*. Pode-se observar que os resíduos do modelo GARCH(1, 1) com distribuição *t* de *Student* têm resíduos que estão mais alinhados com a distribuição especificada, com a maioria dos pontos situados em cima da linha reta e até mesmo os pontos mais distantes nas caudas da distribuição permanecendo dentro do envelope simulado do intervalo de 95% do gráfico de probabilidade. O gráfico de probabilidade do modelo GARCH(1, 1) seguindo a distribuição normal e a distribuição *t* de *Student* pode ser visto na Figura 3.10 modelo.

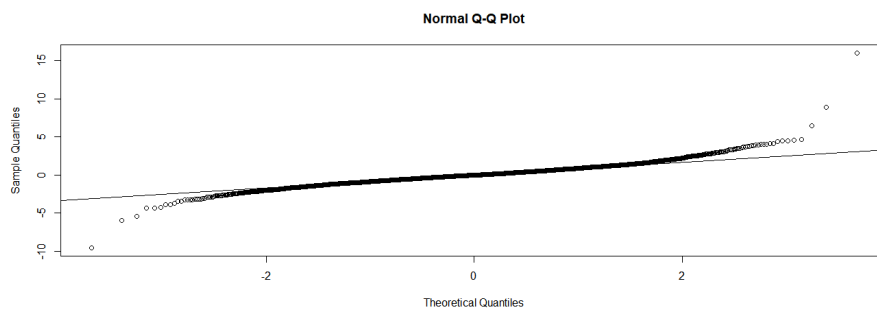


Figura 3.10: Gráfico de probabilidade do modelo GARCH(1,1)

A Figura 3.10 modelo mostra um descolamento dos extremos, ou seja, como a série não está próxima da linha, isso indica que a série não tem uma distribuição normal. Isso se deve aos momentos extremos em que o modelo tem variância condicional em determinados momentos. Assim, a partir dos resultados encontrados na seção 3.4.2, com os testes de autocorrelação, autocorrelação parcial, *Dickey-Fuller* aumentado, *Q-Q plot* normal, bem como a avaliação quadrada dos resíduos e retornos contínuos, pode-se dizer que o uso de $GARCH(1, 1)$ é adequado para prever a série de preços ELET3, juntamente com o modelo ARMA.

3.7 DISCUSSÃO E RESULTADOS

3.7.1 Aplicação da previsão

Inicialmente, a volatilidade condicional (o comportamento de volatilidade é explicado por padrões passados) é demonstrada usando o modelo $GARCH(1,1)$. A Figura 3.11 ilustra a volatilidade condicional estimada pelo modelo $GARCH(1,1)$.



Figura 3.11: Volatilidade condicional modelo $GARCH(1, 1)$

Conforme destacado anteriormente nos gráficos apresentados na Figura 3.3, variações condicionais notáveis ocorreram no início de 2010, bem como em 2017 e 2020. Além disso, a Figura 3.11 mostra que alta volatilidade foi observada no início de 2010. Esse pico foi associado a uma demanda de energia menor do que a esperada naquele ano, que foi ainda mais exacerbada por severas secas no Brasil durante meados de 2010. Essas flutuações provavelmente contribuíram para o aumento da demanda de energia, afetando os preços da energia e resultando nos padrões incomuns representados no gráfico. Um evento de seca semelhante ocorreu em 2017, embora tenha sido menos severo do que o de 2010. Além disso, uma volatilidade significativa nos preços das ações da ELET3 foi observada em 2020, coincidindo com o início das medidas de lockdown e um aumento substancial nos casos de COVID-19 no Brasil, juntamente com os esforços do governo para gerenciar a pandemia.

Outra conclusão que pode ser tirada do gráfico é que períodos caracterizados por maior volatilidade frequentemente influenciados por fatores climáticos como seca e estiagem estão associados a maiores flutuações nos preços das ações. Essa tendência foi evidente de 2010 a 2017, o que pode ser atribuído à maior

demanda por serviços de energia durante períodos de alta produção de energia. No entanto, o início de 2020 apresentou um aumento notável na volatilidade dos dados, juntamente com uma queda acentuada nos preços das ações. Esse comportamento difere do padrão de maior volatilidade observado anteriormente. Assim, fica evidente que a volatilidade condicional do modelo GARCH(1, 1) reflete efetivamente a volatilidade esperada no momento atual de cada período com base em períodos anteriores.

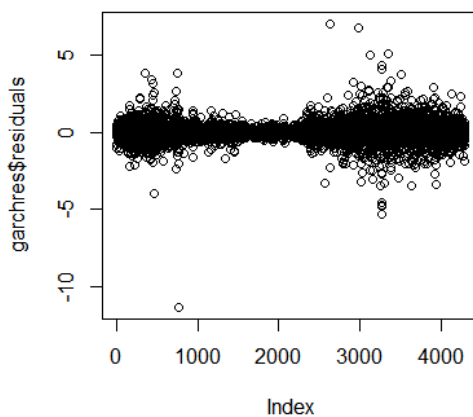
Para prever o modelo GARCH(1, 1), alguns testes serão apresentados para ver qual combinação ARMA (p, q) é melhor para previsão. A Tabela 3.5 mostra os valores dos critérios de informação critério de informação de *akaike* (AIC), critério de informação *bayesiano* (BIC), critério de informação de Shibata (SIC) e critério de informação de *Hannan-Quinn* (HQIC) para os seis melhores modelos ARMA em combinação com o modelo GARCH (1, 1) identificados neste estudo.

Tabela 3.5: Valores ME, RMSE, MAE, MPE, MAPE e MASE para os modelos GARCH

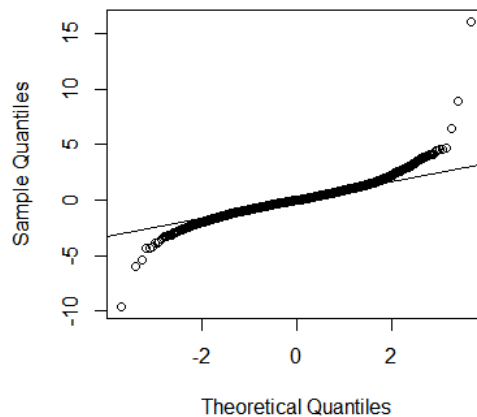
GARCH	ARMA	Akaike	Bayes	Shibata	Hannan-Quinn
(1,1)	(1,1)	1.672949	1.681836	1.672945	1.676088
(1,1)	(1,0)	1.672521	1.679926	1.672518	1.675136
(1,1)	(1,2)	1.673493	1.683861	1.673488	1.677155
(1,1)	(2,0)	1.672817	1.681703	1.672813	1.675955
(1,1)	(2,1)	1.672817	1.681703	1.672813	1.675955
(1,1)	(2,2)	1.672817	1.681703	1.672813	1.675955

Pode-se observar que o modelo ARMA(1, 0) possui os menores critérios de informação AIC e BIC, bem como os menores valores para os testes SIC e HQIC, de acordo com o princípio da parcimônia e há evidências de que foi capaz de remover a autocorrelação significativa na série.

O gráfico representado pela Figura 3.12a mostra se os dados possuem *outliers* ou pontos de alavancagem. Há uma não linearidade nos dados, pois a maioria dos dados está longe da linha horizontal, indicando alguns pontos localizados próximos ao intervalo de 5 e -5, o que é uma indicação de que vários pontos são considerados *outliers*, pois estão fora do intervalo de 3 e -3.



(a) Resíduos Normais GARCH(1,1)



(b) Resíduos padronizados GARCH(1,1)

Figura 3.12: Análise de resíduos

A presença de alavancagem de dados é evidente no $Q-Q$ plot apresentado na Figura 3.12b. O desvio dos pontos extremos da linha padrão sugere que a distribuição t -student fornece um ajuste mais apropriado para prever os dados da amostra. Além disso, o teste de Ljung-Box foi aplicado aos resíduos padronizados quadrados para avaliar a adequação do modelo. Os resultados obtidos por cada modelo GARCH(1, 1) em combinação com os parâmetros ARMA podem ser vistos na Tabela 3.6.

Tabela 3.6: Testes de resíduos quadrados de modelos GARCH

GARCH	Resíduo	X-squared	df	p-value
(1,1)	1	0.47528	1	0.4906
(1,1)	2	0.49433	1	0.4820
(1,1)	3	0.47736	1	0.4896
(1,1)	4	0.47643	1	0.4900
(1,1)	5	0.47643	1	0.4900
(1,1)	6	0.47643	1	0.4900

Os valores p em torno de 0,49 indicam que não há evidência estatística suficiente para rejeitar a hipótese nula do modelo. Além disso, os resíduos quadrados não mostram autocorrelação significativa, sugerindo que o modelo GARCH(1,1) é apropriado para modelar a variância condicional dos dados. Isso é suportado pela falta de qualquer estrutura inexplicada nos resíduos quadrados. Os parâmetros do modelo GARCH(1,1), juntamente com o modelo ARMA(1,0) considerado apropriado, foram usados para prever a série temporal para os próximos 1.460 dias. O gráfico que representa o modelo GARCH(1,1) pode ser encontrado na Figura 3.13.

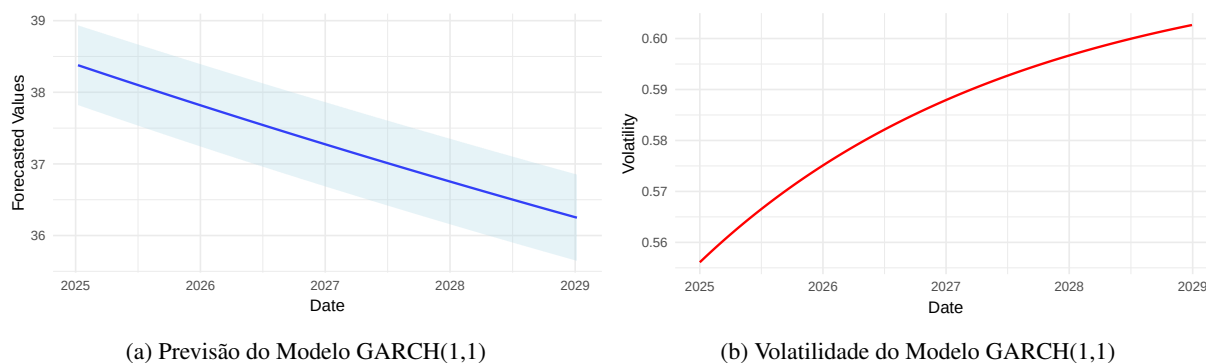


Figura 3.13: Previsões para 1.460 dias.

Após determinar os parâmetros ótimos para o modelo ARMA, os próximos 1.460 dias foram previstos usando o modelo GARCH(1,1) para ELET3. O gráfico que ilustra os preços das ações previstos para os próximos anos (veja a Figura 3.13a) revela uma tendência de queda no preço das ações. No entanto, a variação desses valores deve ser modesta, conforme indicado pela linha azul escura, enquanto a área azul mais clara representa os intervalos de confiança superior e inferior.

No gráfico que representa a volatilidade da série temporal (veja a Figura 3.13b), a previsão gerada pelo modelo GARCH(1,1) sugere uma potencial tendência de alta na volatilidade dos preços das ações da

ELET3.SA nos próximos 1.460 dias. Isso pode indicar uma maior probabilidade de instabilidade neste setor. Outra conclusão tirada dos resultados é que, embora o aumento da volatilidade observado nos dados possa ser atribuído a eventos específicos, como fatores climáticos e COVID-19 (Figura 3.11), também há uma tendência notável de aumento da volatilidade dos dados em prazos mais curtos, embora em menor grau, sem exigir nenhum evento específico para explicar esse aumento. Um resumo estatístico dos valores de previsão e da volatilidade dos dados está disponível na Tabela 3.7.

Tabela 3.7: Resumo da previsão para os próximos 1.460 dias

	Previsão	Sigma
Min	36.25	0.5561
1° Qu.	36.75	0.5751
Mediana	37.26	0.5879
Média	37.28	0.5851
3° Qu	37.81	0.5967
Máx	38.38	0.6027
Período	2024	2025

Os valores das ações não devem ultrapassar 38,38, mas podem atingir valores próximos a 36,25 até o ano de 2029. Os valores de referência da volatilidade apresentam um desvio padrão de 0,5561, que pode atingir o pico de 0,6027.

3.7.2 Correlação com os Preços da Energia Industrial

Essas variáveis devem ser correlacionadas para compreender o impacto dos preços das ações sobre os preços da energia. Os dados atuais dos preços da energia indústria anual foram obtidos da Empresa de Pesquisa Energética (EPE)⁵. Embora essa correlação não precise necessariamente ser positiva, ou seja, quando o valor da variável x_1 aumenta e o valor de y_1 também aumenta, essa relação também pode ser negativa, quando a variável x_1 diminui e o valor de y_1 aumenta. Inicialmente, pode-se afirmar que os preços das ações tendem a subir quando os preços da energia aumentam. No entanto, deve-se levar em conta que a ELET3 é um ativo de uma empresa geradora de energia, de modo que, quando o preço da energia sobe, isso tende a reduzir o consumo. Portanto, a hipótese de não normalidade do modelo deve ser confirmada para determinar qual modelo melhor mede a correlação entre as variáveis. O teste retorna o valor de W com um nível de significância dado pelo p -valor, onde $H_0 : p\text{-valor} > 0.05$ considera que a distribuição apresenta comportamento normal, e $H_1 : p\text{-valor} < 0.05$, considera que a distribuição não apresenta comportamento típico. Os resultados dos testes *Shapiro-Wilk* das amostras podem ser vistos na Tabela 3.8.

⁵<https://www.epe.gov.br/pt/publicacoes-dados-abertos/publicacoes/BEN-Series-Historicas-Completas>. Acesso em: 07 de maio de 2024

Tabela 3.8: Teste *Shapiro-Wilk* para Preço das Ações

Variável	Valor de W	p-valor
Preços das Ações da ELET3 em BRL	0.95578	$< 2.2e - 16$

A partir da Tabela 3.8 e dos valores de p -valor do teste *Shapiro-Wilk*, H_0 é rejeitado, pois p -valor < 0.05 , considerando, assim, que as amostras apresentam comportamento não normal, confirmando os resultados de heterocedasticidade do modelo, já apresentados nas Seções 3.4.2 e 3.7.1. Dessa forma, estima-se a correlação entre o preço médio atual da energia industrial e o preço médio das ações da ELET3. Os testes que podem ser realizados para avaliar a presença de correlação entre variáveis não normais são os testes de *Spearman* e *Kendall*. Aqui, dado o p -valor, tem-se $H_0 : p\text{-valor} > 0.05$, que considera a ausência de correlação, e $H_1 : p\text{-valor} < 0.05$, que indica a presença de correlação entre as variáveis. Os testes de correlação podem ser vistos na Tabela 3.9.

Tabela 3.9: Teste de Correlação de Spearman e Kendall

Teste	p-valor	Correlação
Spearman	0.02004	-0.5823529
Kendall	0.03263	-0.4000000

Os testes de correlação de *Spearman* (ρ) e *Kendall* (τ) apresentam um p -valor < 0.05 , rejeitando, assim, H_0 e aceitando H_1 , o que indica que há correlação entre os modelos, com um valor de $\rho = -0.5823529$ no teste de *Spearman* e um valor de $\tau = -0.4$ no teste de *Kendall*. De acordo com Cohen (1992), a correlação pode ter diferentes intensidades, conforme os intervalos apresentados na Tabela 3.10.

Tabela 3.10: Intensidade da Correlação

Teste	Intervalo	Intensidade da Correlação
ρ ou τ	≥ 0.10	Fraca
ρ ou τ	≥ 0.30	Moderada
ρ ou τ	≥ 0.50	Forte

Essa relação é caracterizada como negativa e de intensidade moderada, sugerindo que os preços das ações podem afetar as variáveis dos preços da energia. Notadamente, quando os preços das ações diminuem, os preços da energia no mercado tendem a subir. A correlação dos preços das ações sobre os preços da energia industrial atual é quantificada como $\rho^2 = 33.91\%$ e $\tau^2 = 16.00\%$. Esses valores indicam possíveis reações percentuais nos preços das ações em relação aos preços da energia, não representando uma relação de causalidade direta, mas sinalizando possíveis tendências no comportamento dos preços de energia. Embora haja evidências de que os preços das ações podem influenciar os preços da energia industrial, é essencial considerar outros fatores que também podem contribuir para a volatilidade dos preços, como a forte influência da ELET3, decorrente de políticas públicas que podem mudar de forma abrupta e são de difícil previsão em períodos futuros.

3.8 DISCUSSÕES FINAIS SOBRE O ESTUDO DE CASO DE VOLATILIDADE DAS AÇÕES DE ENERGIA

As previsões dos preços das ações da ELET3 são de grande importância para a definição dos preços da energia para a indústria brasileira, juntamente com a consolidação dos Objetivos de Desenvolvimento Sustentável, pois auxiliam na tomada de decisões para a construção de políticas públicas e medidas para a redução do custo da eletricidade no Brasil, tornando os setores da economia mais competitivos no mercado internacional. Além disso, isso aumentará a eficiência energética do Brasil, reduzirá custos e permitirá uma melhor gestão dos recursos ambientais por meio da introdução de novas tecnologias e energias renováveis.

Verificou-se que, para a previsão dos preços da energia, o modelo GARCH(1, 1), em conjunto com o ARMA(1, 0), apresentou melhor desempenho do que o modelo ARIMA(1, 1, 2) fornecido pelo *auto.arima*, uma vez que os dados são altamente voláteis. Isso ocorre devido à característica heterocedástica do modelo, ou seja, os dados apresentam um comportamento irregular. Os modelos GARCH(1, 1) e ARMA(1, 0) obtiveram valores de $AIC = 1.672521$ e $BIC = 1.679926$, inferiores às outras combinações testadas, como ARMA(1, 1), com $AIC = 1.672949$ e $BIC = 1.681836$, ARMA(1, 2), com $AIC = 1.673493$ e $BIC = 1.683861$, e os demais modelos ARMA(2, 0), ARMA(2, 1) e ARMA(2, 2), que apresentaram valores de $AIC = 1.672817$ e $BIC = 1.681703$ para a última combinação.

Além disso, a análise de correlação entre os preços das ações da ELET3 e os preços da energia industrial revelou uma correlação negativa de intensidade moderada. Isso foi indicado pelo coeficiente de correlação de *Spearman*, que apresentou um valor de $\rho = -0.5823529$, e pelo coeficiente de correlação de *Kendall*, que mostrou um valor de $\tau = -0.4$. Portanto, à medida que os preços das ações diminuem, espera-se que os preços da energia aumentem, com tamanhos de efeito entre as variáveis de $\rho^2 = 33.91\%$ e $\tau^2 = 16.00\%$. Em contraste, a correlação com o PIB mostrou uma relação positiva com os valores das ações, resultando em tamanhos de efeito de $\rho^2 = 31.57\%$ e $\tau^2 = 12.92\%$. No que se refere ao índice de eletricidade (IEE B3), não foi observada correlação significativa, indicando que a ELET3 possui características únicas em relação a outras ações dentro do mesmo setor.

Assim, embora os preços diários da energia não estejam disponíveis para análise, o que, em princípio, impede um diagnóstico mais preciso do setor energético, pode-se concluir que os preços das ações da ELET3 são altamente representativos no cenário energético. O monitoramento das tendências da ELET3 fornece uma alternativa para uma avaliação mais detalhada do mercado de preços da energia, que foi o objetivo inicial desta pesquisa, tornando-se, portanto, uma alternativa para a indústria na tomada de decisões e no planejamento estratégico das organizações, com base na análise da volatilidade apresentada pelos modelos GARCH e ARMA, resultando em uma tendência ascendente para as ações e descendente para os preços da energia.

Dessa forma, o estudo alcançou seu objetivo inicial de trazer maior clareza ao campo dos preços diários da energia no Brasil, detalhando o protocolo para aplicação dos modelos de heterocedasticidade condicional autorregressiva generalizada (GARCH), demonstrando as análises e critérios para avaliação dos preços da ELET3 e correlacionando tendências com os valores previstos dos preços médios da eletricidade industrial e as previsões do PIB. Ele fornece uma metodologia para que os agentes do setor energético apliquem na estruturação de estratégias e na formulação de políticas públicas para atender ao ODS 7, alinhando-se

ao desenvolvimento econômico por meio da redução de custos.

4 CONCLUSÃO

A presente dissertação encerra um esforço investigativo direcionado à compreensão e avaliação das dinâmicas de eficiência e vulnerabilidade do setor energético brasileiro, à luz de metodologias quantitativas estocásticas aplicadas ao longo de duas décadas de dados. A partir da integração entre a Análise Envoltória de Dados (DEA), algoritmos de reamostragem *bootstrap* e modelos de previsão ARIMA, SARIMA e GARCH, buscou-se construir um arcabouço metodológico robusto e aplicável à realidade energética do país, especialmente diante das necessidades impostas pelos Objetivos de Desenvolvimento Sustentável (ODS), com destaque ao ODS 7 – que trata do acesso confiável, sustentável, moderno e a preço acessível à energia.

A análise de eficiência energética por meio do DEA, ajustada com técnicas *bootstrap* para correção de *outliers*, revelou nuances significativas do uso dos recursos energéticos ao longo do tempo, permitindo inferências mais refinadas quanto ao desempenho relativo entre diferentes períodos e contextos. Tal abordagem demonstrou não apenas a capacidade técnica de captura de variações marginais de eficiência, mas também uma sensibilidade ao contexto sistêmico em que essas variações ocorrem. Os escores de eficiência obtidos, quando interpretados com o suporte de intervalos de confiança, evitaram a armadilha comum da superinterpretação dos valores brutos, oferecendo uma perspectiva mais confiável e estatisticamente sustentada dos resultados.

A modelagem preditiva das séries temporais de demanda energética e do Produto Interno Bruto, com o uso do SARIMA, permitiu observar a importância de se considerar a sazonalidade e o comportamento cíclico das variáveis macroeconômicas. Modelos lineares clássicos, como o ARIMA, embora amplamente utilizados, mostraram-se insuficientes para capturar plenamente as variações periódicas do sistema, especialmente em um país com marcada diferenciação regional e setorial. A boa performance do SARIMA nas métricas RMSE, MAE, MAPE e MASE consolida sua adequação como ferramenta de apoio à tomada de decisão, tanto em nível estratégico quanto operacional.

Ademais, a identificação de padrões consistentes nas séries temporais reforça a importância de uma abordagem prospectiva do planejamento energético. Em vez de depender exclusivamente de projeções baseadas em cenários políticos ou demandas estimadas *ad hoc*, a incorporação de modelos econométricos robustos permite antever gargalos e oportunidades com maior precisão. Essa capacidade preditiva ganha especial importância em contextos de restrição orçamentária, nos quais decisões equivocadas quanto à expansão da oferta podem comprometer a sustentabilidade financeira do sistema energético.

Outro eixo fundamental da pesquisa foi a análise da volatilidade do mercado energético, realizada com base nas ações da Eletrobras (ELET3). A aplicação do modelo GARCH(1,1), ajustado por ARMA(1,0), revelou-se eficaz para capturar a natureza heterocedástica dos retornos dessas ações, revelando períodos de maior instabilidade e seus gatilhos externos. O comportamento dos ativos energéticos, tradicionalmente menos voláteis que os de setores mais sensíveis à conjuntura global, apresentou, no entanto, picos acentuados em momentos de crise hídrica ou sanitária, sugerindo uma crescente correlação entre fatores ambientais e desempenho financeiro do setor.

A correlação negativa observada entre os preços das ações da Eletrobras e o valor médio da energia para o setor industrial, na ordem de $-33,91\%$, fornece uma chave interpretativa relevante para a compreensão das externalidades econômicas associadas à instabilidade do setor. Em outras palavras, quando o mercado antecipa riscos ao desempenho da empresa – por exemplo, diante de previsões de escassez hídrica – os preços das ações caem, ao passo que o custo da energia tende a aumentar. Essa dinâmica é particularmente prejudicial à indústria nacional, que já enfrenta desafios estruturais de competitividade.

O caráter sistêmico dessa vulnerabilidade se intensifica quando se observa que os episódios de maior volatilidade financeira coincidem com momentos de fragilidade institucional e ambiental. Crises de abastecimento, secas prolongadas, falta de investimentos em manutenção e expansão de infraestrutura, e a ausência de sinalizações claras de política energética atuam como catalisadores desses choques. Assim, os modelos GARCH não apenas cumprem papel técnico como instrumentos de previsão, mas se consolidam como ferramentas diagnósticas do grau de exposição do sistema a incertezas.

Ao analisar conjuntamente os resultados obtidos nas três frentes metodológicas DEA, previsão temporal e heterocedasticidade, emerge uma constatação central: o sistema energético brasileiro opera em um equilíbrio precário, onde ganhos de eficiência não são plenamente aproveitados, riscos financeiros não são suficientemente mitigados, e a capacidade preditiva das instituições permanece subutilizada. Em parte, essa fragilidade deriva da própria complexidade do setor, que congrega atores públicos e privados, regulação multissetorial, e dependência de variáveis climáticas difíceis de controlar. Contudo, também é reflexo de uma governança ainda fragmentada, pouco responsiva à necessidade de planejamento de longo prazo.

Nesse contexto, a dissertação contribui com um modelo integrado de análise, que se diferencia da literatura corrente por articular eficiência produtiva, estabilidade econômica e vulnerabilidade sistêmica de forma conjunta e coerente. A replicabilidade do arcabouço metodológico em diferentes contextos, seja em recortes regionais, seja em setores específicos da cadeia energética, amplia o potencial de uso dos achados. A adoção de DEA com ajuste *bootstrap*, aliada a modelos de séries temporais e análise de volatilidade, pode beneficiar estudos sobre consumo urbano, redes de distribuição e até impactos de tarifas sociais, gerando insumos valiosos para políticas públicas mais direcionadas.

Importante também destacar que a pesquisa não se limita a uma abordagem técnica. Ao evidenciar que a vulnerabilidade do setor é tanto técnica quanto institucional, ela convida a uma reflexão mais ampla sobre o papel do Estado, das agências reguladoras e dos agentes econômicos na construção de um sistema energético resiliente. A eficiência energética, nesse sentido, não é uma questão meramente operacional, mas política no sentido mais profundo do termo, pois diz respeito à alocação de recursos escassos, à definição de prioridades, e à garantia de direitos essenciais.

A observação da predominância de fontes hidrelétricas, que responde por mais de 60% da matriz energética brasileira, impõe desafios adicionais. Embora renovável, essa fonte é altamente vulnerável à variabilidade climática, o que torna o país especialmente sensível às mudanças ambientais globais. A integração crescente de fontes intermitentes, como solar e eólica, contribui para diversificar a matriz energética, mas impõe a necessidade de investimentos em armazenamento, redes inteligentes e mecanismos de resposta à demanda. A ausência de uma política energética integrada e de longo prazo compromete a consolidação de um sistema robusto o suficiente para lidar com essas múltiplas variáveis.

Neste sentido, a dissertação aponta para a urgência de políticas públicas que transcendam o ciclo elei-

toral, priorizando o fortalecimento institucional das agências reguladoras, a modernização da infraestrutura energética e o incentivo à pesquisa aplicada. O Brasil dispõe de centros de excelência em energia, como universidades, centros tecnológicos e empresas estatais com elevado capital técnico. Entretanto, falta muitas vezes articulação entre essas instituições, o que compromete a produção de soluções efetivas.

Outro ponto de destaque é a carência de sistemas de monitoramento contínuo de eficiência energética, capazes de fornecer indicadores em tempo real para gestores públicos e privados. A criação de painéis de indicadores dinâmicos, alimentados por dados integrados de consumo, produção, clima e mercado, poderia permitir intervenções mais rápidas e eficazes diante de anomalias ou tendências desfavoráveis.

Finalmente, a presente dissertação reafirma que eficiência e vulnerabilidade não são conceitos opostos, mas sim aspectos interdependentes de uma mesma estrutura. Um sistema eficiente pode ser altamente vulnerável se não possuir mecanismos de resposta a choques externos. Da mesma forma, um sistema robusto, porém ineficiente, pode se tornar insustentável ao longo do tempo. O equilíbrio entre esses dois polos exige não apenas técnica e tecnologia, mas também sensibilidade política, capacidade de coordenação institucional e compromisso com a justiça energética.

A justiça energética, entendida como o direito ao acesso universal, contínuo e sustentável à energia, ganha centralidade neste trabalho. O Brasil, como país em desenvolvimento com desigualdades profundas, não pode dissociar sua estratégia energética da promoção da equidade. A pesquisa conduzida contribui para esse debate ao mostrar que é possível promover eficiência sem ampliar a exclusão, e que a estabilidade do sistema depende, em grande parte, de sua capacidade de se adaptar às transformações sociais, econômicas e ambientais em curso.

Em síntese, os resultados e discussões aqui apresentados fornecem subsídios relevantes para a formulação de uma política energética mais inteligente e resiliente, como a Política Nacional de Transição Energética (PNTE), que tem como base a reestruturação da matriz energética brasileira para torná-la cada vez mais sustentável. Ao combinar robustez analítica com criticidade interpretativa, esta dissertação reforça a importância de integrar ciência, política e sociedade na construção de caminhos possíveis para um futuro energético mais seguro, eficiente, sustentável e justo.

Os resultados também podem servir como subsídio importante para a política da Tarifa Social de Energia Elétrica (TSEE), que visa oferecer descontos à comunidade de baixa renda, bem como a idosos ou pessoas com deficiência que recebam o Benefício de Prestação Continuada da Assistência Social (BPC). Isso porque os resultados apontaram uma tendência de aumento na demanda e nos preços da energia, o que pode impactar significativamente esse público. Portanto, considerar um aumento percentual no desconto concedido a essas famílias pode ser uma alternativa viável.

Outra aplicação relevante seria no âmbito da Política Nacional de Conservação e Uso Racional de Energia, cujo principal objetivo é a alocação eficiente dos recursos energéticos e a preservação do meio ambiente. O uso da metodologia DEA, tal como foi aplicado nesta pesquisa, pode evidenciar a efetividade dessa política no cenário nacional e indicar se, de fato, estamos avançando em direção a uma matriz mais eficiente e sustentável, com o uso racional dos recursos energéticos produzidos no país.

4.1 PERSPECTIVAS PARA PESQUISAS FUTURAS

A presente pesquisa abre espaço para desdobramentos relevantes que podem aprofundar a análise da eficiência e da vulnerabilidade no setor energético brasileiro. Trata-se de um ponto de partida para uma agenda de investigação mais ampla e ambiciosa, capaz de oferecer respostas ainda mais precisas e adaptadas à realidade heterogênea do país. Uma das direções mais promissoras consiste na regionalização da metodologia aplicada, permitindo que os modelos sejam ajustados e aplicados em escalas subnacionais, como unidades federativas, microrregiões ou consórcios intermunicipais. Essa abordagem permitiria revelar nuances espaciais que, em uma análise nacional agregada, tendem a ser suavizadas ou mesmo ignoradas.

A eficiência energética, por exemplo, apresenta variações significativas de acordo com o grau de urbanização, a infraestrutura instalada, os padrões de consumo e os incentivos locais à inovação. Estados com maior penetração de fontes renováveis, como o Nordeste com sua expressiva capacidade eólica e solar, ou o Sul com seu histórico de pequenas centrais hidrelétricas, podem apresentar configurações de fronteira tecnológica distintas daquelas observadas em centros industriais do Sudeste ou em regiões com deficit estrutural de energia, como a Amazônia Legal. Assim, a aplicação de modelos DEA em janelas regionais, calibradas com dados locais e incorporando variáveis socioeconômicas específicas, permitiria uma avaliação mais justa e eficaz da eficiência relativa entre regiões com condições estruturais diversas.

A adoção dessa perspectiva regional também favorece a formulação de políticas públicas mais alinhadas às realidades locais. Os governos estaduais e municipais, frequentemente responsáveis pela implementação de programas de eficiência energética, podem se beneficiar de análises que explicitem suas potencialidades e gargalos. Com isso, seria possível estabelecer metas realistas e calibradas, promover incentivos direcionados e monitorar os efeitos de políticas públicas com maior granularidade. A análise regionalizada, nesse sentido, não apenas enriquece a robustez acadêmica da metodologia, mas a aproxima das necessidades concretas da governança energética em múltiplos níveis.

Paralelamente, futuras pesquisas poderão incorporar variáveis climáticas e institucionais aos modelos preditivos, o que representaria uma importante inovação no campo da modelagem energética. A inclusão de índices de seca, dados de precipitação pluviométrica, médias de temperatura e informações sobre regimes hidrológicos poderia aprimorar significativamente a capacidade dos modelos ARIMA, SARIMA e GARCH em antecipar variações sazonais ou choques climáticos. A seca prolongada no Centro-Sul do país, por exemplo, tem impactos diretos sobre a geração hidrelétrica, e consequentemente sobre o custo marginal de operação do sistema, afetando tanto os preços da energia quanto a necessidade de despacho térmico. Incorporar essas variáveis aos modelos preditivos permite simular cenários de estresse hídrico e prever sua repercussão econômica, ampliando o potencial de resposta preventiva.

Além disso, variáveis institucionais, como alterações regulatórias, eventos políticos, mudanças na retórica de agências reguladoras ou aprovação de novos marcos legais, também afetam diretamente as expectativas de investimento e o comportamento dos agentes de mercado. Modelos enriquecidos com esse tipo de informação podem capturar efeitos estruturais sobre os preços, consumo e confiança do investidor, contribuindo para uma análise mais realista da dinâmica energética. A previsão de rupturas ou transições institucionais torna-se, assim, componente essencial de um sistema energético que se pretende mais resiliente e responsivo.

No que tange à integração metodológica, a combinação de modelos estatísticos tradicionais com técnicas avançadas de inteligência artificial (IA) surge como um dos caminhos mais promissores para elevar a acurácia preditiva em contextos complexos. A utilização de redes neurais profundas (*deep learning*), redes recorrentes (como LSTM e GRU), algoritmos de árvores de decisão (como *Random Forest* ou *XG-Boost*) e séries temporais probabilísticas pode enriquecer substancialmente a modelagem atual. Enquanto os modelos estatísticos oferecem estabilidade, interpretabilidade e fundamento teórico consolidado, os algoritmos de IA contribuem com capacidade de adaptação, detecção de padrões não lineares e robustez frente a grandes volumes de dados.

Modelos híbridos que combinem essas abordagens podem representar um salto qualitativo na análise preditiva, sobretudo em ambientes marcados por alta volatilidade, múltiplas interações e variáveis latentes. Um exemplo seria a integração de um modelo SARIMA para captura da estrutura sazonal da série com uma rede LSTM encarregada de identificar padrões residuais não explicados pelo modelo estatístico. Do ponto de vista computacional, esses modelos exigem infraestrutura adequada, mas já estão ao alcance de centros de pesquisa e universidades brasileiras. Sua aplicação ao setor energético, além de inovadora, aproxima o campo da econometria da moderna ciência de dados.

Na análise de eficiência, recomenda-se a exploração de modelos DEA dinâmicos, com janelas móveis ou estrutura intertemporal. Essa abordagem permitiria avaliar não apenas o desempenho relativo em um determinado ano, mas também a evolução da eficiência ao longo do tempo, considerando o aprendizado tecnológico, as mudanças de escala produtiva e o impacto acumulado de políticas públicas. A análise intertemporal é especialmente relevante em um setor sujeito a ciclos de investimento de longo prazo e dependência histórica de decisões passadas. A utilização de modelos de fronteira estocástica (SFA) com efeitos temporais também pode ser considerada, principalmente em contextos onde a incerteza dos dados observados é elevada.

Além disso, a inclusão de indicadores sociais e ambientais nos modelos de eficiência representa um avanço conceitual na direção de uma avaliação mais holística. Indicadores como o índice de pobreza energética, que mede a dificuldade das famílias em arcar com os custos mínimos de energia para uma vida digna, e a participação de fontes renováveis na matriz local, ajudam a contextualizar os resultados econômicos dentro de uma lógica de justiça energética e sustentabilidade. A eficiência, nesse novo paradigma, deixa de ser apenas uma medida técnica de produtividade e passa a incorporar elementos de equidade, impacto ambiental e coesão territorial.

No tocante à disseminação dos resultados, a criação de painéis interativos (*dashboards*) com dados atualizados, visualizações analíticas e filtros regionais pode transformar os achados da pesquisa em ferramentas práticas de apoio à decisão. Esses painéis, alimentados por bases públicas como a EPE, ONS, ANEEL, IBGE e INMET, poderiam ser acessados por gestores públicos, reguladores, empresas e até cidadãos, promovendo maior transparência, participação e controle social sobre os rumos da política energética. A visualização da eficiência por município, a simulação de impactos de cenários climáticos, ou a evolução dos indicadores ESG por setor, por exemplo, são funcionalidades que elevam o potencial transformador da pesquisa.

Por fim, sugere-se o uso de modelos econométricos estruturais, como Vetores Autorregressivos (VAR) e Modelos de Correção de Erros Vetoriais (VECM), para investigar relações de causalidade entre variáveis

macroeconômicas relevantes: consumo de energia, PIB, inflação, preços de energia, investimentos públicos e privados, entre outros. Esses modelos permitem não apenas observar correlações contemporâneas, mas também mapear os efeitos defasados e as direções causais, fornecendo subsídios mais ricos para o desenho de políticas. A incorporação de métricas ESG (ambientais, sociais e de governança) aos modelos amplia ainda mais sua aplicabilidade, permitindo o alinhamento entre desempenho energético, responsabilidade corporativa e sustentabilidade de longo prazo.

Em síntese, as possibilidades de ampliação aqui exploradas indicam que a presente pesquisa constitui apenas o ponto de partida de uma trajetória investigativa mais abrangente. As direções propostas não apenas aprofundam os aspectos técnicos e analíticos da modelagem, mas também respondem a um imperativo contemporâneo: o de construir um setor energético mais eficiente, transparente, resiliente e justo. Ao integrar variáveis exógenas, métodos inovadores e uma perspectiva territorializada, a pesquisa futura poderá contribuir de maneira decisiva para a formulação de políticas energéticas mais eficazes, orientadas por evidências e comprometidas com o desenvolvimento sustentável do Brasil.

REFERÊNCIAS BIBLIOGRÁFICAS

- T. Ahmad and H. Chen. A review on machine learning forecasting growth trends and their real-time applications in different energy systems. *Sustainable Cities and Society*, 54:102010, 2020. ISSN 2210-6707. doi: 10.1016/j.scs.2019.102010. URL <https://www.sciencedirect.com/science/article/pii/S2210670719335516>.
- Q. A. Al-Haija, O. Mohamed, and W. A. Elhaija. Predicting global energy demand for the next decade: A time-series model using nonlinear autoregressive neural networks. *Energy Exploration & Exploitation*, 41(6):1884–1898, 2023. doi: 10.1177/01445987231181919. URL <https://doi.org/10.1177/01445987231181919>.
- R. Alizadeh, L. Soltanisehat, P. D. Lund, and H. Zamanisabzi. Improving renewable energy policy planning and decision-making through a hybrid mcdm method. *Energy Policy*, 137:111174, 2020. ISSN 0301-4215. doi: 10.1016/j.enpol.2019.111174. URL <https://www.sciencedirect.com/science/article/pii/S0301421519307608>.
- M. H. Alsharif, M. K. Younes, and J. Kim. Time series arima model for prediction of daily and monthly average global solar radiation: The case study of seoul, south korea. *Symmetry*, 11(2), 2019. ISSN 2073-8994. doi: 10.3390/sym11020240. URL <https://www.mdpi.com/2073-8994/11/2/240>.
- A. Ampountolas. Modeling and forecasting daily hotel demand: A comparison based on sarimax, neural networks, and garch models. *Forecasting*, 3(3):580–595, 2021. ISSN 2571-9394. doi: 10.3390/forecast3030037. URL <https://www.mdpi.com/2571-9394/3/3/37>.
- T.-Z. Ang, M. Salem, M. Kamarol, H. S. Das, M. A. Nazari, and N. Prabakaran. A comprehensive study of renewable energy sources: Classifications, challenges and suggestions. *Energy Strategy Reviews*, 43:100939, 2022. ISSN 2211-467X. doi: 10.1016/j.esr.2022.100939. URL <https://www.sciencedirect.com/science/article/pii/S2211467X2200133X>.
- J. J. M. Antunes, J. C. Neves, L. R. C. Elmor, M. F. R. D. Araujo, P. F. Wanke, and Y. Tan. A new perspective on the u.s. energy efficiency: The political context. *Technological Forecasting and Social Change*, 186:122093, 2023. ISSN 0040-1625. doi: 10.1016/j.techfore.2022.122093. URL <https://www.sciencedirect.com/science/article/pii/S004016252200614X>.
- M. Arango, J. Díaz, and Y. Ramírez. Forecast of the energy price in colombia: An econometric application; [pronóstico de precio energético em colombia: Una aplicación económica]. *RISTI - Revista Iberica de Sistemas e Tecnologias de Informacao*, 2020(E27): 663 – 676, 2020. URL <https://www.scopus.com/inward/record.uri?eid=2-s2.0-85080990923&partnerID=40&md5=26ecf073946c4f04935791f779a12ece>.
- S. S. Arnob, A. I. M. S. Arefin, A. Y. Saber, and K. A. Mamun. Energy demand forecasting and optimizing electric systems for developing countries. *IEEE Access*, 11:39751–39775, 2023a. doi: 10.1109/ACCESS.2023.3250110.

- S. S. Arnob, A. I. M. S. Arefin, A. Y. Saber, and K. A. Mamun. Energy demand forecasting and optimizing electric systems for developing countries. *IEEE Access*, 11:39751–39775, 2023b. doi: 10.1109/ACCESS.2023.3250110.
- K. ArunKumar, D. V. Kalaga, C. M. Sai Kumar, G. Chilkoor, M. Kawaji, and T. M. Brenza. Forecasting the dynamics of cumulative covid-19 cases (confirmed, recovered and deaths) for top-16 countries using statistical machine learning models: Auto-regressive integrated moving average (arima) and seasonal auto-regressive integrated moving average (sarima). *Applied Soft Computing*, 103:107161, 2021. ISSN 1568-4946. doi: 10.1016/j.asoc.2021.107161. URL <https://www.sciencedirect.com/science/article/pii/S1568494621000843>.
- K. ArunKumar, D. V. Kalaga, C. Mohan Sai Kumar, M. Kawaji, and T. M. Brenza. Comparative analysis of gated recurrent units (gru), long short-term memory (lstm) cells, autoregressive integrated moving average (arima), seasonal autoregressive integrated moving average (sarima) for forecasting covid-19 trends. *Alexandria Engineering Journal*, 61(10):7585–7603, 2022. ISSN 1110-0168. doi: 10.1016/j.aej.2022.01.011. URL <https://www.sciencedirect.com/science/article/pii/S1110016822000138>.
- K. Aygul, M. Mohammadpourfard, M. Kesici, F. Kucuktezcan, and I. Genc. Benchmark of machine learning algorithms on transient stability prediction in renewable rich power grids under cyber-attacks. *Internet of Things*, 25:101012, 2024. ISSN 2542-6605. doi: <https://doi.org/10.1016/j.iot.2023.101012>. URL <https://www.sciencedirect.com/science/article/pii/S2542660523003359>.
- R. Azhar, F. S. D. Kesumah, Ambya, F. K. Wisnu, and E. Russel. Application of short-term forecasting models for energy entity stock price: evidence from indika energi tbk, jakarta islamic index. *International Journal of Energy Economics and Policy*, 10(1):294 – 301, 2020. doi: 10.32479/ijeep.8715. URL <https://www.scopus.com/inward/record.uri?eid=2-s2.0-85074982195&doi=10.32479%2fijeep.8715&partnerID=40&md5=7a4248da8f8e3efe77d3431e6b191bc9>.
- G. D. Bispo, G. F. Vergara, G. M. Saiki, P. H. d. S. Martins, J. G. Coelho, G. A. P. Rodrigues, M. N. d. Oliveira, L. R. Mosquera, V. P. Gonçalves, C. Neumann, and A. L. M. Serrano. Automatic literature mapping selection: Classification of papers on industry productivity. *Applied Sciences*, 14(9), 2024. ISSN 2076-3417. doi: 10.3390/app14093679.
- P. Bogetoft and L. Otto. *Benchmarking with dea, sfa, and r*, volume 157. Springer Science & Business Media, 2010.
- T. Bollerslev. Generalized autoregressive conditional heteroscedasticity, 1986.
- G. E. Box, G. M. Jenkins, G. C. Reinsel, and G. M. Ljung. *Time series analysis: forecasting and control*. John Wiley & Sons, 2015.
- A. A. Caldeira, M. D. Wilbert, T. B. S. Moreira, and A. L. M. Serrano. Brazilian state debt sustainability: An analysis of net debt and primary balance; [sustainability of brazilian state debt: An analysis of the relationship between net debt and primary result]; [sustainability of the debt of the brazilian states: An

- analysis of the net debt and the primary balance]. *Public Administration Magazine*, 50(2):285–306, 2016. ISSN 0034-7612. doi: 10.1590/0034-7612151140. URL <https://doi.org/10.1590/0034-7612151140>.
- A. Charnes, W. Cooper, and E. Rhodes. Measuring the efficiency of decision making units. *European Journal of Operational Research*, 2(6):429–444, 1978. ISSN 0377-2217. doi: 10.1016/0377-2217(78)90138-8. URL <https://www.sciencedirect.com/science/article/pii/0377221778901388>.
- C. Chen, J. Hu, Q. Meng, and Y. Zhang. Short-time traffic flow prediction with arima-garch model. In *2011 IEEE Intelligent Vehicles Symposium (IV)*, pages 607–612. IEEE, 2011. doi: 10.1109/IVS.2011.5940418.
- Y. Chen, G. Yin, and K. Liu. Regional differences in the industrial water use efficiency of china: The spatial spillover effect and relevant factors. *Resources, Conservation and Recycling*, 167:105239, 2021. ISSN 0921-3449. doi: 10.1016/j.resconrec.2020.105239. URL <https://www.sciencedirect.com/science/article/pii/S0921344920305541>.
- J. Cohen. Statistical power analysis. *Current Directions in Psychological Science*, 1(3):98–101, 1992. doi: <https://doi.org/10.1111/1467-8721.ep10768783>.
- T. Dimri, S. Ahmad, and M. Sharif. Time series analysis of climate variables using seasonal arima approach. *Journal of Earth System Science*, 129:149, 2020. ISSN 0973-774X. doi: 10.1007/s12040-020-01408-x.
- W. Enders. *Applied Econometric Time Series, 2nd Edition*. In: *Wiley Series in Probability and Statistics*. John Wiley and Sons, Inc., Hoboken, 2004.
- R. Engle and A. Patton. What good is a volatility model, 2001.
- R. F. Engle. Autoregressive conditional heteroscedasticity with estimates of the variance of united kingdom inflation. *Econometrica*, 50(4):987–1007, 1982. ISSN 00129682, 14680262. URL <http://www.jstor.org/stable/1912773>.
- N. Entezari and J. A. Fuinhas. Quantifying the impact of risk on market volatility and price: Evidence from the wholesale electricity market in portugal. *Sustainability (Switzerland)*, 16(7), 2024. doi: 10.3390/su16072691. URL <https://www.scopus.com/inward/record.uri?eid=2-s2.0-85190256309&doi=10.3390%2fsu16072691&partnerID=40&md5=6ddcb75d71d742a1a3aa57aad80b60ca>.
- E. R. C. (EPE). National energy balance 2023, 2023. URL <https://www.epe.gov.br/pt/publicacoes-dados-abertos/publicacoes/balanco-energetico-nacional-2023>. Access: 01, April, 2024.
- M. J. Farrell. The measurement of productive efficiency. *Journal of the Royal Statistical Society: Series A (General)*, 120(3):253–281, 1957. doi: 10.2307/2343100. URL <https://rss.onlinelibrary.wiley.com/doi/abs/10.2307/2343100>.

- C. Fernández and M. F. J. Steel. On bayesian modeling of fat tails and skewness. *Journal of the American Statistical Association*, 93(441):359–371, 1998. doi: 10.1080/01621459.1998.10474117. URL <https://www.tandfonline.com/doi/abs/10.1080/01621459.1998.10474117>.
- D. N. Gujarati and D. C. Porter. *Econometria básica-5*. Amgh Editora, 2011.
- V. Gundu and S. P. Simon. Pso-lstm for short term forecast of heterogeneous time series electricity price signals. *Journal of Ambient Intelligence and Humanized Computing*, 12(2):2375–2385, 2021. doi: 10.1007/s12652-020-02353-9. URL <https://doi.org/10.1007/s12652-020-02353-9>.
- F. Gökgöz and E. Yalçın. An environmental, energy, and economic efficiency analysis for the energy market in european union. *Environmental Progress & Sustainable Energy*, 42(4):e14068, 2023. doi: 10.1002/ep.14068. URL <https://aiche.onlinelibrary.wiley.com/doi/abs/10.1002/ep.14068>.
- M. R. Habibi, H. R. Baghaee, T. Dragičević, and F. Blaabjerg. Detection of false data injection cyber-attacks in dc microgrids based on recurrent neural networks. *IEEE Journal of Emerging and Selected Topics in Power Electronics*, 9(5):5294 – 5310, 2021. doi: 10.1109/JESTPE.2020.2968243. URL <https://www.scopus.com/inward/record.uri?eid=2-s2.0-85116550986&doi=10.1109%2fJESTPE.2020.2968243&partnerID=40&md5=f279a2e36706a52cae6fb4db361ae20b>.
- J. D. Hamilton and R. Susmel. Autoregressive conditional heteroskedasticity and changes in regime. *Journal of Econometrics*, 64(1):307–333, 1994. ISSN 0304-4076. doi: [https://doi.org/10.1016/0304-4076\(94\)90067-1](https://doi.org/10.1016/0304-4076(94)90067-1). URL <https://www.sciencedirect.com/science/article/pii/0304407694900671>.
- Y. Huang, X. Dai, Q. Wang, and D. Zhou. A hybrid model for carbon price forecasting using garch and long short-term memory network. *Applied Energy*, 285, 2021. ISSN 03062619. doi: <https://doi.org/10.1016/j.apenergy.2021.116485>.
- I. International Energy Agency. Electricity 2024. analysis and forecast to 2026., 2023. URL <https://www.iea.org/reports/electricity-2024>. Access: 25, April, 2024.
- M. Ismail, M. F. Shaaban, M. Naidu, and E. Serpedin. Deep learning detection of electricity theft cyber-attacks in renewable distributed generation. *IEEE Transactions on Smart Grid*, 11(4):3428 – 3437, 2020. doi: 10.1109/TSG.2020.2973681. URL <https://www.scopus.com/inward/record.uri?eid=2-s2.0-85081133842&doi=10.1109%2fTSG.2020.2973681&partnerID=40&md5=76f54f1c5e91ac405371db2f5deba8b1>.
- R. Jamil. Hydroelectricity consumption forecast for pakistan using arima modeling and supply-demand analysis for the year 2030. *Renewable Energy*, 154:1–10, 2020. ISSN 0960-1481. doi: <https://doi.org/10.1016/j.renene.2020.02.117>. URL <https://www.sciencedirect.com/science/article/pii/S0960148120303128>.
- G. Kapoor and N. Wichitaksorn. Electricity price forecasting in new zealand: A comparative analysis of statistical and machine learning models with feature selection. *Applied Energy*, 347, 2023.

- doi: 10.1016/j.apenergy.2023.121446. URL <https://www.scopus.com/inward/record.uri?eid=2-s2.0-85162079596&doi=10.1016%2fj.apenergy.2023.121446&partnerID=40&md5=20154c03c312b401a4dab41167fb023d>.
- L. Kong, G. Li, W. Rafique, S. Shen, Q. He, M. R. Khosravi, R. Wang, and L. Qi. Time-aware missing healthcare data prediction based on arima model. *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, pages 1–10, 2022. ISSN 1545-5963. doi: 10.1109/TCBB.2022.3205064.
- V. I. Kontopoulou, A. D. Panagopoulos, I. Kakkos, and G. K. Matsopoulos. A review of arima vs. machine learning approaches for time series forecasting in data driven networks. *Future Internet*, 15(8), 2023. ISSN 1999-5903. doi: 10.3390/fi15080255. URL <https://www.mdpi.com/1999-5903/15/8/255>.
- W.-T. Lin, G. Chen, and X. Zhou. Privacy-preserving federated learning for detecting false data injection attacks on power system. *Electric Power Systems Research*, 229:110150, 2024. ISSN 0378-7796. doi: <https://doi.org/10.1016/j.epsr.2024.110150>. URL <https://www.sciencedirect.com/science/article/pii/S0378779624000397>.
- X. Liu, P. Guo, and S. Guo. Assessing the eco-efficiency of a circular economy system in china’s coal mining areas: Emergy and data envelopment analysis. *Journal of Cleaner Production*, 206: 1101–1109, 2019. ISSN 0959-6526. doi: 10.1016/j.jclepro.2018.09.218. URL <https://www.sciencedirect.com/science/article/pii/S0959652618329512>.
- X. Liu, Z. Lin, and Z. Feng. Short-term offshore wind speed forecast by seasonal arima - a comparison against gru and lstm. *Energy*, 227:120492, 2021. ISSN 0360-5442. doi: 10.1016/j.energy.2021.120492. URL <https://www.sciencedirect.com/science/article/pii/S0360544221007416>.
- Z. Liu and S. Huang. Carbon option price forecasting based on modified fractional brownian motion optimized by garch model in carbon emission trading. *North American Journal of Economics and Finance*, 55, 2021. doi: 10.1016/j.najef.2020.101307. URL <https://www.scopus.com/inward/record.uri?eid=2-s2.0-85094873548&doi=10.1016%2fj.najef.2020.101307&partnerID=40&md5=1fbad676457fcc17c81610f15bac6131>.
- Z. Liu, M. Hajiali, A. Torabi, B. Ahmadi, and R. Simoes. Novel forecasting model based on improved wavelet transform, informative feature selection, and hybrid support vector machine on wind power forecasting. *Journal of Ambient Intelligence and Humanized Computing*, 9:1919–1931, 2018. doi: 10.1007/s12652-018-0886-0.
- S. Loizidis, A. Kyprianou, and G. E. Georghiou. Electricity market price forecasting using elm and bootstrap analysis: A case study of the german and finnish day-ahead markets. *Applied Energy*, 363, 2024. doi: 10.1016/j.apenergy.2024.123058. URL <https://www.scopus.com/inward/record.uri?eid=2-s2.0-85188577023&doi=10.1016%2fj.apenergy.2024.123058&partnerID=40&md5=02badaaa541267591d6c83bd648d7280>.

- S. Makridakis, E. Spiliotis, V. Assimakopoulos, A.-A. Semenoglou, G. Mulder, and K. Nikolopoulos. Statistical, machine learning and deep learning forecasting methods: Comparisons and ways forward. *Journal of the Operational Research Society*, 74(3):840–859, 2023. doi: 10.1080/01605682.2022.2118629. URL <https://doi.org/10.1080/01605682.2022.2118629>.
- Z. Malki, E.-S. Atlam, A. Ewis, G. Dagnew, A. R. Alzighaibi, G. ELmarhomy, M. A. Elhosseini, A. E. Hassanien, and I. Gad. Arima models for predicting the end of covid-19 pandemic and the risk of second rebound. *Neural Computing and Applications*, 33:2929–2948, 2021. ISSN 1433-3058. doi: 10.1007/s00521-020-05434-0.
- A. Mardani, H. Liao, M. Nilashi, M. Alrasheedi, and F. Cavallaro. A multi-stage method to predict carbon dioxide emissions using dimensionality reduction, clustering, and machine learning techniques. *Journal of Cleaner Production*, 275:122942, 2020. ISSN 0959-6526. doi: <https://doi.org/10.1016/j.jclepro.2020.122942>. URL <https://www.sciencedirect.com/science/article/pii/S0959652620329875>.
- D. Markovics and M. J. Mayer. Comparison of machine learning methods for photovoltaic power forecasting based on numerical weather prediction. *Renewable and Sustainable Energy Reviews*, 161:112364, 2022. ISSN 1364-0321. doi: <https://doi.org/10.1016/j.rser.2022.112364>. URL <https://www.sciencedirect.com/science/article/pii/S136403212200274X>.
- A. L. Marques Serrano, G. M. Saiki, C. Rosano-Penã, G. A. P. Rodrigues, R. d. O. Albuquerque, and L. J. García Villalba. Bootstrap method of eco-efficiency in the brazilian agricultural industry. *Systems*, 12(4), 2024. ISSN 2079-8954. doi: 10.3390/systems12040136. URL <https://www.mdpi.com/2079-8954/12/4/136>.
- K. Matsumoto and Y. Chen. Industrial eco-efficiency and its determinants in china: A two-stage approach. *Ecological Indicators*, 130:108072, 2021. ISSN 1470-160X. doi: 10.1016/j.ecolind.2021.108072. URL <https://www.sciencedirect.com/science/article/pii/S1470160X21007378>.
- Ministério de Minas e Energia. Consumo de energia deverá crescer em média 2,1% ao ano nos próximos dez anos, mostra projeção, out 2024. URL <https://www.gov.br/mme/pt-br/assuntos/noticias/consumo-de-energia-devera-crescer-em-media-2-1-ao-ano-nos-proximos-dez-anos-mo>
- P. Muniain and F. Ziel. Probabilistic forecasting in day-ahead electricity markets: Simulating peak and off-peak prices. *International Journal of Forecasting*, 36(4):1193 – 1210, 2020. doi: 10.1016/j.ijforecast.2019.11.006. URL <https://www.scopus.com/inward/record.uri?eid=2-s2.0-85078304861&doi=10.1016%2fj.ijforecast.2019.11.006&partnerID=40&md5=4ab21231c3ff5e9ee9ab2cc97a419328>.
- M. Narajewski and F. Ziel. Econometric modelling and forecasting of intraday electricity prices. *Journal of Commodity Markets*, 19:100107, 2020. ISSN 2405-8513. doi: <https://doi.org/10.1016/j.jcomm.2019.100107>. URL <https://www.sciencedirect.com/science/article/pii/S2405851319300728>.
- D. Nelson. Conditional heteroskedasticity in asset returns: A new approach, 1991.

- Y. Ning, H. Kazemi, and P. Tahmasebi. A comparative machine learning study for time series oil production forecasting: Arima, lstm, and prophet. *Computers & Geosciences*, 164:105126, 2022. ISSN 0098-3004. doi: <https://doi.org/10.1016/j.cageo.2022.105126>. URL <https://www.sciencedirect.com/science/article/pii/S009830042200084X>.
- S. Pascoe. On the use of data envelopment analysis for multi-criteria decision analysis. *Algorithms*, 17(3), 2024. ISSN 19994893. doi: 10.3390/a17030089. URL <https://www.scopus.com/inward/record.uri?eid=2-s2.0-85188826727&doi=10.3390%2fa17030089&partnerID=40&md5=6ebf69cafb1f2dcbe91405be25068e2a>. Cited by: 0; All Open Access, Gold Open Access.
- P. Pawłowski. Carbon emissions futures price forecasting with random forest. *Rynek Energii*, 2021.
- M. Pokushko, A. Stupina, I. Medina-Bulo, S. Ezhemanskaya, R. Kuzmich, and R. Pokushko. Algorithm for application of a basic model for the data envelopment analysis method in technical systems. *Algorithms*, 16(10), 2023. ISSN 19994893. doi: 10.3390/a16100460. Cited by: 0; All Open Access, Gold Open Access.
- I. Priyadarshini, A. Alkhayyat, A. Gehlot, and R. Kumar. Time series analysis and anomaly detection for trustworthy smart homes. *Computers and Electrical Engineering*, 102:108193, 2022. ISSN 0045-7906. doi: 10.1016/j.compeleceng.2022.108193. URL <https://www.sciencedirect.com/science/article/pii/S0045790622004347>.
- S. Rafayal, M. Cevik, and D. Kici. An empirical study on probabilistic forecasting for predicting city-wide electricity consumption. In *AI*, 2022.
- Y. Ramírez, M. Arango, and J. Díaz. Model for forecasting the price of the electrical market in colombia through wavelet transformed.; [modelo para pronosticar el precio del mercado eléctrico en colombia por medio de transformada wavelet]. *RISTI - Revista Iberica de Sistemas e Tecnologias de Informacao*, 2021(E41):448 – 461, 2021. URL <https://www.scopus.com/inward/record.uri?eid=2-s2.0-85105530560&partnerID=40&md5=28a538a1493b2d68c4d798561e1d0aa9>.
- D. d. Reis, R. Mattos, F. Perobelli, and W. Corrêa. Options pricing for petrobras share: A comparison between realized volatility and conditional heteroskedastic estimations, 2013.
- S. R. Riady and R. Apriani. Multivariate time series with prophet facebook and lstm algorithm to predict the energy consumption. In *2023 International Conference on Computer Science, Information Technology and Engineering (ICCoSITE)*, pages 805–810, 2023. doi: 10.1109/ICCoSITE57641.2023.10127735.
- C. Rosano-Peña, M. D. R. Pensado-Leglise, A. L. M. Serrano, A. A. Bernal-Campos, and M. Hernández-Cayetano. Agricultural eco-efficiency and climate determinants: application of dea with bootstrap methods in the tropical montane cloud forests of puebla, mexico. *Sustainable Environment*, 8(1): 2138852, 2022. doi: 10.1080/27658511.2022.2138852. URL <https://doi.org/10.1080/27658511.2022.2138852>.
- F. Rundo, F. Trenta, A. L. di Stallo, and S. Battiato. Machine learning for quantitative finance applications: A survey. *Applied Sciences*, 9(24), 2019. ISSN 2076-3417. doi: 10.3390/app9245574. URL <https://www.mdpi.com/2076-3417/9/24/5574>.

- S. Sahoo, T. Dragičević, and F. Blaabjerg. Cyber security in control of grid-tied power electronic converters—challenges and vulnerabilities. *IEEE Journal of Emerging and Selected Topics in Power Electronics*, 9(5):5326–5340, 2021. doi: 10.1109/JESTPE.2019.2953480.
- G. Saiki, A. Serrano, G. Rodrigues, C. Rosano, F. Pompermayer, and P. Albuquerque. An analysis of the eco-efficiency of the agricultural industry in the brazilian amazon biome. *Sustainability*, 16:5731, 07 2024. doi: 10.3390/su16135731.
- R. Sala-Garrido, M. Mocholi-Arce, A. Maziotis, and M. Molinos-Senante. The carbon and production performance of water utilities: Evidence from the english and welsh water industry. *Structural Change and Economic Dynamics*, 64:292–300, 2023. ISSN 0954-349X. doi: 10.1016/j.strueco.2023.01.003. URL <https://www.sciencedirect.com/science/article/pii/S0954349X23000036>.
- S. Sengar and X. Liu. Ensemble approach for short term load forecasting in wind energy system using hybrid algorithm. *Journal of Ambient Intelligence and Humanized Computing*, 11:5297–5314, 2020. doi: 10.1007/s12652-020-01866-7.
- A. L. M. Serrano, G. A. P. Rodrigues, P. H. d. S. Martins, G. M. Saiki, G. P. R. Filho, V. P. Gonçalves, and R. d. O. Albuquerque. Statistical comparison of time series models for forecasting brazilian monthly energy demand using economic, industrial, and climatic exogenous variables. *Applied Sciences*, 14(13), 2024. ISSN 2076-3417. doi: 10.3390/app14135846. URL <https://www.mdpi.com/2076-3417/14/13/5846>.
- A. Shadab, S. Ahmad, and S. Said. Spatial forecasting of solar radiation using arima model. *Remote Sensing Applications: Society and Environment*, 20:100427, 2020. ISSN 2352-9385. doi: 10.1016/j.rsase.2020.100427. URL <https://www.sciencedirect.com/science/article/pii/S2352938520302731>.
- C.-W. Shyu. A framework for ‘right to energy’ to meet un sdg7: Policy implications to meet basic human energy needs, eradicate energy poverty, enhance energy justice, and uphold energy democracy. *Energy Research & Social Science*, 79:102199, 2021. ISSN 2214-6296. doi: 10.1016/j.erss.2021.102199. URL <https://www.sciencedirect.com/science/article/pii/S2214629621002929>.
- S. Siami-Namini, N. Tavakoli, and A. Siami Namin. A comparison of arima and lstm in forecasting time series. In *2018 17th IEEE International Conference on Machine Learning and Applications (ICMLA)*, pages 1394–1401, 2018. doi: 10.1109/ICMLA.2018.00227.
- S. Siami-Namini, N. Tavakoli, and A. S. Namin. A comparative analysis of forecasting financial time series using arima, lstm, and bilstm, 2019.
- L. Simar and P. W. Wilson. Sensitivity analysis of efficiency scores: How to bootstrap in nonparametric frontier models. *Management Science*, 44(1):49–61, 1998. doi: 10.1287/mnsc.44.1.49. URL 10.1287/mnsc.44.1.49.
- E. Spiliotis, S. Makridakis, A.-A. Semenoglou, and V. Assimakopoulos. Comparison of statistical and machine learning methods for daily sku demand forecasting. *Operational Research*, 22(3):3037–3061, 2022. doi: <https://doi.org/10.1007/s12351-020-00605-2>.

- E. D. Spyrou, I. Tsoulos, and C. Stylios. Applying and comparing lstm and arima to predict co levels for a time-series measurements in a port area. *Signals*, 3(2):235–248, 2022. ISSN 2624-6120. doi: 10.3390/signals3020015. URL <https://www.mdpi.com/2624-6120/3/2/15>.
- B. D. Stosic. Technical efficiency of the brazilian municipalities. *Journal of Productivity Analysis*, pages 157–181, 2005. doi: 10.1007/s11123-005-4702-4.
- B. D. Stošić and M. d. C. S. de Sousa. Jackstrapping dea scores for robust efficiency measurement. *Econometric Society 2004 Latin American Meetings*, 2003.
- Suripto and Supriyanto. The effect of the covid-19 pandemic on stock prices with the event window approach: A case study of state gas companies, in the energy sector. *International Journal of Energy Economics and Policy*, 11(3):155 – 162, 2021. doi: 10.32479/ijee.10999. URL <https://www.scopus.com/inward/record.uri?eid=2-s2.0-85104626235&doi=10.32479%2fijee.10999&partnerID=40&md5=ae2412e19fa766d541056744ae8134b7>.
- A. Thangavel and V. Govindaraj. Forecasting energy demand using conditional random field and convolution neural network. *Elektronika ir Elektrotechnika*, 28(5):12–22, Oct. 2022. doi: 10.5755/j02.eie.30740. URL <https://eejournal.ktu.lt/index.php/elt/article/view/30740>.
- R. Tsay. Analysis of financial time series. *John Wiley and Sons*, 2005. doi: <http://dx.doi.org/10.1002/0471746193>.
- P. A. Verwiebe, S. Seim, S. Burges, L. Schulz, and J. Müller-Kirchenbauer. Modeling energy demand—a systematic literature review. *Energies*, 14(23), 2021. ISSN 1996-1073. doi: 10.3390/en14237859. URL <https://www.mdpi.com/1996-1073/14/23/7859>.
- K. Vijaya Chandrakala and P. Rajan. A hybrid machine learning approach for price forecasting in electricity market with smart bidding strategies and wind energy influence. *International Review on Modelling and Simulations*, 15(6):414 – 424, 2022. doi: 10.15866/iremos.v15i6.22653. URL <https://www.scopus.com/inward/record.uri?eid=2-s2.0-85158843905&doi=10.15866%2firemos.v15i6.22653&partnerID=40&md5=8bfadb87fe0eec008f228f3d9fe685d5>.
- F. Weschenfelder, G. de Novaes Pires Leite, A. C. Araújo da Costa, O. de Castro Vilela, C. M. Ribeiro, A. A. Villa Ochoa, and A. M. Araújo. A review on the complementarity between grid-connected solar and wind power systems. *Journal of Cleaner Production*, 257:120617, 2020. ISSN 0959-6526. doi: 10.1016/j.jclepro.2020.120617. URL <https://www.sciencedirect.com/science/article/pii/S0959652620306648>.
- P. W. Wilson. Detecting influential observations in data envelopment analysis. *Journal of Productivity Analysis*, pages 57–45, 1995. doi: 10.1007/BF01073493.
- P. W. Wilson. Fear: A software package for frontier efficiency analysis with r. *Socio-Economic Planning Sciences*, 42(4):247–254, 2008. ISSN 0038-0121. doi: 10.1016/j.seps.2007.02.001. URL <https://www.sciencedirect.com/science/article/pii/S0038012107000109>.

- Z. Xu, M. Mohsin, K. Ullah, and X. Ma. Using econometric and machine learning models to forecast crude oil prices: Insights from economic history. *Resources Policy*, 83, 2023. ISSN 0301-4207. doi: <https://doi.org/10.1016/j.resourpol.2023.103614>.
- P. T. Yamak, L. Yujian, and P. K. Gadosey. A comparison between arima, lstm, and gru for time series forecasting. In *Association for Computing Machinery, ACAI '19*, page 49–55, New York, NY, USA, 2020. ISBN 9781450372619. doi: 10.1145/3377713.3377722.
- R. Yasmeen, X. Yao, I. Ul Haq Padda, W. U. H. Shah, and W. Jie. Exploring the role of solar energy and foreign direct investment for clean environment: Evidence from top 10 solar energy consuming countries. *Renewable Energy*, 185:147–158, 2022. ISSN 0960-1481. doi: 10.1016/j.renene.2021.12.048. URL <https://www.sciencedirect.com/science/article/pii/S0960148121017687>.
- S. Yu, T. Lu, X. Hu, L. Liu, and Y.-M. Wei. Determinants of overcapacity in china’s renewable energy industry: Evidence from wind, photovoltaic, and biomass energy enterprises. *Energy Economics*, 97: 105056, 2021. ISSN 0140-9883. doi: <https://doi.org/10.1016/j.eneco.2020.105056>. URL <https://www.sciencedirect.com/science/article/pii/S0140988320303960>.
- H. Yildirim and F. V. Bekun. Predicting volatility of bitcoin returns with arch, garch and egarch models. *Future Business Journal*, 2023. ISSN 2314-7210. doi: <https://doi.org/10.1186/s43093-023-00255-8>.
- K. Zhou, W. Y. Wang, T. Hu, and C. H. Wu. Comparison of time series forecasting based on statistical arima model and lstm with attention mechanism. *Journal of Physics: Conference Series*, 1631(1):012141, sep 2020. doi: <https://dx.doi.org/10.1088/1742-6596/1631/1/012141>.
- J. Zhu, H. Long, J. Deng, and W. Wu. Risk measurement of oil price based on bayesian nonlinear quantile regression model. *Alexandria Engineering Journal*, 60(6):5567 – 5578, 2021. doi: 10.1016/j.aej.2021.04.029. URL <https://www.scopus.com/inward/record.uri?eid=2-s2.0-85107296406&doi=10.1016%2fj.aej.2021.04.029&partnerID=40&md5=5abfa14cc6f7ca7ee37ed537873eef70>.