



UNIVERSIDADE FEDERAL RURAL DE PERNAMBUCO
UNIDADE ACADÊMICA DO CABO DE SANTO AGOSTINHO
BACHARELADO EM ENGENHARIA ELÉTRICA

BRUNA LAVÍNIA SANTOS FORTES

Previsão de consumo de energia elétrica no brasil: uma comparação entre
métodos

CABO DE SANTO AGOSTINHO - PE
2023

BRUNA LAVÍNIA SANTOS FORTES

Previsão de consumo de energia elétrica no brasil: uma comparação entre
métodos

Monografia apresentada ao Curso de Graduação em Engenharia Elétrica da Unidade Acadêmica do Cabo de Santo Agostinho da Universidade Federal Rural de Pernambuco para obtenção do título de Bacharel em Engenharia Elétrica.

Área de concentração: Consumo de Energia Elétrica.

Orientador: Prof. Dr. Felipe Alberto Barbosa Simão Ferreira

CABO DE SANTO AGOSTINHO - PE

2023

Dados Internacionais de Catalogação na Publicação
Sistema Integrado de Bibliotecas da UFRPE
Bibliotecário(a): Ana Catarina Macêdo – CRB-4 1781

F738p Fortes, Bruna Lavinia Santos.
Previsão de consumo de energia elétrica no Brasil: uma comparação entre métodos / Bruna Lavinia Santos Fortes. – Cabo de Santo Agostinho, 2023.
80 f.; il.

Orientador(a): Felipe Alberto Barbosa Simão Ferreira.

Trabalho de Conclusão de Curso (Graduação) – Universidade Federal Rural de Pernambuco, Unidade Acadêmica Cabo de Santo Agostinho - UACSA, Bacharelado em Engenharia Elétrica, Cabo de Santo Agostinho, BR-PE, 2025.

Inclui referências.

1. Energia elétrica - Consumo. 2. Análise de séries temporais. 3. Inteligência artificial. I. Ferreira, Felipe Alberto Barbosa Simão, orient. II. Título

CDD 621.3

AGRADECIMENTOS

Agradeço de coração a Deus por me dar força, sabedoria e perseverança durante esta jornada acadêmica. Sua orientação e bênçãos estiveram sempre presentes, iluminando meu caminho e me dando a força necessária para enfrentar os desafios.

Minha profunda gratidão vai para minha amada família, meus pais e avô, que representaram o alicerce sólido sobre o qual este trabalho acadêmico se apoiou. Seu encorajamento, amor incondicional e apoio constante foram fundamentais para o meu sucesso, e sou grato por sempre acreditarem em mim.

Gostaria de expressar minha sincera gratidão ao meu dedicado orientador, Felipe Alberto Barbosa, cuja orientação, paciência e dedicação foram fundamentais para o desenvolvimento deste trabalho.

Neste momento de realização, quero agradecer a todos que fizeram parte desta jornada. Sem a presença de Deus, o apoio incansável da minha família e a orientação do meu orientador, este projeto não teria sido possível.

RESUMO

A previsão do consumo de energia elétrica desempenha um papel importante no gerenciamento eficaz dos recursos energéticos. Este estudo investiga métodos de previsão de consumo de energia elétrica aplicados no contexto brasileiro. Quatro técnicas de previsão amplamente reconhecidas foram exploradas: Suavização Exponencial, ARIMA (do inglês, Autoregressive Integrated Moving Average), Máquinas de vetores de suporte (MVS) e Redes Neurais Artificiais (RNA). A análise teve início com a decomposição da série temporal, permitindo a identificação de tendências, sazonalidades e variações aleatórias nos dados. A Suavização Exponencial de Holt-Winters e o modelo ARIMA foram utilizados para modelar os aspectos temporais subjacentes. A técnica das Máquinas de Vetores de Suporte (MVS) foi aprimorada por meio da otimização de parâmetros, resultando em previsões de alta precisão e confiabilidade. Embora exigisse esforços computacionais significativos, o MVS demonstrou notável capacidade na captura de relações complexas. As RNA foram investigadas e emergiram como a abordagem mais eficaz, fornecendo previsões excepcionalmente precisas e robustas. O método se destacou na captura de padrões complexos e sequenciais, sendo a escolha preferencial para prever o consumo de energia elétrica no Brasil. No entanto, a escolha final do modelo deve considerar não apenas métricas de avaliação, como MSE (do inglês, *Mean Squared Error*), MAE (do inglês, *Mean Absolute Error*), e RMSE (do inglês, *Root Mean-Squared Error*), mas também a adaptação aos dados específicos e as demandas computacionais.

Palavras-chave: consumo de energia elétrica; séries temporais; arima; suavização exponencial; inteligência artificial.

ABSTRACT

The prediction of electricity consumption plays a crucial role in the effective management of energy resources. This study investigates electricity consumption forecasting methods applied in the Brazilian context. Four widely recognized forecasting techniques were explored: Exponential Smoothing, ARIMA (Autoregressive Integrated Moving Average), Support Vector Machines (SVM), and Artificial Neural Networks (ANN). The analysis began with the decomposition of the time series, allowing for the identification of trends, seasonality, and random variations in the data. The Holt-Winters Exponential Smoothing and the ARIMA model were used to model the underlying temporal aspects. The Support Vector Machines (SVM) technique was enhanced through parameter optimization, resulting in highly accurate and reliable forecasts. Although it required significant computational efforts, SVM demonstrated remarkable capability in capturing complex relationships. ANN was investigated and emerged as the most effective approach, providing exceptionally precise and robust predictions. This method stood out in capturing complex and sequential patterns, making it the preferred choice for forecasting electricity consumption in Brazil. However, the final choice of the model should consider not only evaluation metrics such as MSE (Mean Squared Error), MAE (Mean Absolute Error), and RMSE (Root Mean-Squared Error) but also adaptation to specific data and computational demands.

Keywords: electrical energy consumption; time series; arima; exponential smoothing; artificial intelligence.

LISTA DE ILUSTRAÇÕES

| | |
|--|----|
| Figura 1 - Consumo de energia elétrica no Brasil (TWh) entre 2012 e 2021 | 18 |
| Figura 2 - Consumo de energia elétrica no Brasil por setor entre 2010 e 2022 | 21 |
| Figura 3 - Representação esquemática da metodologia de Box-Jenkins..... | 36 |
| Figura 4 - Representação de MVS com margem larga e MVS com margem estreita | 38 |
| Figura 5 - Representação de um neurônio artificial..... | 41 |
| Figura 6 - Consumo de energia Elétrica no Brasil (2004-2023)..... | 49 |
| Figura 7 - Distribuição dos Dados | 51 |
| Figura 8 - Tendência Decomposta | 52 |
| Figura 9 - Sazonalidade Decomposta | 52 |
| Figura 10 - Resíduo Decomposto..... | 53 |
| Figura 11 - Média Móvel e Desvio Padrão Móvel..... | 56 |
| Figura 12 - Fluxograma da Proposta de Previsão de Consumo de energia Elétrica no Brasil | 62 |
| Figura 13 - Previsão versus Consumo Efetivo de energia Elétrica: Suavização Exponencial..... | 64 |
| Figura 14 - Comparação entre Previsão (laranja) e Consumo Real (azul) de energia Elétrica pelo método ARIMA | 66 |
| Figura 15 - Previsão utilizando o método MVS | 68 |
| Figura 16 - Comparação entre Previsões e Valores Reais de Consumo de energia Elétrica | 70 |

LISTA DE TABELAS

| | |
|--|----|
| Tabela 1 - Equações do Método Aditivo e Multiplicativo do Holt-Winters..... | 33 |
| Tabela 2 - Distribuição dos Registros nos Conjuntos de Treinamento e Teste | 50 |
| Tabela 3 - Parâmetros Otimizados na Metodologia de Suavização Exponencial..... | 55 |
| Tabela 4 - Resultados do Teste Dickey-Fuller e Valores Correspondentes | 56 |
| Tabela 5 - Avaliação de Desempenho da Previsão utilizando Suavização Exponencial..... | 63 |
| Tabela 6 - Avaliação de Desempenho da Metodologia ARIMA..... | 65 |
| Tabela 7– Métricas de Avaliação MVS..... | 67 |
| Tabela 8 - Métricas de Desempenho da Previsão com RNA | 69 |

LISTA DE SIGLAS

| | |
|------------|--|
| ARIMA | Autoregressive Integrated Moving Average |
| IA | Inteligência Artificial |
| ML | Machine Learning |
| RNA | Rede Neural Artificial |
| MVS | Máquinas de Vetores de Suporte |
| MAE | Mean Absolute Error |
| MSE | Mean Squared Error |
| RMSE | <i>Root Mean-Squared Error</i> |
| SEP | Sistema Elétrico de Potência |
| ISO | International Organization for Standardization |
| PIB | Produto Interno Bruto |
| ARMA | Auto Regressivo de Média Móvel |
| RL | Regressão Linear |
| RLM | Regressão Linear Múltipla |
| SAM | Sistema de Acompanhamento do Mercado |
| MPL | Multilayer Perceptron |
| TCC | Trabalho de Conclusão de Curso |
| PIB | Produto Interno Bruto |
| BMLP | <i>Bridged Multilayer Perceptron</i> |
| FCC | <i>Fully Connected Cascade</i> |
| <i>RBF</i> | <i>Radial Basis Function</i> |
| <i>LVQ</i> | <i>Learning Vector Quantization</i> |
| RNN | <i>Recurrent Neural Network</i> |
| SEP | Sistema Elétrico de Potência |
| Procel | Programa Nacional de Conservação de energia Elétrica |
| RL | Regressão Linear |
| RLM | Regressão Linear Múltipla |
| LSSVM | <i>Least Squares Support Vector Machine</i> |
| FOA | <i>Fruit Fly Optimization Algorithm</i> |
| ACF | <i>Autocorrelation Function</i> |
| PACF | <i>Partial Autocorrelation Function</i> |
| EPE | Empresa de Pesquisa Energética |
| Proálcool | Programa Nacional do Alcool |

SUMÁRIO

| | |
|---|-----------|
| 1 INTRODUÇÃO | 12 |
| 1.1 MOTIVAÇÃO | 14 |
| 1.2 OBJETIVOS | 15 |
| 1.2.1 Objetivo Geral | 15 |
| 1.2.2 Objetivos Específicos | 15 |
| 1.3 ESTRUTURA DO TRABALHO | 15 |
| 2 CENÁRIO DE CONSUMO E DEMANDA DE ENERGIA ELÉTRICA NO BRASIL.. | 17 |
| 2.1. COMPREENDENDO O CONSUMO E A DEMANDA NO SISTEMA ELÉTRICO DE POTÊNCIA NO BRASIL | 17 |
| 2.2 CONSUMO DE ENERGIA ELÉTRICA NO BRASIL POR SETOR | 18 |
| 2.3 INFLUÊNCIAS NO CONSUMO DE ENERGIA ELÉTRICA | 21 |
| 2.4 EFICIÊNCIA ENERGÉTICA E O CONSUMO | 23 |
| 3 FUNDAMENTAÇÃO TEÓRICA | 25 |
| 3.1 SÉRIES TEMPORAIS: CONCEITOS E CLASSIFICAÇÃO | 25 |
| 3.1.1 Características da série temporal no consumo de energia elétrica | 26 |
| 3.2 MÉTODOS DE PREVISÃO | 27 |
| 3.3 MÉTODO DE SUAVIZAÇÃO EXPONENCIAL | 30 |
| 3.4 MODELO ARIMA | 33 |
| 3.5 MÁQUINAS DE VETORES DE SUPORTE | 37 |
| 3.6 REDE NEURAL ARTIFICIAL..... | 40 |
| 4 METODOLOGIA | 44 |
| 4.1 PROCEDIMENTOS METODOLÓGICOS..... | 44 |
| 4.2 BIBLIOTECAS E FERRAMENTAS PARA ANÁLISE DE DADOS | 45 |
| 4.3 COLETA DE DADOS PARA ANÁLISE E PREVISÃO | 47 |
| 4.3.1 Pré-processamento dos Dados | 48 |

| | |
|--|-----------|
| 4.3.2 Porcionamento dos Dados | 50 |
| 4.3.3 Decomposição da Série Temporal | 51 |
| 4.4 DETERMINAÇÃO DOS PARÂMETROS..... | 54 |
| 4.4.1 Suavização Exponencial | 54 |
| 4.4.2 ARIMA..... | 55 |
| 4.4.3 Máquinas de Vetores de Suporte | 57 |
| 4.4.4 Rede Neural Artificial..... | 58 |
| 4.5 MÉTRICAS DE AVALIAÇÃO..... | 60 |
| 5 RESULTADOS E DISCUSSÕES | 61 |
| 5.1 SUAVIZAÇÃO EXPONENCIAL..... | 62 |
| 5.2 ARIMA..... | 64 |
| 5.3 MÁQUINAS DE VETORES DE SUPORTE | 66 |
| 5.4 REDE NEURAL ARTIFICIAL..... | 68 |
| 5.5 COMPARAÇÃO ENTRE OS MÉTODOS | 70 |
| 6 CONCLUSÃO | 73 |
| REFERÊNCIAS..... | 75 |

1 INTRODUÇÃO

A preocupação com as mudanças climáticas se intensifica a cada dia no cenário global, e a busca por soluções que levem à neutralidade climática tem se tornado uma prioridade para muitos países. No ano de 2023, o Governo Federal do Brasil publicou diretrizes abrangentes para uma estratégia nacional voltada à neutralidade climática. A principal proposta é adotar medidas abrangentes para reduzir as emissões de gases de efeito estufa e transformar a matriz energética (Brasil, 2020). Uma das áreas-chave contempladas nessas diretrizes é o setor de energia. Reconhecendo sua importância como um dos principais contribuintes para as emissões de gases de efeito estufa, uma série de medidas específicas foi proposta visando a transição para uma matriz energética mais sustentável e renovável (Brasil, 2020).

A matriz elétrica brasileira desempenha um papel importante na economia do país e abrange a geração, transmissão e distribuição de energia elétrica. O Brasil tem um enorme potencial de geração de energia renovável, principalmente eólica e solar, e tem feito avanços significativos nessa área. Em janeiro de 2022, a capacidade instalada de energia eólica ultrapassou 20 GW (Machado, 2022). Dessa forma, o país ocupa a sétima posição em capacidade de energia eólica instalada.

Apesar dos avanços na geração de energia renovável, o setor de energia precisa de melhoria contínua, aprimorando a gestão de seus recursos. Uma das estratégias de gestão de recursos é a adoção de práticas de eficiência energética, que envolvem a implementação de medidas para reduzir o consumo de energia sem comprometer a qualidade da entrega. O padrão ISO (do inglês, *International Organization for Standardization*) 50001 estabelece as melhores práticas para gestão e eficiência energética e incentiva as organizações a desenvolver políticas e processos para o uso mais eficiente de energia (ISO, 2018). Uma das maneiras de atingir esse objetivo é utilizando métodos que podem auxiliar na previsão de consumo.

A previsão do consumo de energia elétrica é uma ferramenta importante para controlar a demanda e o consumo de eletricidade. Ela permite antecipar a evolução do consumo de energia e evitar o desperdício (Klyuev *et al.*, 2022). Isso é fundamental para o gerenciamento eficaz dos recursos energéticos e o planejamento de um fornecimento adequado. A projeção pode ser utilizada para evitar interrupções no

fornecimento de energia elétrica, otimizar a produção e reduzir custos operacionais (Carvalho, 2019).

No passado, técnicas como regressão linear, média móvel e ARIMA foram utilizadas para prever o consumo de energia. Essas técnicas funcionavam bem em situações simples e estáveis, nas quais os padrões de consumo não mudavam com o tempo. No entanto, à medida que os dados se tornaram mais complexos e novos padrões surgiram, essas técnicas começaram a apresentar limitações (Mohamed; Bodger, 2005).

As técnicas de inteligência artificial (IA) destacam-se por sua habilidade em lidar com a complexidade dos dados de consumo energético. Elas conseguem aprender e se adaptar aos padrões dos dados históricos, permitindo previsões mais precisas. A IA e o ML (do inglês, *Machine Learning*) têm o potencial de revolucionar a engenharia elétrica, trazendo inovações e melhorando a eficácia em diversas aplicações (Ahmad *et al.*, 2014).

Neste Trabalho de Conclusão de Curso (TCC), utilizaremos dados históricos do consumo de energia elétrica no Brasil para modelar e prever seu comportamento futuro. O objetivo principal deste estudo é explorar uma variedade de abordagens, utilizando tanto técnicas tradicionais de previsão (regressão linear e ARIMA) quanto técnicas avançadas de IA, como RNA e MVS.

Para entender melhor o comportamento do consumo de energia elétrica ao longo do tempo, utilizaremos a regressão linear inicialmente. Esse método nos permitirá identificar tendências e padrões lineares nos dados históricos. O ARIMA será empregado para modelar séries temporais estacionárias, permitindo-nos capturar a conexão entre os registros históricos e o comportamento atual do consumo de energia elétrica. Essa técnica é especialmente útil para lidar com padrões sazonais e ciclos que podem ter um impacto significativo no consumo de energia elétrica (Hyndman; Athanasopoulos, 2021).

Utilizaremos também as técnicas RNA e MVS, que são métodos mais avançados. A RNA, inspirada no funcionamento do cérebro humano, possui várias camadas de neurônios interconectados, permitindo a identificação de padrões não triviais nos dados e a previsão de comportamentos futuros. As MVS buscam encontrar um hiperplano ótimo para separar diferentes classes de dados, proporcionando previsões mais precisas (Haykin, 1998).

Inicialmente, cada uma dessas técnicas será implementada e treinada utilizando uma fração dos dados para teste. Durante o processo, ajustaremos os hiperparâmetros e otimizaremos o desempenho dos modelos para que possam lidar adequadamente com os dados reais de consumo de energia elétrica no Brasil. Isso inclui a seleção adequada de parâmetros, como neurônios em redes neurais, e parâmetros de regularização em modelos de Máquinas de Vetores de Suporte (Du *et al.*, 2018).

Depois, avaliaremos os resultados usando as métricas de desempenho de MAE, RMSE e coeficiente de determinação (R^2). Dessa maneira, iremos analisar o desempenho das técnicas de RNA e MVS em relação à abordagem padrão e verificar se elas trazem melhorias significativas na precisão das previsões do consumo energético. Avaliaremos a precisão, a habilidade em capturar padrões complexos nos dados e a eficiência computacional de cada técnica.

Ao concluir este TCC, pretende-se realizar uma análise comparativa abrangente entre as técnicas de RNA, MVS, ARIMA e suavização exponencial, aplicadas à previsão de consumo de energia elétrica no cenário brasileiro. O processo auxiliará na compreensão das vantagens e limitações de cada técnica, oferecendo perspectivas valiosas para aperfeiçoar as práticas de previsão e colaborar com a eficácia e sustentabilidade do setor elétrico.

1.1 MOTIVAÇÃO

A análise comparativa de métodos de previsão para o consumo de energia elétrica oferece informações importantes para as empresas de energia no Brasil, permitindo que tomem decisões fundamentadas sobre a adoção de tecnologias avançadas. Ao considerar a capacidade de adaptação dessas técnicas às mudanças contínuas no setor, como variações na demanda, expansão das fontes renováveis de energia e mudanças climáticas, é possível obter projeções mais precisas e tomar decisões mais eficientes.

A utilização de técnicas avançadas de IA na previsão de consumo de energia elétrica no Brasil traz benefícios significativos, tanto em termos de precisão quanto de eficiência operacional. À medida que o setor energético continua evoluindo, é fundamental explorar e adotar abordagens inovadoras para enfrentar os desafios

futuros e garantir um fornecimento confiável e sustentável de energia elétrica. Neste contexto, este estudo realiza previsões de carga para a série histórica de consumo de energia no Brasil, adotando abordagens estatísticas, como o ARIMA, e técnicas de inteligência computacional, como o MVS, de maneira separada.

1.2 OBJETIVOS

1.2.1 Objetivo Geral

O objetivo principal deste trabalho é analisar e comparar técnicas de IA com métodos tradicionais no âmbito da previsão de consumo de energia elétrica no Brasil, a fim de determinar a mais adequada ao cenário nacional.

1.2.2 Objetivos Específicos

- Comparar as técnicas de RNA e MVS com abordagens tradicionais na previsão do consumo de energia elétrica no Brasil.
- Comparar a acurácia das técnicas de IA com a abordagem padrão, utilizando as métricas de avaliação, como *Mean Absolute Error* (MAE).
- Analisar a capacidade das técnicas de IA em capturar padrões complexos e sutis nos dados de consumo de energia elétrica, permitindo previsões mais precisas e confiáveis.
- Fornecer informações relevantes que possam contribuir significativamente para o aprimoramento das práticas de previsão de energia elétrica, além de colaborar para aumentar a eficiência e promover a sustentabilidade do setor elétrico.

1.3 ESTRUTURA DO TRABALHO

O presente trabalho está dividido em seis capítulos. No primeiro capítulo, apresentamos a contextualização, os motivos e os objetivos do estudo sobre a previsão do consumo de energia elétrica no Brasil, destacando a importância do tema e a estrutura do trabalho. No segundo capítulo, realizamos uma análise sobre o consumo e a demanda de energia elétrica no Brasil. No terceiro capítulo, oferecemos uma fundamentação teórica abrangente, apresentando o embasamento necessário para a aplicação de técnicas de previsão energética globais. No quarto capítulo, detalhamos a metodologia, incluindo os procedimentos de coleta e análise de dados. No quinto capítulo, apresentamos o projeto detalhado das técnicas aplicadas em previsões energéticas, com ênfase nos modelos, hiperparâmetros ajustados, métricas de desempenho e outras considerações relativas à modelagem. No sexto e último capítulo, concluímos o trabalho apresentando uma síntese geral dos resultados obtidos ao longo do estudo.

2 CENÁRIO DE CONSUMO E DEMANDA DE ENERGIA ELÉTRICA NO BRASIL

2.1. COMPREENDENDO O CONSUMO E A DEMANDA NO SISTEMA ELÉTRICO DE POTÊNCIA NO BRASIL

A rede elétrica moderna depende do Sistema Elétrico de Potência (SEP) para garantir a transmissão, distribuição e geração eficientes de energia elétrica. Uma das características essenciais do SEP atual é o seu funcionamento trifásico (Weedy, 2012).

Ao longo da história, o desenvolvimento do SEP foi impulsionado por avanços tecnológicos e pela crescente demanda por eletricidade. No final do século XIX, a eletricidade começou a ser gerada em larga escala por meio de usinas termoeletricas e hidreletricas. A distribuição inicial da eletricidade era limitada, ocorrendo principalmente em áreas urbanas. No entanto, com o tempo, a demanda por eletricidade aumentou, o que exigiu melhorias nas redes de transmissão e distribuição (Weedy, 2012).

O consumo e a demanda de energia elétrica são conceitos fundamentais no SEP, com implicações diretas no fornecimento e no planejamento energético. Compreender a diferença entre esses dois termos é essencial para analisar o uso de energia elétrica e suas tendências. Embora complementares, esses conceitos apresentam diferenças (Oliveira; Silveira; Braga, 2000).

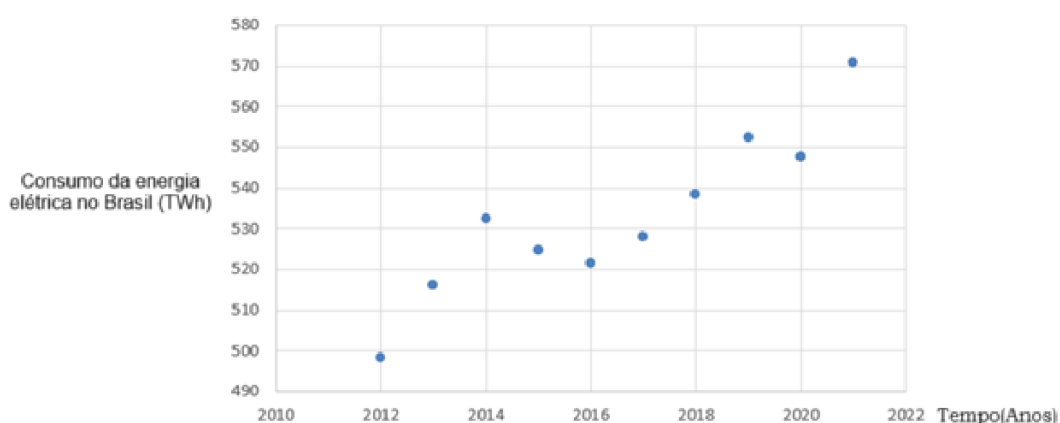
No contexto do SEP, tanto o consumo quanto a demanda de energia elétrica são considerados para garantir a estabilidade e a confiabilidade do sistema. Os operadores precisam monitorar e prever a demanda, levando em conta os diferentes perfis de consumo ao longo do dia e do ano, a fim de planejar e operar o sistema de forma eficiente. A demanda pode variar ao longo do tempo devido a fatores sazonais, mudanças climáticas, atividades econômicas e eventos imprevisíveis (Weedy, 2012).

De acordo com Santos (2004), as mudanças significativas no SEP, devido à inserção crescente da geração distribuída (GD), têm implicações diretas no

consumo de energia elétrica. Essas transformações não apenas aumentam a complexidade do SEP, mas também podem influenciar os padrões de consumo e demanda de energia elétrica.

Em relação ao consumo de energia elétrica no Brasil, dados da Empresa de Pesquisa Energética (EPE, 2023), ilustrados na Figura 1, mostram uma tendência de aumento ao longo dos anos. No entanto, é importante observar que em períodos, como os anos de 2014 a 2016, e entre 2019 e 2020 houve redução no consumo. Entre 2014 e 2016, a redução do Produto Interno Bruto (PIB) pode ter impactado negativamente o consumo de energia elétrica (Gonçalves, 2004). No entanto, a partir de 2017, com a retomada do crescimento econômico, o consumo voltou a aumentar.

Figura 1 - Consumo de energia elétrica no Brasil (TWh) entre 2012 e 2021



Fonte: elaborada pela própria autora

Esses exemplos destacam a dinamicidade do consumo de energia elétrica, que é influenciado por fatores econômicos, sociais e ambientais (Oliveira; Silveira; Braga, 2000). A introdução da geração distribuída, como a energia solar fotovoltaica, oferece uma alternativa mais sustentável e localizada para suprir parte das demandas de energia, contribuindo para a diversificação da matriz energética e para a redução das emissões de gases de efeito estufa (Gonçalves, 2004).

2.2 CONSUMO DE ENERGIA ELÉTRICA NO BRASIL POR SETOR

O consumo de energia elétrica no Brasil pode ser dividido em três classes de consumidores: residencial, comercial e industrial. Diversos estudos foram conduzidos para analisar a demanda e o consumo de energia elétrica em cada uma dessas classes, comparando os padrões de consumo e demanda em diferentes países (Hansen, 2000).

No contexto brasileiro do consumo residencial, o crescimento econômico está intimamente relacionado ao uso de energia elétrica. Esse crescimento tem sido impulsionado pelo avanço tecnológico e pela redução dos custos de eletrodomésticos, o que tem levado a melhorias na qualidade de vida e no conforto das populações. No entanto, é importante ressaltar que o aumento do consumo de energia elétrica não tem acompanhado de forma adequada o crescimento da capacidade de geração e distribuição do país (Ferreira; Borelli, 2015).

O consumo residencial de energia elétrica refere-se à utilização de eletricidade em residências e suas áreas adjacentes. Isso engloba o funcionamento de eletrodomésticos e outros dispositivos elétricos relacionados ao uso doméstico (Hansen, 2000). Segundo a pesquisa mencionada por Hobby e Tucci (2011), o consumo de eletricidade em residências é grandemente impactado pelo uso de sistemas de aquecimento e resfriamento. Em regiões com climas mais quentes e úmidos, o uso de ar-condicionado é mais intenso, levando a um aumento no consumo de eletricidade.

Segundo Schmidt e Lima (2004), diferentemente da demanda residencial, que pode ser interpretada como um problema de maximização de utilidade sujeita a uma restrição orçamentária, a demanda comercial e industrial é impulsionada pela necessidade de operar aparelhos elétricos e máquinas específicas para a produção de bens. O consumo industrial de energia elétrica pode ser descrito como um problema de minimização de custo, sujeito a um certo nível de produção.

A demanda de energia elétrica no âmbito industrial está diretamente ligada à produção e às operações das indústrias, estando relacionada à Quarta Revolução Industrial. O conceito se refere à atual transformação tecnológica em larga escala, caracterizada pela convergência de avanços em áreas como

inteligência artificial, automação e outras tecnologias. Essa transformação tem implicações no consumo de energia, uma vez que está ocorrendo a automação de processos e a adoção de tecnologias capazes de aumentar a eficiência energética e diminuir o consumo de energia elétrica (Dias *et al.*, 2023). As inovações estão relacionadas à resolução do problema de minimização de custos, no qual as empresas buscam utilizar a energia elétrica de forma eficiente para produzir (Schmidt; Lima, 2004).

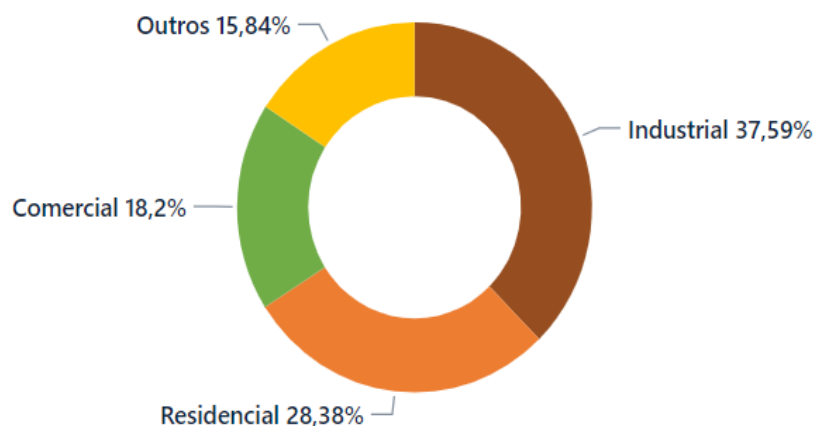
O consumo de energia no setor comercial abrange uma variedade de atividades e serviços. Esses setores dependem de energia elétrica para manter seus sistemas e equipamentos em funcionamento, desde iluminação e climatização até equipamentos específicos para cada tipo de negócio. O consumo de energia varia de acordo com o tamanho dos equipamentos, o tempo de operação diária e a eficiência energética dos estabelecimentos (Ferreira; Borelli, 2015).

A citação de Carvalho *et al.* (2010), é relevante para entender a alta demanda de energia no setor comercial, especialmente em relação ao uso de sistemas de ar condicionado e iluminação. No Brasil, os sistemas de ar condicionado representam cerca de 70% (Carvalho; La Rovere; Gonçalves, 2010) do consumo de energia em edifícios comerciais, devido à necessidade de manter ambientes confortáveis para as atividades comerciais. Além disso, os sistemas de iluminação também desempenham um papel significativo, sendo responsáveis por até 86% (Carvalho; La Rovere; Gonçalves, 2010) do consumo total de energia em locais como bancos e escritórios. Essa alta demanda evidencia a importância de adotar estratégias de eficiência energética e tecnologias mais sustentáveis para minimizar o impacto ambiental e os custos com energia no setor comercial.

A Figura 2 (EPE, [202-?]) ilustra o consumo de energia elétrica por setor entre 2010 e 2022. Observa-se que o setor industrial é o maior consumidor de energia elétrica; no entanto, quando analisado em termos percentuais, há uma tendência de diminuição dessa participação. Em 2004, o setor industrial representava 47,1% do consumo total de energia elétrica, enquanto, no período de 2010 a 2022, sua participação mostrou uma redução. Em contraste, o setor

residencial apresentou o maior crescimento, com um aumento de aproximadamente 7% entre 2004 e o período de 2010 a 2022 (EPE, [202-?]).

Figura 2 - Consumo de energia elétrica no Brasil por setor entre 2010 e 2022



Fonte: EPE, [202-?]

Com base na Figura 2, podemos concluir que o setor industrial continua a deter a maior demanda de energia elétrica no Brasil. No entanto, observa-se uma tendência de diminuição dessa participação percentual quando comparado ao ano de 2002. Em contraste, o setor residencial experimenta um crescimento expressivo, ampliando sua participação no consumo total de energia elétrica. Essa mudança destaca a crescente importância do setor residencial no consumo energético do país, refletindo transformações no perfil de consumo e, possivelmente, a evolução das condições socioeconômicas e demográficas.

É importante destacar a necessidade de acompanhar e planejar adequadamente o suprimento de energia elétrica para atender a essa crescente demanda residencial, considerando a eficiência energética e a sustentabilidade como diretrizes fundamentais. A análise do consumo de energia elétrica por setor evidencia uma mudança na composição do consumo ao longo do tempo, com uma diminuição relativa do setor industrial e um crescimento expressivo do setor residencial.

2.3 INFLUÊNCIAS NO CONSUMO DE ENERGIA ELÉTRICA

Diversos aspectos exercem influência sobre o consumo de energia elétrica, afetando diretamente as escolhas e demandas dos consumidores. Podemos categorizar esses fatores em dois grupos principais: aqueles relacionados à economia e demografia, e aqueles específicos do setor elétrico (Abrahão, 2020).

Uma das influências mais evidentes é a variação nos preços das fontes energéticas e a disponibilidade de recursos naturais. Segundo Rodrigues *et al.* (2015), a matriz energética de um país está sujeita a mudanças devido às flutuações nos preços das fontes de energia e à quantidade de recursos naturais disponíveis. A disponibilidade de recursos naturais, como rios para geração hidrelétrica, influencia a diversidade de fontes energéticas utilizadas. O governo, em resposta à diminuição da acessibilidade financeira das fontes energéticas, muitas vezes opta por aumentar o valor da taxa de luz, o que pode ter um impacto direto no consumo de energia elétrica (Correia-Silva; Rodrigues, 2015).

Os preços das despesas com energia elétrica têm um impacto direto no consumo, particularmente nos setores comercial e residencial, pois variações nos preços podem afetar a demanda por energia elétrica e estimular práticas mais eficientes (Oliveira; Silveira; Braga, 2000). A maioria dos consumidores brasileiros, cerca de 85%, está tomando medidas para reduzir o consumo de energia, priorizando também a diminuição dos custos energéticos, de acordo com relatório da empresa EY (do inglês, Ernst & Young) (Brasil, 2023).

Outra influência importante é o desenvolvimento de tecnologias de produção e consumo eficientes. A crescente preocupação com questões ambientais tem levado a uma busca cada vez maior por eficiência no uso dos recursos naturais. O desenvolvimento de tecnologias mais eficientes para a produção e o consumo de energia elétrica pode reduzir o consumo total de energia, uma vez que equipamentos mais eficientes consomem menos energia para realizar as mesmas tarefas (Abrahão, 2020).

O contexto político e as questões ambientais também desempenham um papel significativo no consumo de energia elétrica. Políticas governamentais podem incentivar o uso de fontes renováveis de energia, como foi o caso do Programa Nacional do Álcool (Proálcool) no Brasil (Correia-Silva; Rodrigues,

2015). A conscientização ambiental da população influencia as escolhas individuais em relação ao consumo de energia elétrica, levando a práticas mais sustentáveis (Abrahão, 2020).

O desenvolvimento socioeconômico de um país está intimamente ligado à demanda por energia elétrica. À medida que a economia cresce e as condições de vida da população melhoram, a demanda por energia elétrica tende a aumentar. Isso ocorre devido ao maior consumo de bens e serviços, bem como ao uso mais frequente de equipamentos elétricos e eletrônicos (Francisco, 2010). É importante ressaltar que essas influências são complexas e inter-relacionadas. As variações nos preços das fontes energéticas, a disponibilidade de recursos naturais, o desenvolvimento tecnológico, as políticas governamentais e o contexto socioeconômico são apenas algumas das muitas variáveis que podem afetar o consumo de energia elétrica. Compreender essas influências e promover práticas sustentáveis de consumo de energia elétrica são desafios fundamentais para garantir o uso eficiente dos recursos e a sustentabilidade ambiental (Altoé *et al.*, 2017).

2.4 EFICIÊNCIA ENERGÉTICA E O CONSUMO

Eficiência energética é um princípio e conjunto de práticas que têm como objetivo otimizar o uso dos recursos energéticos, visando alcançar um desempenho máximo e resultados finais com o menor consumo de energia possível (Ferreira; Borelli, 2015). Essa abordagem abrange a adoção de tecnologias mais eficientes, o aprimoramento dos processos produtivos, a implementação de políticas públicas e a conscientização dos consumidores. Seu propósito é assegurar um uso sustentável de energia, reduzindo custos, minimizando impactos ambientais e promovendo a segurança energética. Ao buscar a eficiência energética, busca-se um equilíbrio entre a demanda crescente por energia e a necessidade de preservação dos recursos naturais e do meio ambiente (Altoé *et al.*, 2017).

No Brasil, foi criado o Programa de Eficiência Energética (PEE) da Agência Nacional de Energia Elétrica (ANEEL) pela Lei nº 9.991/2000, com o objetivo de promover o uso eficiente de energia elétrica em todos os setores da economia. Seu propósito é demonstrar a importância e a viabilidade econômica da melhoria da eficiência energética em equipamentos, processos e usos finais de energia (Ferreira; Borelli, 2015).

Outro exemplo sobre como promover a eficiência energética utilizando políticas e programas governamentais é o Programa Nacional de Conservação de Energia Elétrica (Procel). O seu objetivo é promover o uso racional de energia elétrica e estimular a adoção de tecnologias mais eficientes. Atua em diversas frentes, como certificação de equipamentos, etiquetagem de eficiência energética, capacitação técnica e campanhas de conscientização (Brasil, 2023). O Selo Procel de Economia de Energia, que certifica produtos e equipamentos que apresentam alto desempenho em termos de eficiência energética (Ferreira; Borelli, 2015), auxilia os consumidores na escolha de produtos mais eficientes e contribui para a redução do consumo de energia elétrica no país.

No âmbito das novas tecnologias, podemos citar a Robótica Avançada e a IA; ambos têm desempenhado um papel cada vez mais relevante na busca pela eficiência energética. Utilizando a flexibilidade e adaptabilidade dos robôs, é possível automatizar processos industriais, substituindo tarefas que demandam maior consumo de energia por operações mais eficientes e precisas (Dias *et al.*, 2023). A aplicação de algoritmos de inteligência artificial permite o desenvolvimento de sistemas inteligentes capazes de otimizar o uso de energia em tempo real, levando em consideração as demandas e as condições específicas do ambiente (Mat Daut *et al.*, 2017). Essas tecnologias contribuem significativamente para a redução do consumo energético e para a promoção de uma produção mais sustentável.

3 FUNDAMENTAÇÃO TEÓRICA

3.1 SÉRIES TEMPORAIS: CONCEITOS E CLASSIFICAÇÃO

Uma série temporal representa a evolução de uma variável ao longo do tempo, com observações registradas em momentos distintos. Em outras palavras, é uma coleção de pontos de dados coletados em intervalos regulares ao longo do tempo. A análise de séries temporais envolve o estudo dos padrões e tendências nos dados para fazer previsões sobre valores futuros (Brown, 1959).

As séries temporais podem ser divididas em três tipos principais: multivariadas, contínuas e discretas (Hamilton, 1994). Em séries temporais multivariadas, várias variáveis são observadas simultaneamente ao longo do tempo, permitindo a análise de suas interações complexas. Utilizadas em monitoramentos em tempo real, as séries temporais contínuas são compostas por observações constantes e sem intervalos fixos ao longo do tempo (Kirchgässner; Wolters; Hassler, 2013). As séries temporais discretas possuem índices discretos e representam observações feitas em pontos específicos no tempo. Já as séries temporais contínuas têm índices contínuos e representam observações registradas ao longo de um intervalo contínuo de tempo. Por sua vez, as séries temporais multivariadas consistem em múltiplas variáveis de interesse e requerem métodos estatísticos específicos para análise e previsão (Brown, 1959).

O objetivo da análise de séries temporais é compreender como a série é gerada e realizar previsões sobre seu comportamento futuro, empregando modelos apropriados. É uma ferramenta útil para a tomada de decisões e o planejamento estratégico em diversas áreas, permitindo identificar tendências, sazonalidades e anomalias, além de avaliar o impacto de eventos externos na série. A escolha do modelo mais adequado depende dos objetivos da análise e do tipo de série temporal em questão (Hamilton, 1994).

3.1.1 Características da série temporal no consumo de energia elétrica

Uma série temporal é composta por diferentes características que descrevem os padrões presentes nos dados ao longo do tempo: tendência, sazonalidade e ciclo (Brockwell; Davis, 2002). A tendência é um padrão de longo prazo observado na série temporal, podendo ser positiva, indicando um aumento gradual ao longo do tempo, ou negativa, representando uma queda gradual (Holt, 2004). No contexto do consumo de energia elétrica, a tendência pode ser observada no aumento progressivo da eficiência energética em equipamentos e sistemas elétricos, o que leva a uma diminuição na demanda e no preço de energia para a mesma produção ou serviço (Nogales *et al.*, 2002).

A sazonalidade ocorre quando a série temporal apresenta flutuações regulares em determinados períodos de tempo, como mensalmente ou trimestralmente. No consumo de energia, a sazonalidade é influenciada pelo clima e pelos hábitos de consumo das pessoas, apresentando um aumento durante os meses mais frios ou mais quentes devido ao uso de aquecedores ou ar-condicionado, por exemplo (Brown, 1959).

As flutuações irregulares no consumo de energia elétrica ao longo do dia, com aumento durante as horas de pico de atividade e queda durante a noite, quando o consumo é menor, não são consideradas ciclos, mas sim variações diárias (Morettin; Toloj, 2020). Os ciclos, por sua vez, ocorrem em intervalos de tempo de maneira irregular, mas não estão necessariamente associados a medidas temporais específicas.

O componente aleatório é imprevisível e representa a variabilidade aleatória presente na série temporal. É a parte que não pode ser explicada pela tendência ou pela sazonalidade e pode ser causada por flutuações aleatórias ou erros de medição (Kirchgässner; Wolters; Hassler, 2013). No consumo de energia, o componente aleatório pode ser observado em eventos imprevisíveis, como a ocorrência de apagões ou falhas no fornecimento de energia elétrica (EPE, [202-?]).

A necessidade de identificar padrões históricos e fatores que influenciam o consumo, como sazonalidade, torna a análise de séries temporais uma técnica importante (Nogales *et al.*, 2002). As informações obtidas com a análise são utilizadas no desenvolvimento de modelos para prever com precisão a demanda futura por eletricidade em diferentes períodos de tempo, garantindo a estabilidade do sistema elétrico e evitando interrupções no fornecimento de energia.

3.2 MÉTODOS DE PREVISÃO

Os métodos de previsão são técnicas e abordagens utilizadas para fazer estimativas ou previsões sobre eventos ou quantidades futuras (Hyndman; Athanasopoulos, 2021).

Segundo Hyndman (2011), os métodos de previsão se dividem em dois grupos: quantitativos e qualitativos. Os quantitativos aplicam-se com informações numéricas disponíveis do passado, assumindo que padrões passados continuarão no futuro. Os qualitativos, por outro lado, são usados quando há falta de informações numéricas ou relações claras entre as variáveis, ajustando previsões quantitativas com base em informações não incorporadas ao modelo, utilizando abordagens estruturadas baseadas em julgamentos especializados. As previsões utilizando métodos de ambos os grupos são importantes para vários setores (Hyndman; Athanasopoulos, 2021).

No contexto da previsão de consumo de energia elétrica, a análise de séries temporais é um método quantitativo (Hyndman, 2010), utilizado em boa parte das técnicas tradicionais para entender o comportamento passado e extrapolar esses padrões para fazer previsões futuras (Mohamed; Bodger, 2005). O método de suavização exponencial é um exemplo de abordagem que usa séries temporais para previsão (DAUT *et al.*, 2017). Esse método explora as informações contidas nos dados históricos, como as tendências e os padrões sazonais, para estimar o consumo futuro de energia elétrica (Klyuev *et al.*, 2022).

Os modelos autoregressivos e de média móvel, quando combinados em um modelo Auto Regressivo de Média Móvel (ARMA), são amplamente explorados nessa abordagem (Cai, 2020). O modelo ARMA é utilizado para capturar a dependência linear entre os valores passados da carga e os valores futuros. A versão aprimorada do ARMA, conhecida como ARIMA, é aplicada quando o processo de carga é não estacionário, exigindo uma etapa de diferenciação para tornar o processo estacionário (Hyndman; Athanasopoulos, 2021).

O método de regressão é amplamente utilizado na abordagem baseada no relacionamento entre variáveis. Esse método permite estabelecer uma relação entre uma variável dependente (resposta) e uma ou mais variáveis independentes (preditoras). Na previsão de carga elétrica, a variável dependente pode ser representada pelo consumo de energia ou preço da eletricidade, enquanto as variáveis independentes estão geralmente relacionadas ao clima, como temperatura ou precipitação (Dong; Krzyzak; Suen, 2005).

Existem duas formas comuns de regressão utilizadas nessa abordagem: Regressão Linear (RL) e Regressão Linear Múltipla (RLM). A RL é a forma mais simples, em que apenas uma variável independente é considerada. Já a regressão múltipla envolve o uso de várias variáveis independentes para estimar a variável dependente por meio de uma equação de regressão (Dong; Krzyzak; Suen, 2005). De acordo com Li *et al.* (2020), o RLM pode ser utilizado para prever os componentes de baixa frequência do consumo de energia, obtidos por meio da decomposição da carga elétrica. Esses componentes representam a tendência geral das mudanças no consumo e podem ser previstos de forma rápida pelo RLM.

Outra abordagem mencionada por Klyuev *et al.* (Klyuev *et al.*, 2022) é a análise de correlação Gray, apresentada em um estudo sobre previsão do consumo de eletricidade por setor industrial na China. O modelo de Gray é uma técnica de modelagem baseada em equações diferenciais de primeira ordem para realizar previsões (Lin; Liu, 2005). O modelo levou em consideração fatores socioeconômicos que afetam a produção industrial. Essas previsões foram

criadas para cada setor da economia, levando em consideração diferentes cenários de desenvolvimento (Xiao *et al.*, 2018).

O modelo Gray permite a construção de previsões com base em uma pequena quantidade de dados, superando métodos clássicos de previsão (Brockwell; Davis, 2002). O modelo diferencia-se dos métodos estatísticos tradicionais ao introduzir o conceito de derivadas de Gray, permitindo que se estabeleçam modelos semelhantes a equações diferenciais para sequências de dados discretos (Lin; Liu, 2005). O modelo leva em consideração fatores socioeconômicos que afetam a produção industrial (Brockwell; Davis, 2002).

A árvore de decisão é um modelo de aprendizado de máquina e um método para realizar previsões, utilizando a análise de dados e na tomada de decisões. Ela é construída com base em uma estrutura em forma de árvore, em que cada nó representa uma decisão ou um atributo, e cada ramo representa uma possível resposta ou resultado dessa decisão (Cai, 2020). A metodologia é reconhecida como uma forma eficiente de suporte à decisão em sistemas de produção. De acordo com Tso (2007), é possível coletar detalhes sobre os níveis de posse e classificações de potência dos eletrodomésticos, e os padrões de uso registrados em diários para determinar o consumo de energia em *Hong Kong*. A árvore de decisão pode ser comparada com outras abordagens, como as redes neurais, para avaliar sua precisão na previsão do consumo de energia elétrica (Tso; Yau, 2007).

Alves, Lotufo e Lopes (2013) mencionam a utilização de redes neurais como uma importante metodologia baseada em algoritmos de aprendizado de máquina. Elas são destacadas como uma abordagem poderosa para mapeamento de variáveis, capaz de capturar padrões complexos e comportamentos não lineares em dados de séries temporais (Nogales *et al.*, 2002). Outro método abordado por Li *et al.* (2012) são as MVS, o que amplia as possibilidades de análise e previsão em diferentes contextos..

Os algoritmos e técnicas abordadas anteriormente podem ser combinadas entre si e com outros métodos para obter um desempenho superior na previsão de consumo de energia elétrica. O estudo sobre previsão anual de carga elétrica propôs um modelo híbrido para prever o consumo anual de

eletricidade usando os algoritmos de LSSVM (do inglês, *Least Squares Support Vector Machine*) e de FOA (do inglês, *Fruit Fly Optimization Algorithm*). O modelo híbrido superou os modelos de regressão, o modelo LSSVM sem a aplicação de abordagens metaheurísticas para a otimização de parâmetros e outros métodos. O FOA se mostrou um poderoso método evolutivo para otimização de funções em um domínio contínuo, como confirmado no estudo descrito em (Li *et al.*, 2012).

A previsão de consumo de energia elétrica é um campo em constante evolução, impulsionado pelos avanços na área de inteligência artificial e aprendizado de máquina. A combinação de abordagens tradicionais e técnicas mais avançadas tem se mostrado promissora na obtenção de previsões mais precisas (Mohamed; Bodger, 2005). À medida que novas pesquisas são realizadas e novas técnicas são desenvolvidas, espera-se que a precisão das previsões melhore ainda mais, possibilitando uma gestão mais eficiente e sustentável do consumo de energia elétrica (DAUT *et al.*, 2017). A incorporação de métodos de inteligência artificial pode aprimorar significativamente as capacidades de previsão e análise em várias áreas de estudo (Barak; Sadegh, 2016). A seguir, será apresentada uma revisão bibliográfica mais detalhada dos métodos de previsão que serão aplicados neste trabalho.

3.3 MÉTODO DE SUAVIZAÇÃO EXPONENCIAL

Proposta no final da década de 1950 por Brown, Holt e Winters (Hyndman; Athanasopoulos, 2021), a suavização exponencial é um método de previsão amplamente utilizado em séries temporais. O método consiste em criar previsões de valores futuros por meio de médias ponderadas de observações anteriores (Winters, 1960). As observações mais recentes possuem um peso maior, atribuído por parâmetros, na determinação das previsões do que as observações mais antigas. O método pode ser dividido em três tipos diferentes: Suavização Exponencial Simples, Suavização Exponencial de Holt e Suavização Exponencial de Holt-Winters.

A Suavização Exponencial Simples é um modelo básico indicado quando não há tendência pronunciada ou sazonalidade evidente nos registros da série

(Brown, 1959). Nesses casos, esse método é adequado para obter previsões, pois se baseia na média ponderada das observações passadas, sem considerar explicitamente elementos de tendência ou sazonalidade. A fórmula matemática geral para a suavização exponencial simples é retratada na Equação 1.

$$\hat{y}_{t+1} = \alpha y_t + \alpha(1 - \alpha)y_{t-1} + \alpha(1 - \alpha)^2 y_{t-2} + \dots \quad (1)$$

No qual \hat{y}_{t+1} , é a previsão, y_t , y_{t-1} e y_{t-2} são os valores previamente observados da série temporal e α é o fator de suavização. O fator de suavização, α , é responsável por controlar a taxa de decrescimento exponencial de observações anteriores. O valor de α varia entre 0 e 1 e é calculado com base na natureza da série e no objetivo da previsão. Quanto maior o valor de α , mais peso é dado às observações recentes, e vice-versa para valores menores de α (Makridakis; Hibon, 1997).

Quando entendemos como a Suavização Exponencial Simples funciona, podemos utilizá-la como uma ferramenta inicial para prever séries temporais com características específicas. É uma abordagem simples e eficaz, mas devemos lembrar que a escolha desse modelo deve considerar a natureza dos dados e verificar se existem tendências, sazonalidades ou outros padrões relevantes na série temporal em questão (Morettin; Toloi, 2020).

Em 1957, Charles Holt ampliou o modelo de suavização exponencial simples para incluir um componente de tendência. O método de Holt pode ser expresso de forma semelhante à suavização exponencial simples (Makridakis; Hibon, 1997). Essa extensão adiciona um componente de tendência, permitindo que o modelo capture mudanças ao longo do tempo e gere previsões mais realistas. Em vez de previsões estáticas, o método considera a tendência da série temporal, resultando em uma previsão linear baseada no passo de tempo. Esse avanço tornou o método mais adequado para séries temporais com tendências visíveis (Hyndman; Athanasopoulos, 2021). As equações para calcular a previsão com componente de tendência, b_t , são dadas pela Equação 2 e Equação 3.

$$\hat{y}_{t+h} = l_t + hb_t, l_t = \alpha y_t + (1 - \alpha)(l_{t-1} + b_{t-1}) \quad (2)$$

$$b_t = \beta(l_t - l_{t-1}) + (1 - \beta)b_{t-1}. \quad (3)$$

Na Equação 3, b_t representa o componente de tendência previsto no tempo t , enquanto b_{t-1} é a tendência prevista no instante anterior. O parâmetro β é o fator de suavização de tendência, variando entre 0 e 1. A equação de tendência é calculada com base na variação passo a passo no componente de nível da série temporal. Com a inclusão do componente de tendência multiplicado pelo passo de tempo h na equação $\hat{y}_{t+h} = l_t + hb_t$, as previsões deixam de ser estáticas e passam a ser uma função linear de h (HO; XIE; GOH, 2002).

Os modelos de previsão apresentados anteriormente não consideravam a sazonalidade dos dados, o que limitava sua precisão em séries temporais com padrões sazonais. Para superar essa limitação, foi desenvolvido o método Holt-Winters, uma evolução do método de Holt que incorpora um índice sazonal às equações de previsão (Hyndman; Athanasopoulos, 2021). Esse modelo utiliza equações de suavização para estimar três componentes principais: o nível, a tendência e o componente sazonal dos dados. As equações são ajustadas por meio de parâmetros de suavização, incluindo o fator de suavização da sazonalidade (γ), além dos parâmetros de nível e tendência previamente mencionados (Winters, 1960).

Esse método apresenta dois modelos: o aditivo e o multiplicativo. A principal diferença entre eles está no tamanho das flutuações de sazonalidade. No modelo aditivo, essas flutuações são mais constantes, enquanto no modelo multiplicativo elas são proporcionais ao valor da série temporal naquele momento específico (Hyndman; Athanasopoulos, 2021).

A equação de nível calcula uma média ponderada entre a observação ajustada sazonalmente e a previsão não sazonal para cada tempo t . A equação de tendência é idêntica ao método linear de Holt. Já a equação sazonal calcula uma média ponderada entre o índice sazonal atual e o índice sazonal da mesma estação do ano passado. A frequência da sazonalidade, representada por m ,

indica o número de períodos em um ciclo sazonal. Na Tabela 1, estão apresentadas as equações gerais de nível, tendência e sazonal utilizadas em ambos os métodos (Hyndman; Athanasopoulos, 2021).

Tabela 1 - Equações do Método Aditivo e Multiplicativo do Holt-Winters

| Equação | Método Aditivo | Método Multiplicativo |
|-----------|---|--|
| Geral | $\hat{y}_{t+h t} = l_t + hb_t + s_{t+h-m(k+1)}$ | $\hat{y}_{t+h t} = (l_t + hb_t)s_{t+h-m(k+1)}$ |
| Nível | $l_t = \alpha(y_t - s_{t-m}) + (1 - \alpha)(l_{t-1} + b_{t-1})$ | $l_t = \alpha\left(\frac{y_t}{s_{t-m}}\right) + (1 - \alpha)(l_{t-1} + b_{t-1})$ |
| Tendência | $b_t = \beta(l_t - l_{t-1}) + (1 - \beta)b_{t-1}$ | $b_t = \beta(l_t - l_{t-1}) + (1 - \beta)b_{t-1}$ |
| Sazonal | $s_t = \gamma(y_t - l_{t-1} - b_{t-1}) + (1 - \gamma)s_{t-m}$ | $s_t = \gamma \frac{y_t}{(l_{t-1} + b_{t-1})} + (1 - \gamma)s_{t-m}$ |

Fonte: elaborada pela própria autora

O método de suavização exponencial é amplamente utilizado em diversas áreas, como economia, finanças e previsão de demanda, sendo uma ferramenta valiosa para auxiliar na tomada de decisões e no planejamento estratégico. Ao fornecer previsões precisas e confiáveis, esse método se torna uma opção importante para muitas empresas. No entanto, é fundamental compreender as equações e parâmetros envolvidos para aplicá-lo de forma eficaz (Werner; Ribeiro, 2003).

3.4 MODELO ARIMA

O modelo ARIMA é uma classe de modelos matemáticos utilizados para analisar e prever séries temporais. O método combina componentes autoregressivos, de média móvel e de diferenciação (Werner; Ribeiro, 2003). Esses componentes foram desenvolvidos para capturar a autocorrelação entre os valores da série temporal, abrangendo tendências, sazonalidades e flutuações aleatórias nos dados. Isso possibilita realizar previsões futuras com base no comportamento histórico da série (Werner; Ribeiro, 2003). O modelo ARIMA é amplamente aplicado em diversas áreas, como economia e engenharia, para analisar tendências, padrões sazonais e comportamentos complexos em dados sequenciais ao longo do tempo (Brown, 1959).

Os termos autoregressivo e de média móvel são componentes fundamentais do modelo ARIMA e são usados para modelar a dependência linear entre as observações de uma série temporal. O termo autoregressivo é responsável por capturar a relação linear entre os valores passados e presentes da série temporal. O termo de média móvel é utilizado para capturar a relação linear entre os erros passados e presentes da série temporal (Morettin; Toloi, 2020).

A estacionaridade do método é utilizada para prever o futuro com base em dados anteriores, assumindo que a média e a variância dos valores são constantes ao longo do tempo, utilizando conceitos de diferenciação (Barak; Sadegh, 2016). A estacionaridade é importante em séries temporais, pois muitos modelos de previsão assumem que a série é estacionária. No entanto, muitas séries temporais apresentam tendências e sazonalidades que as tornam não estacionárias. A diferenciação é um processo de transformação da série temporal para torná-la estacionária. O método mais comum envolve subtrair cada valor da série pelo valor anterior (Meloni, 2017). O modelo matemático resultante do método é a Equação 4.

$$y'_t = c + \phi_1 y'_{t-1} + \theta_1 \varepsilon_{t-1} + \dots + \theta_q \varepsilon_{t-q} + \varepsilon_t. \quad (4)$$

No qual, y'_t , representa a série diferenciada e captura $\phi_1 y'_{t-1}$, que reflete a dependência do valor anterior de y . O termo $\theta_1 \varepsilon_{t-1}$ considera o impacto dos erros passados sobre a variável atual, enquanto ε é o erro branco no tempo t . (Hyndman; Athanasopoulos, 2021). O modelo ARIMA utiliza três valores inteiros para realizar a parametrização: p , d e q . Esses valores são importantes porque definem as características da série temporal que serão consideradas no modelo, e variam de acordo com o cenário (Berwick, 200-?).

1. Parâmetro p : O número de termos auto-regressivos (Hyndman; Athanasopoulos, 2021). No contexto da previsão de consumo de energia, o parâmetro p pode representar o número de valores anteriores de consumo de energia que são utilizados para prever o consumo futuro. Por exemplo, se p for igual a 2, o modelo ARIMA levará em consideração os

dois valores anteriores de consumo de energia para prever o consumo futuro (Box *et al.*, 2015).

2. Parâmetro d : O número de vezes que a série temporal é diferenciada. Na previsão do consumo de energia, o parâmetro d denota a quantidade de diferenciações realizadas na série temporal de consumo de energia para alcançar a estacionariedade (Barak; Sadegh, 2016). A estacionariedade é importante para a aplicação do modelo ARIMA, pois permite que os padrões da série temporal sejam mais facilmente identificados e utilizados para prever valores futuros (Hyndman; Athanasopoulos, 2021).

3. Parâmetro q : O número de termos de média móvel (Hyndman; Athanasopoulos, 2021). Na previsão do consumo de energia, o parâmetro q pode corresponder ao número de erros passados da série temporal de consumo de energia utilizados para fazer a previsão do consumo futuro (Barak; Sadegh, 2016). Por exemplo, se q for igual a 1, o modelo ARIMA levará em consideração o erro mais recente da previsão anterior para prever o consumo futuro (Box *et al.*, 2015).

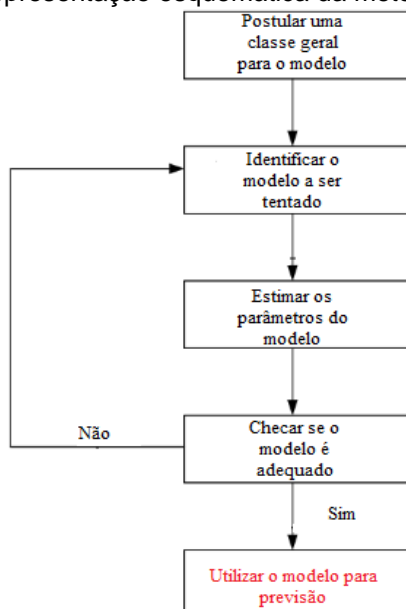
Para definir os parâmetros, é necessário realizar uma análise cuidadosa dos dados, identificar a ordem dos termos auto-regressivos, integrados e de média móvel, e avaliar a estacionariedade e a presença de sazonalidade na série temporal. A metodologia Box-Jenkins fornece um conjunto de técnicas e critérios para auxiliar na escolha dos parâmetros mais adequados para o modelo ARIMA (Makridakis; Hibon, 1997).

A metodologia Box-Jenkins é uma abordagem iterativa para construção de modelos de previsão de séries temporais. Ela é dividida em quatro etapas principais: identificação, estimação, verificação e previsão. Primeiramente, na etapa de identificação, são exploradas as funções ACF (do inglês, *Autocorrelation Function*) e PACF (do inglês, *Partial Autocorrelation Function*) para determinar o modelo mais adequado que descreve o comportamento da série temporal. Em seguida, na etapa de estimação, os parâmetros do modelo ARIMA são estimados (HO; XIE; GOH, 2002).

A partir das análises realizadas nas duas primeiras etapas, é possível determinar o valor de d , necessário para tornar a série estacionária. Em seguida, é possível utilizar as funções de ACF e a de PACF para identificar os valores de p e q . O valor de p é determinado pelo número de termos de autocorrelação significativos no autocorrelação parcial, enquanto o valor de q é determinado pelo número de termos de autocorrelação significativos na autocorrelação (Berwick, 200-?). É importante ressaltar que a escolha dos valores de p , d e q pode ser um processo iterativo, sendo recomendável testar diferentes combinações de parâmetros e avaliar o desempenho do modelo por meio de métricas como o MAE e o MSE (Berwick, 200-?).

Em seguida, realiza-se a etapa de verificação, na qual avalia-se se o modelo estimado consegue representar de maneira adequada o comportamento dos dados. Essa verificação é crucial, pois, caso o modelo não seja satisfatório, é necessário retornar à fase de identificação e repetir o ciclo até que um modelo adequado seja obtido (Makridakis; Hibon, 1997).

Figura 3 - Representação esquemática da metodologia de Box-Jenkins



Fonte: Adaptado de Makridakis e Hibon (1997)

Na Figura 3, é apresentada a representação esquemática do modelo de Box-Jenkins, que ilustra as principais etapas do processo de análise e previsão

de séries temporais. Essa metodologia oferece uma abordagem sistemática e iterativa para a análise e previsão de séries temporais, proporcionando resultados confiáveis que apoiam a tomada de decisões (Makridakis; Hibon, 1997).

Embora o ARIMA seja uma técnica poderosa, sua eficácia depende da qualidade dos dados e da escolha adequada dos parâmetros do modelo. É importante lembrar que as previsões geradas pelo ARIMA são apenas estimativas e não garantem resultados precisos. É uma ferramenta útil para análise e previsão de séries temporais, mas deve ser utilizada com cuidado e avaliada juntamente com outras técnicas e informações relevantes (Barak; Sadegh, 2016).

3.5 MÁQUINAS DE VETORES DE SUPORTE

De acordo com Jakkula (2006), as Máquinas de Vetores de Suporte foram propostas pela primeira vez em 1992 e são um tipo de modelo de aprendizado de máquina supervisionado. A principal ideia por trás das MVS é encontrar um hiperplano de separação que possa dividir os dados em diferentes classes (Haykin, 1998). Para entender melhor esse princípio, é importante compreender o conceito de hiperplano de separação e sua importância na classificação.

Um hiperplano é uma superfície que divide o espaço nas dimensões nas quais os dados estão localizados. Em um problema de classificação binária, existem apenas duas classes, e um hiperplano de separação pode ser uma linha, um plano ou uma superfície que separa as classes de dados. É importante ressaltar que o hiperplano não é uma linha que passa necessariamente pelos dados, mas sim uma superfície que separa as classes (Du *et al.*, 2018).

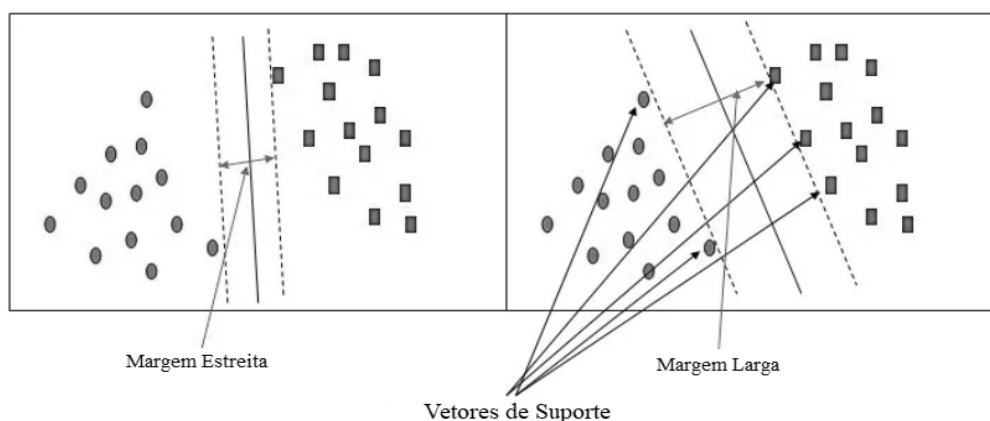
A margem de separação é a distância entre o hiperplano de separação e os pontos mais próximos de cada classe. O objetivo das MVS é encontrar o hiperplano que maximize essa margem de separação. Isso é considerado importante porque um maior espaço entre as classes resulta em uma melhor generalização do modelo e maior tolerância a erros de classificação (Meloni,

2017). O princípio da margem máxima é o conceito central das MVS. Para isso, as MVS maximizam a margem enquanto garantem que o hiperplano de separação não cometa erros de classificação nesses pontos mais próximos de cada classe, chamados vetores de suporte (Mcculloch; Pitts, 1943).

Os vetores de suporte são os pontos de dados de treinamento que estão mais próximos do hiperplano de separação. Esses pontos são críticos para a determinação do hiperplano, pois definem sua posição e orientação. Os vetores de suporte são chamados assim porque são os pontos que "sustentam" ou "suportam" o hiperplano de separação (Lorena; Carvalho, 2003).

A localização dos vetores de suporte é fundamental na MVS, pois eles influenciam o cálculo e a determinação do hiperplano de separação. Todos os outros pontos de treinamento que não são vetores de suporte não afetam diretamente o posicionamento do hiperplano, mas afetam a margem e a capacidade de generalização do modelo (Meloni, 2017).

Figura 4 - Representação de MVS com margem larga e MVS com margem estreita



Fonte: elaborada pela própria autora

Na Figura 4, podemos observar duas classes de pontos de dados diferentes, sendo que cada classe é representada por uma cor distinta. No lado direito, temos uma representação com uma margem estreita (Berwick, 200-?). Isso pode causar *overfitting* (sobreajuste) no modelo, ou seja, ele pode se ajustar excessivamente aos dados de treinamento, mas apresentar um desempenho inferior ao lidar com dados novos ou de teste. Na imagem à direita, vemos uma máquina de vetores de suporte com uma margem larga, e podemos visualizar

os vetores de suporte, o que significa que é possível separar as classes de pontos de dados de maneira mais eficaz. Essa margem mais ampla permite que a máquina de vetores de suporte seja mais tolerante a erros de classificação e consiga classificar mais pontos de dados corretamente (Haykin, 1998).

Uma das principais vantagens das MVS é a capacidade de lidar com dados de alta dimensão, o que significa que elas podem lidar com conjuntos de dados com muitas características. O hiperplano maximizado permite que as MVS manipulem dados não lineares. Além disso, elas são eficientes em termos de tempo de execução, o que significa que podem ser usadas em grandes conjuntos de dados. Isso é possível porque as MVS utilizam apenas um subconjunto dos pontos de dados para encontrar o hiperplano que separa as classes, utilizando os vetores de suporte (Lorena; Carvalho, 2003).

Para lidar com espaços de alta dimensão, as MVS utilizam o conceito de *kernels*, que são funções que mapeiam os dados para espaços de alta dimensão (Haykin, 2002). Isso permite que a MVS encontre um hiperplano que separa os dados em um espaço dimensional superior, mesmo que os dados originais não sejam linearmente separáveis. Existem vários tipos de *kernels*, incluindo *kernels* lineares e polinomiais. Cada tipo de *kernel* tem suas próprias vantagens e desvantagens, e diferentes *kernels* podem ser usados dependendo das características dos dados.

No entanto, é importante experimentar diferentes tipos de *kernels* para encontrar o que funciona melhor para o problema específico (Berwick, 200-?). A capacidade de classificação das MVS em diversas categorias torna essa técnica útil em várias aplicações, incluindo a previsão de diferentes grandezas, como o consumo de energia. Ao lidar com dados complexos e variados, as MVS não fazem suposições rígidas sobre a distribuição dos dados, o que as torna uma abordagem valiosa em cenários em que os padrões podem ser altamente não lineares e imprevisíveis (Fleck *et al.*, 2016)..

A utilização das MVS na previsão de diferentes grandezas é uma ferramenta inestimável, uma vez que permite a criação de modelos resilientes diante de imperfeições e erros de medição, proporcionando resultados mais precisos e confiáveis. Sua capacidade de prever várias categorias de dados

ajuda a otimizar o gerenciamento dos recursos e a tomar decisões estratégicas em diversas áreas (Ahmad *et al.*, 2014). É uma abordagem poderosa e confiável para enfrentar os desafios complexos da previsão, oferecendo uma contribuição significativa para a eficiência e confiabilidade dos sistemas. Sua aplicação em diversas áreas evidencia sua extrema relevância, proporcionando benefícios significativos para a sociedade em várias situações.

3.6 REDE NEURAL ARTIFICIAL

O cérebro humano é composto por bilhões de neurônios que se comunicam por meio de sinapses excitatórias e inibitórias. Quando o efeito cumulativo das sinapses excitatórias que chegam a um neurônio excede um valor limite, o neurônio dispara e envia um sinal para outros neurônios. Esse processo de comunicação entre neurônios é fundamental para a realização de funções como pensamento, movimento, percepção e emoção. O cérebro humano é capaz de se adaptar e mudar ao longo do tempo, criando novas conexões sinápticas e modificando as existentes em resposta a estímulos e experiências (Haykin, 1998).

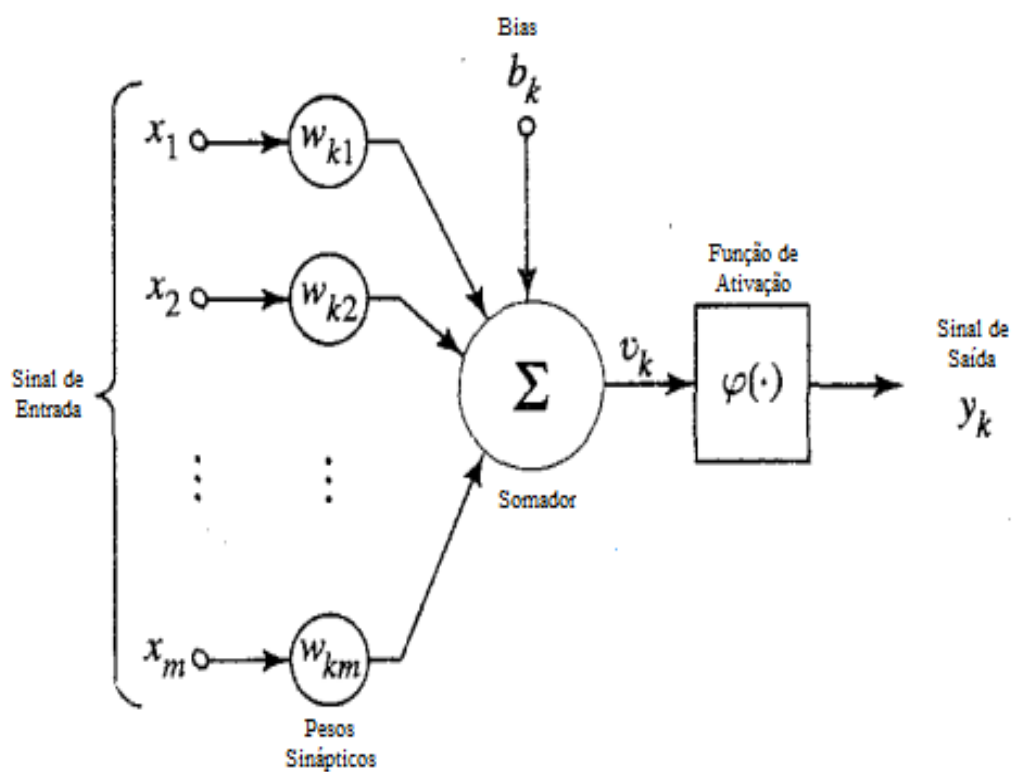
De acordo com Haykin (1998), as redes neurais são modelos computacionais inspirados no cérebro humano, capazes de aprender e executar tarefas complexas, como classificação, reconhecimento de padrões e previsão. A história dessas redes remonta aos anos 1940, quando Warren McCulloch e Walter Pitts publicaram um artigo pioneiro descrevendo um modelo de neurônio artificial que serviria de base para a construção de redes neurais. Desde então, o desenvolvimento e a aplicação das redes neurais têm sido uma área de grande interesse e progresso na ciência e tecnologia (McCulloch; Pitts, 1943).

A estrutura fundamental de uma rede neural é composta por neurônios, camadas e conexões. Os neurônios desempenham o papel crucial de unidades de processamento nas redes neurais. Eles atuam como receptores de um ou mais sinais de entrada, realizam cálculos e transformações internas nesses sinais e, em resposta, geram um sinal de saída correspondente. As camadas são agrupamentos de neurônios responsáveis por trabalhar em conjunto para

processar os sinais de entrada e produzir sinais de saída, os quais são encaminhados para a próxima camada. As conexões, por sua vez, são os canais que interligam os neurônios, permitindo a transmissão dos sinais de uma camada para outra (Du *et al.*, 2018).

A representação de um neurônio artificial é composta por diversos elementos, como é possível observar na Figura 5. Primeiramente, o neurônio recebe um conjunto de valores de entrada, representados por um vetor de números, sendo que cada valor é associado a um peso sináptico. Esses pesos indicam a importância relativa de cada dado para o processamento do neurônio, podendo ser positivos ou negativos (Haykin, 1998).

Figura 5 - Representação de um neurônio artificial



Fonte: Adaptado de Haykin (1998)

Após receber as entradas ponderadas pelos pesos, o neurônio realiza uma operação de agregação, somando os valores resultantes. Essa soma ponderada é o produto da combinação linear das entradas e pesos sinápticos, e ela representa a "excitação" ou "inibição" do neurônio com base nas informações recebidas (Hyndman; Athanasopoulos, 2021).

Em seguida, é aplicada uma função de ativação que, de acordo com Fleck *et al.* (2016), estabelece o limite para os valores de saída de um neurônio. Uma função de ativação é aplicada a cada neurônio para determinar se ele deve ser ativado ou não. Ela introduz não-linearidade na rede, o que é importante para que a rede possa aprender a partir de dados não lineares. Existem várias funções de ativação que podem ser utilizadas em redes neurais, como a função sigmoide e a função tangente hiperbólica. A escolha da função de ativação pode ter um impacto significativo no desempenho da rede neural (Haykin, 1998). A saída do neurônio é, então, o resultado da aplicação da função de ativação. Esse valor de saída será transmitido para os neurônios da próxima camada da rede neural, caso existam, dando continuidade ao processamento das informações.

Outro parâmetro importante é o *bias* (viés), presente no neurônio. O viés é adicionado à soma ponderada antes da aplicação da função de ativação, permitindo ajustar o ponto de partida da função. Isso proporciona maior controle sobre o comportamento e a flexibilidade do neurônio durante o processo de aprendizado de padrões complexos. A inclusão do viés é essencial para garantir que o neurônio possa se adaptar melhor aos dados de entrada e realizar suas tarefas de forma mais eficiente e precisa (Du *et al.*, 2018).

Em conjunto, essas etapas de processamento do neurônio artificial constituem a base do funcionamento das redes neurais, permitindo a realização de tarefas de aprendizado, classificação, previsão e resolução de problemas em diversas áreas da inteligência artificial. Conforme Wilamowski (2009) e Seabold (2010), os tipos de arquiteturas de redes neurais citados são:

- *Multilayer Perceptron* (MLP): A arquitetura MLP é a mais antiga e uma das mais comumente utilizadas entre as topologias de redes neurais. Ela consiste em várias camadas de neurônios interconectados, incluindo uma camada de entrada, uma ou mais camadas ocultas e uma camada de saída. No entanto, as arquiteturas MLP raramente produzem resultados satisfatórios.
- *Bridged Multilayer Perceptron* (BMLP): A arquitetura BMLP é uma variação do MLP em que são permitidas conexões entre as camadas,

representadas por linhas pontilhadas na rede. Essa arquitetura é considerada mais poderosa do que o MLP tradicional.

- *Fully Connected Cascade (FCC)*: A arquitetura FCC é outro tipo de topologia de rede neural que inclui apenas conexões diretas (*feed-forward*). Ela é relativamente simples e eficiente, especialmente para problemas menores.
- *Radial Basis Function (RBF), Counterpropagation, ou Learning Vector Quantization (LVQ) networks*: As RBF usam funções de base radial como ativação, sendo adequadas para aproximação de funções e classificação. As redes de contrapropagação possuem uma camada de entrada e uma camada competitiva de saída, usadas para agrupamento e classificação. Já as redes LVQ são supervisionadas e úteis no reconhecimento de padrões, treinadas para mapear vetores de entrada em classes específicas.
- *Recurrent Neural Network (RNN)*: As RNN são um tipo de rede neural artificial que possui uma memória interna e são projetada para lidar com dados sequenciais, como séries temporais, texto e fala, nos quais a ordem dos dados é importante. Essa memória interna permite que as RNN considerem informações anteriores para afetar as saídas atuais, tornando-as adequadas para tarefas sequenciais. O texto parece se concentrar na explicação dos fundamentos das RNN, incluindo sua formulação formal a partir de equações diferenciais e técnicas associadas, como o "*unrolling*" (desenrolamento) de uma RNN.

A escolha da arquitetura de uma rede neural pode influenciar significativamente o processo de treinamento, sendo essencial considerar diferentes topologias e algoritmos de aprendizado para alcançar resultados otimizados (Wilamowski, 2009). Os algoritmos de treinamento são responsáveis por ajustar os pesos das conexões entre os neurônios de modo que a rede possa aprender a partir de dados de entrada (Haykin, 1998).

4 METODOLOGIA

Neste capítulo apresentaremos a metodologia utilizada no trabalho para prever o consumo de energia elétrica no Brasil, utilizando a aplicação de quatro técnicas amplamente reconhecidas: Suavização Exponencial, ARIMA, MVS e RNA.

4.1 PROCEDIMENTOS METODOLÓGICOS

Para propor o procedimento metodológico do trabalho, foi essencial considerar os pontos de convergência e as semelhanças presentes em cada um dos métodos selecionados: Suavização Exponencial, ARIMA, MVS e RNA. Ao utilizar as mesmas séries temporais para aplicar essas técnicas, buscamos obter uma avaliação justa e comparativa do desempenho de cada modelo na previsão do consumo de energia elétrica no Brasil.

Ao analisar os pontos em comum entre os métodos, pudemos identificar as etapas fundamentais que compõem o procedimento metodológico. Iniciamos com a escolha das ferramentas e bibliotecas apropriadas para a aplicação das técnicas de previsão selecionadas. Em seguida, a coleta de dados sobre o consumo de energia elétrica no Brasil. Prosseguimos com o pré-processamento e análise dos dados, nas quais nos certificamos de preparar as informações de maneira adequada para o treinamento dos modelos.

Ao utilizar as mesmas séries temporais, garantimos a igualdade nas condições de teste e avaliação de cada modelo. Isso é fundamental para uma análise do desempenho relativo de cada técnica e possibilita uma escolha mais fundamentada do modelo mais adequado para a previsão do consumo de energia elétrica.

Considerar os pontos semelhantes entre os métodos também nos permitiu identificar as particularidades de cada abordagem. Dessa forma, foi possível explorar as vantagens e limitações individuais de cada técnica, bem como compreender como elas podem se complementar para fornecer resultados mais robustos e precisos.

Ao levar em conta os pontos de convergência e semelhanças entre as técnicas e ao utilizar as mesmas séries temporais para sua aplicação, pudemos desenvolver um procedimento metodológico coeso e abrangente para prever o consumo de energia elétrica no Brasil. Essa abordagem nos permite avaliar as técnicas de forma justa e embasar nossas conclusões de maneira sólida, garantindo a confiabilidade dos resultados obtidos no trabalho.

4.2 BIBLIOTECAS E FERRAMENTAS PARA ANÁLISE DE DADOS

Ao escolher ferramentas para aplicar algoritmos em análise de dados e aprendizado de máquina com *Python*, foi essencial considerar a natureza e o tamanho dos conjuntos de dados, os tipos de análises a serem realizadas e os requisitos específicos do projeto. A comunidade *Python* oferece uma ampla variedade de bibliotecas e ferramentas que podem ser combinadas para atender às necessidades específicas de cada tarefa (Stančin; Jović, 2019).

O *Python* é uma linguagem de programação extremamente versátil, sendo amplamente utilizada na comunidade de ciência de dados e aprendizado de máquina. Sua sintaxe simples e legível torna o código mais compreensível e facilita a colaboração em projetos complexos (Stančin; Jović, 2019).

Ao aplicar técnicas de previsão em *Python*, utilizaremos algumas bibliotecas para implementação. Abaixo está uma lista das bibliotecas utilizadas e uma breve descrição de cada uma delas, segundo Seabold (2010) e Stančin (2019):

- *Statsmodels*: Oferece uma implementação robusta e flexível de modelos estatísticos, incluindo modelos de suavização exponencial e modelos ARIMA. Ela também oferece recursos adicionais, como diagnósticos de resíduos e previsões de intervalo de confiança (Seabold; Perktold, 2010).
- *Scikit-learn*: Oferece uma variedade de algoritmos para tarefas como classificação. A biblioteca é valorizada por sua eficiência e facilidade de uso, tornando o processo de implementação de algoritmos

de aprendizado de máquina mais acessível para os usuários (Stančin; Jović, 2019).

- *Keras*: Focada em redes neurais, e é conhecida por ser fácil de usar, especialmente para a construção e prototipagem de redes neurais. *Keras* fornece uma *Application Programming Interface* (“Interface de Programação de Aplicação”) de alto nível, que pode ser executada em cima de outras bibliotecas de aprendizado de máquina, como o *TensorFlow*, permitindo criar redes neurais de forma rápida e intuitiva (Stančin; Jović, 2019).
- *Pandas*: É essencial para análise de dados em *Python*. Ela oferece estruturas de dados poderosas, como a *Series* e o *DataFrame*, que permitem manipular e analisar dados de maneira eficiente. Com o *Pandas*, é possível carregar dados de diferentes fontes, prepará-los, limpar dados ausentes, fazer seleções, filtrações e agregações, tornando todo o processo de análise de dados mais produtivo (Stančin; Jović, 2019).
- *NumPy*: Oferece suporte para *arrays* multidimensionais eficientes e funções matemáticas de alto desempenho para trabalhar com esses *arrays*. O *NumPy* é amplamente utilizado para realizar operações numéricas em larga escala, sendo a base para várias outras bibliotecas de ciência de dados (Stančin; Jović, 2019).
- *Matplotlib*: É uma biblioteca para criação de gráficos em *Python*. Ela permite criar uma ampla variedade de gráficos estáticos. É uma biblioteca robusta que facilita a visualização de dados e resultados de análises de forma clara e informativa (Stančin; Jović, 2019).
- *Plotly*: É outra biblioteca de visualização em *Python*. No entanto, diferentemente do *Matplotlib*, ela permite criar gráficos interativos e visualizações para web. Essa biblioteca é especialmente útil para criar gráficos interativos, que podem ser explorados e manipulados pelo usuário (Stančin; Jović, 2019).

Essas bibliotecas foram selecionadas com base em sua eficiência, integração e na possibilidade de utilização no Jupyter Notebook. Com o suporte

dessas bibliotecas, o Jupyter Notebook oferece uma base sólida para a exploração e pré-processamento de dados (Kluyver *et al.*, 2016).

4.3 COLETA DE DADOS PARA ANÁLISE E PREVISÃO

A coleta dos dados foi realizada a partir da base de dados da EPE (EPE, [202-?]), uma instituição governamental brasileira responsável por produzir estudos e análises no setor de energia elétrica (Batista, 2003). Essa fonte fornece informações cruciais para a análise e previsão do consumo de energia, permitindo o acesso aos dados em formatos adequados para a aplicação das técnicas de previsão.

A atualização desses dados é realizada pela equipe de Estatísticas Energéticas na Superintendência de Estudos Econômicos e Energéticos da EPE. Esse processo ocorre regularmente, com base nas informações de consumo fornecidas mensalmente pelos agentes do setor elétrico, por meio do Sistema de Acompanhamento do Mercado (SAM). As reuniões da Comissão Permanente de Análise e Acompanhamento do Mercado de energia Elétrica (COPAM) também desempenham um papel importante na coleta dos dados (EPE, [202-?]).

A base de dados da EPE contém informações no formato XML, posteriormente convertidas em um CSV. O arquivo contém uma série histórica que inicia em 2004, com informações detalhadas sobre o consumo mensal e anual de energia elétrica em níveis nacional, regional e por subsistemas. Esses dados abrangem as classes residencial, industrial, comercial e outros (Seabold; Perktold, 2010).

Para estruturar a base de dados com os 6.000 registros, divididos por região, mês e tipo de consumo de energia elétrica, adotamos a seguinte estrutura de colunas para o treinamento dos modelos:

- Região: Uma coluna que representa a região geográfica na qual o consumo de energia ocorre.
- Data: Uma coluna que indica o mês em que o consumo de energia foi registrado. Cada registro terá um valor correspondente ao mês específico e ao ano específico.
- Tipo de consumo: Uma coluna que descreve o tipo de consumo de energia elétrica. Isso pode incluir as categorias: Residencial, Comercial, Industrial e Outros.
- Consumo de energia: Uma coluna que registra a quantidade de energia elétrica consumida em cada registro.

Essa estrutura permitiu a organização dos dados de consumo de energia elétrica de acordo com a região, o mês e o tipo de consumo. Com esses dados, foi possível realizar análises e obter informações para prever o consumo de energia elétrica em todo o país.

4.3.1 Pré-processamento dos dados

Os dados obtidos a partir da base de dados da EPE têm uma finalidade essencial: facilitar a análise e previsão do consumo de energia elétrica no Brasil. Entretanto, é fundamental reconhecer que esses dados podem conter desafios comuns que afetam a qualidade das análises e modelagens.

Um desses desafios é a presença de ruído nos dados, representando valores incorretos ou irrelevantes, e inconsistências, que se referem a discrepâncias ou contradições nos registros. O processo de limpeza dos dados é fundamental para identificar e solucionar esses problemas, visando evitar que prejudiquem negativamente os resultados das análises (Batista, 2003).

Outra questão relevante é a existência de valores ausentes nos registros, especialmente em campos como o consumo de energia. O tratamento adequado

desses valores faltantes se torna imperativo para assegurar que não impacte negativamente a análise e a previsão.

Adicionalmente, a base de dados pode apresentar classes desbalanceadas, ou seja, uma distribuição desproporcional entre as diferentes categorias de consumo de energia. O desbalanceamento das classes pode gerar resultados enviesados e pouco representativos. Assim sendo, técnicas de balanceamento de classes devem ser empregadas, garantindo uma análise mais precisa e abrangente.

Para assegurar a confiabilidade das análises e previsões relacionadas ao consumo de energia elétrica, é primordial conduzir um pré-processamento minucioso dos dados. Isso envolve não apenas a limpeza para eliminação de ruídos e inconsistências, como também o tratamento de valores ausentes e a equalização das classes (Batista, 2003). A estruturação dos dados em um formato organizado, tal como detalhado anteriormente, desempenha um papel crucial ao facilitar a aplicação de técnicas analíticas e algorítmicas.

A etapa de limpeza assume um papel fundamental ao detectar e tratar questões comuns, tais como valores ausentes, duplicados, *outliers* (valores discrepantes) e erros. Valores faltantes são preenchidos com base em médias ou estimativas derivadas de outros dados. Outliers são tratados por meio de exclusão, garantindo que não prejudiquem as análises. Posteriormente, ocorre a transformação dos dados, voltada para a conversão dos dados brutos em um formato mais adequado para análise. Isso engloba a conversão de tipos de dados, e a padronização das colunas que contêm informações de datas.

Figura 6 - Consumo de energia Elétrica no Brasil (2004-2023)



Fonte: elaborada pela própria autora

Com o término do processo de pré-processamento dos dados, eles se transformam em uma fonte confiável e eficaz para análises futuras. Com os dados devidamente preparados, foi possível criar um gráfico preciso sobre o consumo real de energia elétrica no Brasil, conforme exemplificado na Figura 6.

4.3.2 Porcionamento dos dados

A separação dos conjuntos de treinamento e teste é um passo crucial na análise de séries temporais, permitindo avaliar a capacidade de generalização do modelo de previsão. Essa separação é realizada por meio de dois passos essenciais: determinar a quantidade de dados a serem utilizados no treinamento do modelo e no teste. Optamos por utilizar 70% da massa de dados para o treinamento e os 30% restantes para o teste. A porcentagem utilizada para a divisão foi ajustada conforme a natureza do problema e os dados disponíveis.

Em seguida, dividimos a série temporal em dois grupos: um conjunto de treinamento, com os primeiros dados até um tamanho predefinido, e um conjunto de teste, com os dados restantes, conforme a Tabela 4.1. Esse procedimento assegura que o modelo seja treinado com informações históricas conhecidas e, posteriormente, testado em dados mais recentes, permitindo avaliar sua capacidade de realizar previsões precisas.

Tabela 2 - Distribuição dos Registros nos Conjuntos de Treinamento e Teste

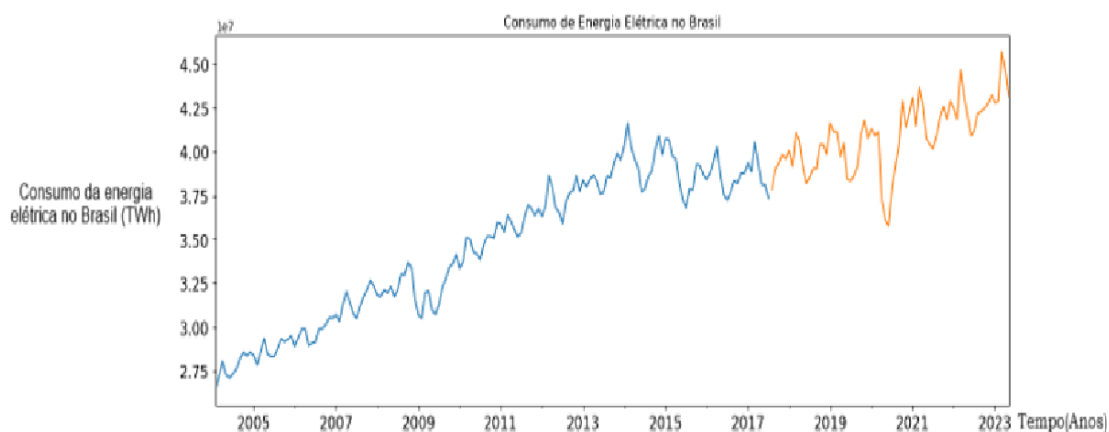
| Conjunto | Quantidade de Registros |
|-------------|-------------------------|
| Treinamento | 4200 |
| Teste | 1800 |

Fonte: elaborada pela própria autora

A Figura 7 exibe visualmente essa distribuição de dados. No gráfico, a cor azul representa o conjunto de treinamento, composto pelos 4200 registros, enquanto a cor laranja simboliza o conjunto de teste, com os 1800 registros

correspondentes. Essa visualização permite uma compreensão clara da divisão dos dados e como eles foram distribuídos nos conjuntos de treinamento e teste.

Figura 7 - Distribuição dos Dados



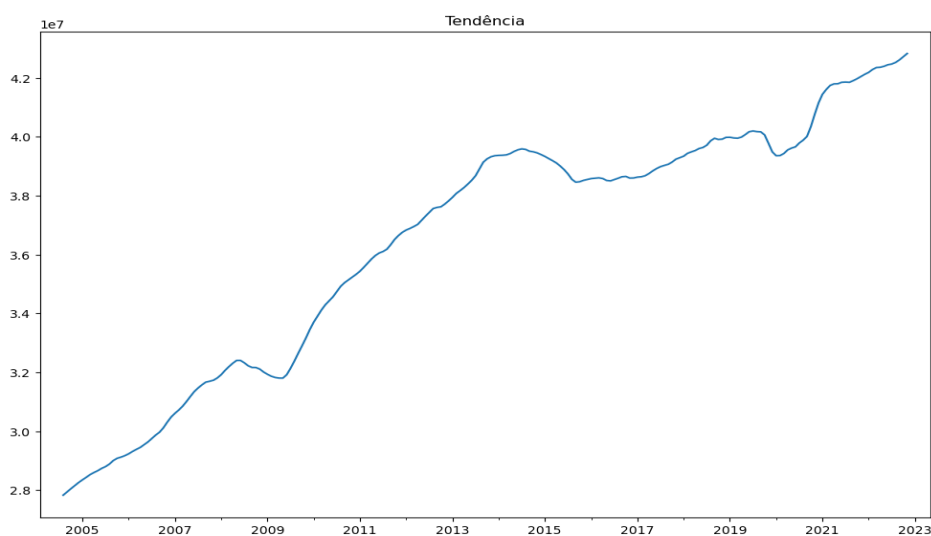
Fonte: elaborada pela própria autora

4.3.3 Decomposição da Série Temporal

Após a etapa de preparação dos dados, avançamos para a decomposição da série, na qual examinamos detalhadamente sua tendência, sazonalidade e ruído. Para atingir esse objetivo, empregamos transformações ou técnicas de diferenciação. A decomposição da série é executada por meio da função *seasonal_decompose* da biblioteca *statsmodels*. Nossa escolha recaiu sobre o modelo multiplicativo, que presume que os componentes da série podem ser multiplicados para reconstruir a série original. Uma vez que aplicamos essa função aos dados, obtemos os componentes fundamentais da série temporal.

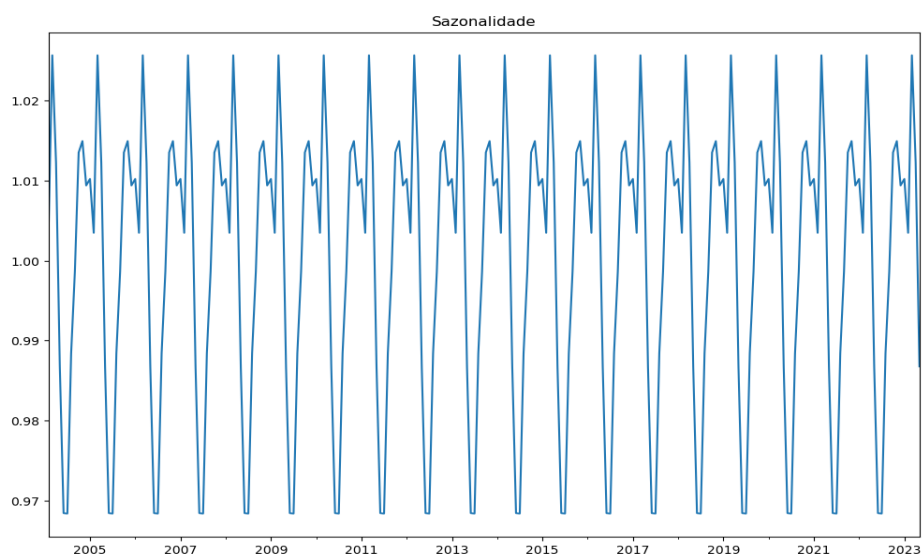
Para plotar a tendência, os valores da série de tendência são definidos em função do tempo, criando uma linha contínua que descreve a direção geral dos dados. A Figura 8 representa a componente de tendência obtida após a decomposição da série temporal e, aparentemente, possui uma tendência crescente. Em outras palavras, os dados estão mostrando uma direção geral de crescimento ao longo do período analisado. Isso pode indicar um aumento contínuo em algum fenômeno ou variável ao longo do tempo.

Figura 8 - Tendência Decomposta



Fonte: elaborada pela própria autora

Figura 9 - Sazonalidade Decomposta



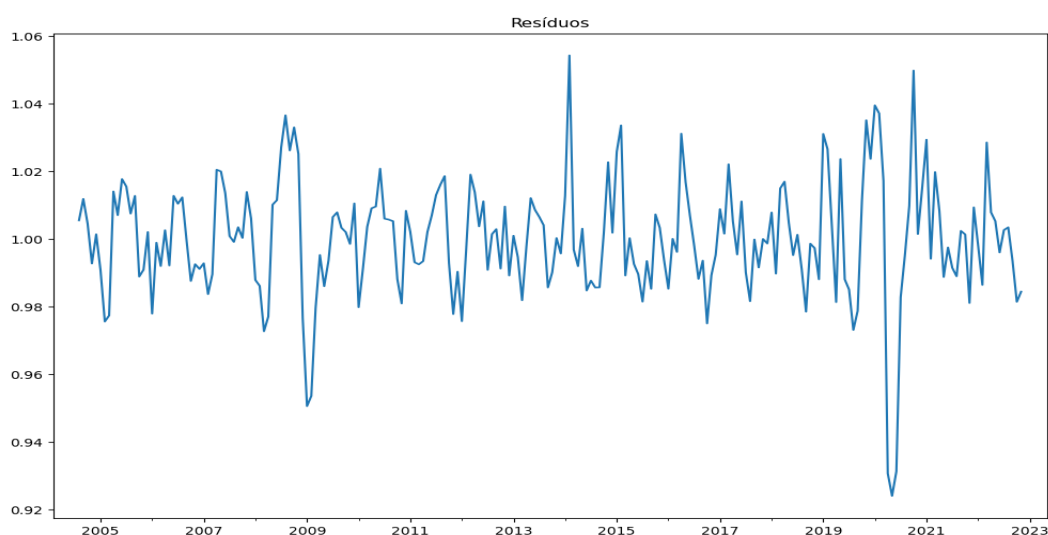
Fonte: elaborada pela própria autora

Ao identificar esse padrão de crescimento na tendência, é importante considerar seu significado e impacto na análise dos dados. Esse comportamento pode ter implicações importantes na formulação de estratégias ou na realização de previsões futuras, pois indica uma tendência de aumento em direção a valores mais altos (Brown, 1959). No gráfico de sazonalidade presente na Figura 9, podemos identificar ciclos regulares que se repetem ao longo do período

analisado. A amplitude desses ciclos indica o quão significativa é a sazonalidade (Brockwell; Davis, 2002).

Os resíduos em uma série temporal representam o que resta dos dados após a eliminação da tendência e da sazonalidade, essencialmente capturando as flutuações aleatórias e imprevisíveis. A análise desses resíduos é crucial, pois eles devem assemelhar-se a um "ruído branco", ou seja, não devem mostrar nenhum padrão aparente ou tendência discernível. Se padrões ou comportamentos não aleatórios forem identificados nos resíduos, é necessário refinamento do modelo ou consideração de outros fatores na análise. Analisando a Figura 10, é possível perceber que essa condição é satisfeita, pois não é possível perceber sazonalidade ou tendência.

Figura 10 - Resíduo Decomposto



Fonte: elaborada pela própria autora

A análise dos componentes da série temporal, após a decomposição, nos proporcionou a identificação da tendência de crescimento e a sazonalidade, o que nos permitiu entender melhor o comportamento dos dados ao longo do tempo. A natureza do ruído branco nos resíduos sugere que a série foi bem-modelada, o que é essencial para a aplicação de modelos estatísticos eficazes.

4.4 DETERMINAÇÃO DOS PARÂMETROS

4.4.1 Suavização Exponencial

Com base na análise detalhada dos padrões extraídos da série temporal, optamos pelo método de Suavização Exponencial de Holt-Winters, uma abordagem amplamente reconhecida por sua capacidade de lidar com dados que exibem complexidades em suas tendências e sazonalidades. A escolha do método Holt-Winters foi orientada pela natureza dos padrões identificados na decomposição da série temporal (Holt, 2004; Winters, 1960).

A adequação desse método reside na capacidade de incorporar os componentes de tendência e sazonalidade. Dessa forma, foi possível capturar de forma precisa as variações ao longo do tempo e realizar previsões mais precisas (Winters, 1960). A seleção do método de previsão, portanto, é um reflexo direto da complexidade dos padrões presentes nos dados, proporcionando uma estrutura sólida para a geração de previsões robustas e informadas sobre o consumo futuro de energia elétrica. A inicialização dos parâmetros desempenha um papel crítico no processo. Esses parâmetros são essenciais para a maneira como o modelo interpreta e captura os padrões presentes nos dados. A inicialização é o processo de atribuir valores iniciais a esses parâmetros antes de ajustar o modelo aos dados de treinamento.

No método *ExponentialSmoothing* do pacote *statsmodels*, a inicialização dos parâmetros é feita automaticamente durante o ajuste do modelo aos dados de treinamento. Ao fornecer os dados de treinamento, o número de períodos sazonais, a abordagem para modelar a tendência e a sazonalidade, o método otimiza automaticamente os valores iniciais para o nível, tendência e fatores de suavização. O resumo gerado com o método *hwes.summary()* exibe os valores otimizados para esses parâmetros, fornecendo uma visão clara de como o modelo foi ajustado aos dados de treinamento, os valores obtidos estão retratados na Tabela 3.

Tabela 3 - Parâmetros Otimizados na Metodologia de Suavização Exponencial

| Parâmetro | Valor | Código |
|----------------------|------------|--------------|
| Smoothing Level | 0,7828571 | <i>Alpha</i> |
| Smoothing Trend | 0,0001 | <i>Beta</i> |
| Smoothing Seasonal | 0,1837363 | <i>Gamma</i> |
| Nível Inicial | 2,7763e+07 | <i>L.0</i> |
| Tendência Inicial | 1,0033779 | <i>B.0</i> |
| Sazonalidade Inicial | 0,9895800 | <i>S.0</i> |

Fonte: elaborada pela própria autora

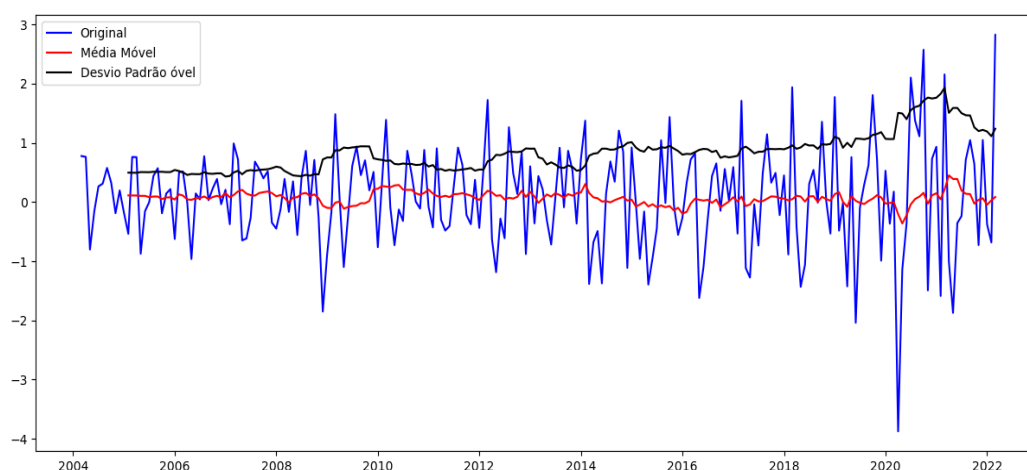
4.4.2 ARIMA

Primeiramente, é definida uma função chamada *test_stationarity(timeseries)*, a qual recebe como entrada a série temporal para analisar a estacionariedade. Dentro da função, são calculadas duas estatísticas importantes: a média móvel e o desvio padrão móvel da série. Isso é feito utilizando a aplicação de uma janela deslizando de tamanho 12 sobre a série, ou seja, a cada 12 pontos consecutivos, a média e o desvio padrão são recalculados, gerando assim curvas suavizadas que mostram as tendências e a variabilidade da série ao longo do tempo.

O próximo passo é a aplicação do teste Dickey-Fuller sobre a série temporal. Esse teste é utilizado para verificar se a série é estacionária ou não. Ele fornece um valor de teste estatístico e um valor p , selecionamos o valor de 0,05 como valor de referência para p . Os resultados do teste Dickey-Fuller indicaram que o valor p de 0,601008, é maior do que 0,05, o que sugere que a série não é estacionária. Portanto, foi necessário realizar ajustes adicionais na série para torná-la estacionária antes de prosseguir com a análise de séries temporais.

Aplicamos uma técnica de diferenciação à série temporal original. Na etapa seguinte, visualizamos, de forma gráfica, as estatísticas de média móvel e desvio padrão móvel, juntamente com a série temporal original. Essa representação visual é de suma importância para reconhecer tendências, padrões e variações que ocorrem ao longo do tempo, abrangendo tanto tendências de longo prazo quanto alterações na amplitude das oscilações. A Figura 11 apresenta a série original (em azul), a média móvel (em vermelho) e o desvio padrão móvel (em preto), criando uma representação visual clara das transformações temporais.

Figura 11 - Média Móvel e Desvio Padrão Móvel



Fonte: elaborada pela própria autora

Tabela 4 - Resultados do Teste Dickey-Fuller e Valores Correspondentes

| Resultados | Valores |
|----------------------|-----------|
| Estatística do Teste | -4,469555 |
| Valor p | 0,000223 |

Fonte: elaborada pela própria autora

Após a diferenciação, uma nova aplicação do teste Dickey-Fuller é conduzida na série temporal transformada. Os resultados do teste Dickey-Fuller subsequente à diferenciação são apresentados na Tabela 4, exibindo

parâmetros essenciais como o valor do teste estatístico e o valor p . É possível identificar a presença de estacionariedade na série temporal após a aplicação da diferenciação.

Para determinar os parâmetros p , d e q , utilizamos a função AutoARIMA para determinar os parâmetros de forma automatizada. Ela realiza iterações ao testar diferentes combinações de valores de p , d e q , selecionando os parâmetros que resultam no melhor ajuste do modelo. Ao usar a função AutoARIMA, o processo de escolha dos parâmetros é simplificado e automatizado. Isso evita a necessidade de uma exploração manual extensa dos gráficos de autocorrelação, tornando o processo mais eficiente e guiado por critérios estatísticos. Dessa forma, determinamos os parâmetros para a técnica do ARIMA.

4.4.3 Máquinas de Vetores de Suporte

O objetivo principal é desenvolver um modelo capaz de fazer previsões precisas com base nos padrões identificados nos dados de treinamento, e posteriormente avaliar seu desempenho nas amostras de teste utilizando. Primeiro, as bibliotecas necessárias são importadas, incluindo a classe MVS do módulo Scikit-Learn para criar o modelo.

Determinaremos o *kernel* a ser utilizado, levando em consideração a modelagem de relações complexas e não lineares entre variáveis. Nessa abordagem, o modelo foi cuidadosamente configurado com parâmetros essenciais: *gamma*, *C* e *epsilon*.

O parâmetro *gamma* desempenha um papel crucial no ajuste do modelo, controlando a influência dos pontos de treinamento. Quanto maior o valor de *gamma*, mais os pontos de treinamento têm um impacto localizado, resultando em um modelo mais sensível às variações nos dados de entrada. Por outro lado, um valor menor de *gamma* permite uma influência mais ampla dos pontos de treinamento, resultando em um ajuste mais suave (Jakkula, 2011).

O parâmetro C é um fator de regularização que equilibra a busca pela precisão do modelo com a sua complexidade. Valores mais altos de C tendem a priorizar um ajuste mais exato aos dados de treinamento, mas podem resultar em um modelo mais complexo e suscetível a *overfitting*. Em contrapartida, valores menores de C tendem a favorecer um modelo mais simples, evitando um ajuste excessivo aos dados de treinamento (Jakkula, 2011).

O *epsilon*, por sua vez, desempenha um papel importante ao determinar a margem de erro aceitável para as previsões do modelo. Ele estabelece um limite para a diferença entre as previsões e os valores reais, permitindo que o modelo seja mais flexível na acomodação das variações nos dados (Jakkula, 2011).

Vale destacar que os valores específicos dos parâmetros *gamma*, C e *epsilon* foram otimizados por meio da técnica de *Grid Search*, a qual explora diversas combinações desses parâmetros para encontrar a configuração mais adequada ao problema em questão. Essa abordagem visa obter um equilíbrio ideal entre precisão e generalização do modelo, resultando em previsões mais confiáveis e robustas para os dados de interesse.

Em seguida, o modelo passou por um processo de treinamento utilizando os dados de treinamento e suas saídas desejadas. Após a conclusão do treinamento, o modelo foi utilizado para fazer previsões tanto nos conjuntos de treinamento quanto nos de teste, gerando matrizes de resultados. Para facilitar a comparação com os valores reais, essas previsões foram revertidas para a escala original por meio de um processo de escalonamento. Com o objetivo de possibilitar uma análise visual, as datas correspondentes aos conjuntos de dados de treinamento e teste foram extraídas. Isso permitiu a criação de gráficos que ilustram as previsões feitas pelo modelo em relação aos valores reais.

4.4.4 Rede Neural Artificial

Para realizar a previsão de séries temporais, utilizamos uma rede neural (RNA) com o framework TensorFlow/Keras. O processo começa com o

carregamento das bibliotecas necessárias para processar e treinar a rede neural. Em seguida, os dados são carregados a partir de um arquivo CSV utilizando o *pandas*. A coluna 'Consumo de energia' foi selecionada, representando o volume de consumo de energia elétrica em intervalos mensais.

Para garantir que a rede neural seja capaz de aprender com os dados, eles são normalizados usando a função *MinMaxScaler*. A normalização é importante para evitar que valores maiores dominem o treinamento, o que poderia prejudicar o desempenho da rede.

O código faz uso da classe *TimeseriesGenerator* do *Keras* para criar um gerador de séries temporais. Esse gerador divide os dados de treinamento em sequências de tamanho definidos na seção 3.3.2, especificamente, *n_input*, e também os dados de teste. Cada uma dessas sequências é utilizada como entrada para o modelo, e a próxima observação subsequente é usada como a saída correspondente.

Ao dividir os dados em sequências, o gerador estabelece uma relação direta entre as entradas e as saídas subsequentes. Isso capacita a rede neural a entender a dependência entre os valores passados e futuros da série temporal. Esse processo de treinamento sequencial é fundamental para modelar e capturar a dinâmica temporal presente nos dados de série temporal (Li *et al.*, 2020). A abordagem de usar sequências como entrada e saída permite que a rede neural aprenda a reconhecer padrões complexos ao longo do tempo, aprimorando assim suas capacidades de previsão.

A etapa central é a criação do modelo RNA. Um modelo sequencial é criado usando o *Keras*, que é uma pilha linear de camadas. Primeiramente, uma camada densa com 100 neurônios é adicionada para capturar os padrões temporais nos dados. Em seguida, uma camada de saída com um único neurônio é incorporada. O modelo é treinado usando o gerador de séries temporais criado previamente, e o processo de treinamento é realizado por 50 épocas, ajustando os pesos da rede com o objetivo de minimizar o erro quadrático médio entre as previsões e os valores reais. Esse número de épocas foi escolhido com base em uma combinação de experimentação e considerações práticas.

Após o treinamento, o código realiza previsões sequenciais para o conjunto de teste. A cada iteração do *loop*, a rede neural prevê a próxima observação com base nas observações anteriores. As previsões são, então, transformadas de volta para a escala original, a fim de facilitar a interpretação e a comparação com os valores reais.

4.5 MÉTRICAS DE AVALIAÇÃO

De acordo com as considerações de Chicco, Warrens e Jurman (2021), as métricas de avaliação mencionadas são positivamente incorporadas à metodologia da pesquisa para uma análise abrangente e precisa do desempenho dos modelos.

Mean Absolute Error. É uma métrica que mede o erro absoluto médio entre os valores reais e os valores previstos pelo modelo de regressão. Ela é calculado como a média das diferenças absolutas entre os valores reais e previstos. O MAE é útil para avaliar a magnitude média dos erros do modelo, mas não leva em consideração a direção dos erros (Chicco; Warrens; Jurman, 2021).

Mean Squared Error. É uma métrica que mede o erro quadrático médio entre os valores reais e os valores previstos pelo modelo de regressão. Ela é calculado como a média das diferenças ao quadrado entre os valores reais e previstos. O MSE é sensível a *outliers*, uma vez que os erros são elevados ao quadrado. Valores mais baixos de MSE indicam um melhor ajuste do modelo aos dados (Chicco; Warrens; Jurman, 2021).

Root Mean-Squared Error. É a raiz quadrada do MSE e representa a média da raiz quadrada das diferenças ao quadrado entre os valores reais e previstos. O RMSE é uma medida de erro semelhante ao MSE, mas está na mesma escala das unidades originais dos dados, sendo interpretável mais facilmente. Assim como o MSE, valores mais baixos de RMSE indicam um melhor ajuste do modelo (Chicco; Warrens; Jurman, 2021).

5 RESULTADOS E DISCUSSÕES

A técnica proposta para prever o consumo de energia elétrica no Brasil é baseada em uma abordagem abrangente e comparativa, envolvendo a aplicação de quatro técnicas de previsão amplamente reconhecidas: Suavização Exponencial, ARIMA, MVS e RNA. Essa metodologia foi desenvolvida com o objetivo de proporcionar uma análise robusta e precisa das tendências de consumo de energia elétrica, permitindo uma compreensão mais profunda dos padrões temporais e sazonais presentes nos dados.

A decomposição da série temporal é realizada para analisar a tendência, a sazonalidade e o ruído presentes nos dados. Esse processo é realizado utilizando o método de Suavização Exponencial de Holt-Winters, que é especialmente adequado para lidar com dados que exibem complexidades em suas tendências e sazonalidades. O método ARIMA é aplicado para modelar os componentes autoregressivos e de média móvel da série temporal, a fim de identificar padrões temporais e sazonais.

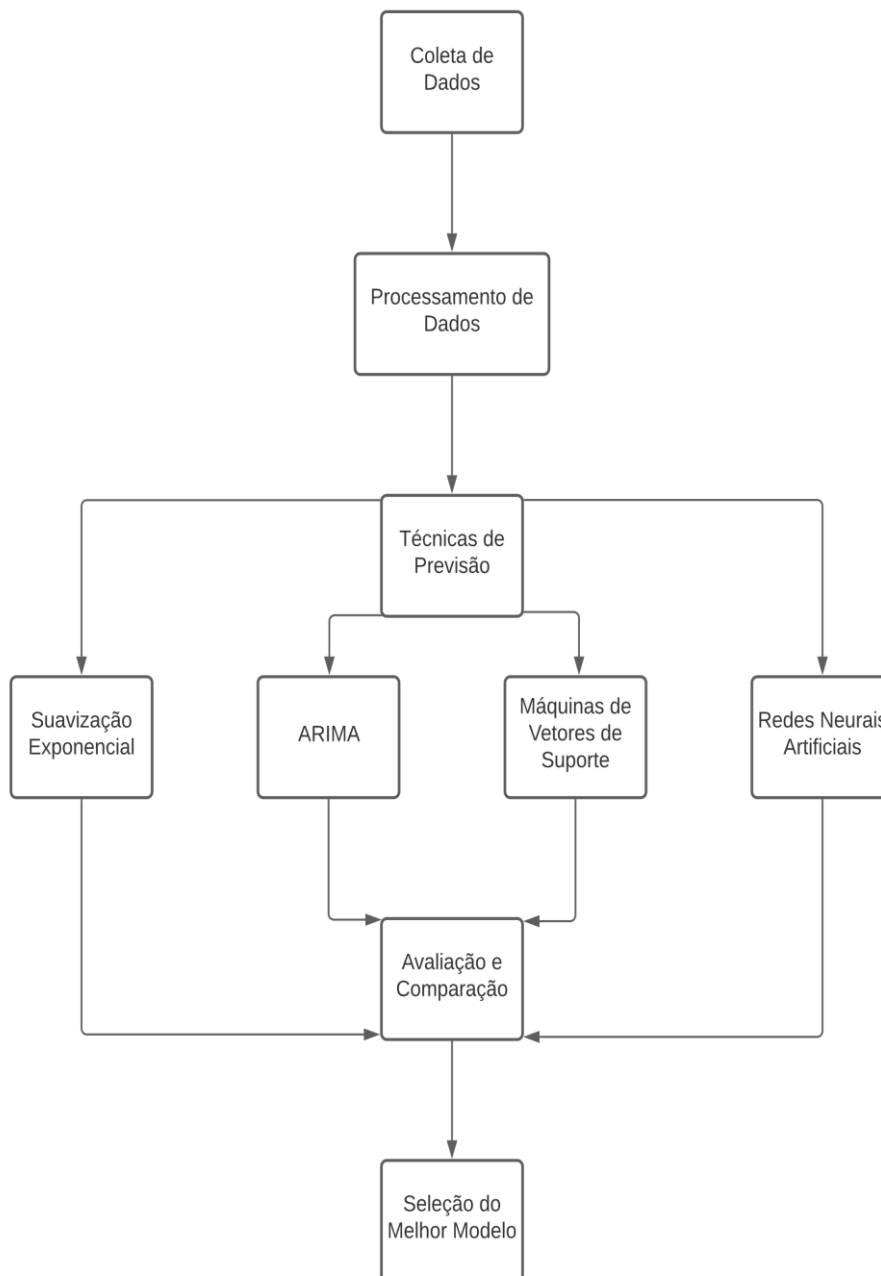
São determinados parâmetros críticos, como *gamma*, *C* e *epsilon*, para criar um modelo que possa fazer previsões precisas e equilibradas. Essa técnica é particularmente útil para capturar relações complexas e não-lineares entre as variáveis de consumo de energia elétrica. A técnica de RNA utiliza uma rede neural para modelar os padrões temporais nos dados de consumo de energia elétrica. Essa arquitetura permite que a rede aprenda a dependência entre os valores passados e futuros da série temporal, melhorando sua capacidade de fazer previsões precisas.

A avaliação das técnicas é conduzida por meio de métricas de desempenho, que permitem comparar o desempenho dos modelos em relação aos valores reais de consumo de energia elétrica foi feita utilizando somente o conjunto de teste.

A proposta de Previsão de Consumo de energia Elétrica no Brasil é delineada por meio de um conjunto de etapas profundamente interconectadas, todas direcionadas ao objetivo central de avaliar a precisão de diferentes técnicas de previsão. O diagrama esquemático que ilustra a sequência de

execução dessas etapas pode ser visualizado na Figura 12. Cada etapa do fluxograma desempenha um papel fundamental na construção e refinamento da análise dos modelos.

Figura 12 - Fluxograma da Proposta de Previsão de Consumo de energia Elétrica no Brasil



Fonte: elaborada pela própria autora

5.1 SUAVIZAÇÃO EXPONENCIAL

Nesta seção, exploramos os resultados da previsão do consumo de energia elétrica no Brasil utilizando a técnica de Suavização Exponencial de Holt-Winters. Embora a Suavização Exponencial seja uma abordagem amplamente utilizada para modelar séries temporais e tenha mostrado eficácia em capturar padrões sazonais e de tendência, nossas descobertas destacam algumas limitações importantes associadas a essa técnica específica.

Ao aplicar a Suavização Exponencial para a previsão do consumo de energia elétrica, obtivemos resultados promissores, porém, uma análise mais profunda revela que a técnica pode não ser tão confiável quanto inicialmente esperávamos. A Tabela 5 apresenta os resultados detalhados da previsão utilizando Suavização Exponencial:

Tabela 5 - Avaliação de Desempenho da Previsão utilizando Suavização Exponencial

| Métrica | Valor |
|---------------------------------------|-----------|
| <i>Mean Absolute Error (MAE)</i> | 0,656307 |
| <i>Mean Squared Error (MSE)</i> | 0,0391339 |
| <i>Root Mean-Squared Error (RMSE)</i> | 0,1979314 |

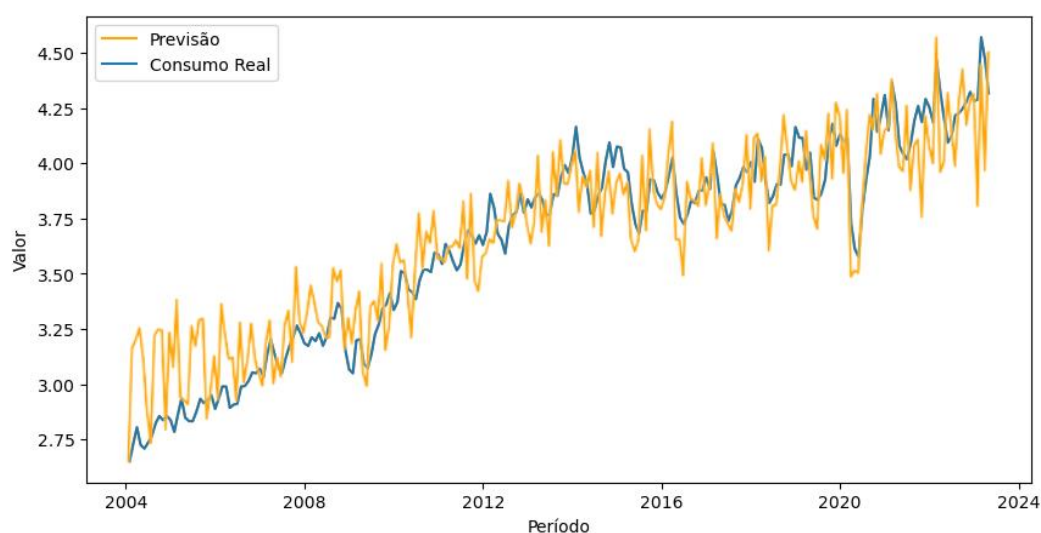
Fonte: elaborada pela própria autora

Os resultados obtidos a partir dos valores do MAE revelam que as previsões derivadas da técnica de Suavização Exponencial apresentam um desvio médio de 0,1656 unidades em relação aos valores reais do consumo de energia elétrica. Esse resultado sugere que as estimativas geradas por essa abordagem podem estar consideravelmente afastadas da realidade.

Avaliando o desempenho por meio da MSE e da RMSE é ainda mais evidente a limitação da técnica de Suavização Exponencial. Os valores elevados dessas métricas destacam que as previsões frequentemente apresentam desvios significativos em relação aos valores reais, resultando em um nível considerável de erro quadrático médio.

A Figura 13 complementa essa análise ao apresentar o gráfico comparativo das previsões, em laranja, e dos valores reais, em azul. Nota-se que as discrepâncias na previsão são mais notáveis, especialmente no início do período analisado. Essa observação reforça as limitações apontadas nas métricas de erro mencionadas anteriormente e sublinha a necessidade de aprimoramentos na técnica de Suavização Exponencial para melhor capturar os padrões complexos presentes nos dados.

Figura 13 - Previsão versus Consumo Efetivo de energia Elétrica: Suavização Exponencial



Fonte: elaborada pela própria autora

A análise do gráfico e das métricas destaca a necessidade de considerar cuidadosamente as limitações da Suavização Exponencial ao utilizar o método para prever o consumo de energia elétrica. Embora a técnica tenha mostrado promessas iniciais, sua falta de confiabilidade em produzir previsões precisas e consistentes levanta preocupações sobre sua aplicabilidade prática em cenários de previsão de longo prazo.

5.2 ARIMA

Os resultados obtidos ao aplicar a metodologia ARIMA para prever o consumo de energia elétrica no Brasil são apresentados na Tabela 6. A Tabela

oferece uma visão abrangente do desempenho da abordagem ARIMA na previsão, revelando importantes métricas de avaliação:

Tabela 6 - Avaliação de Desempenho da Metodologia ARIMA

| Métrica | Valor |
|---------------------------------------|-----------|
| <i>Mean Absolute Error</i> (MAE) | 0,0923561 |
| <i>Mean Squared Error</i> (MSE) | 0,0187832 |
| <i>Root Mean-Squared Error</i> (RMSE) | 0,1347583 |

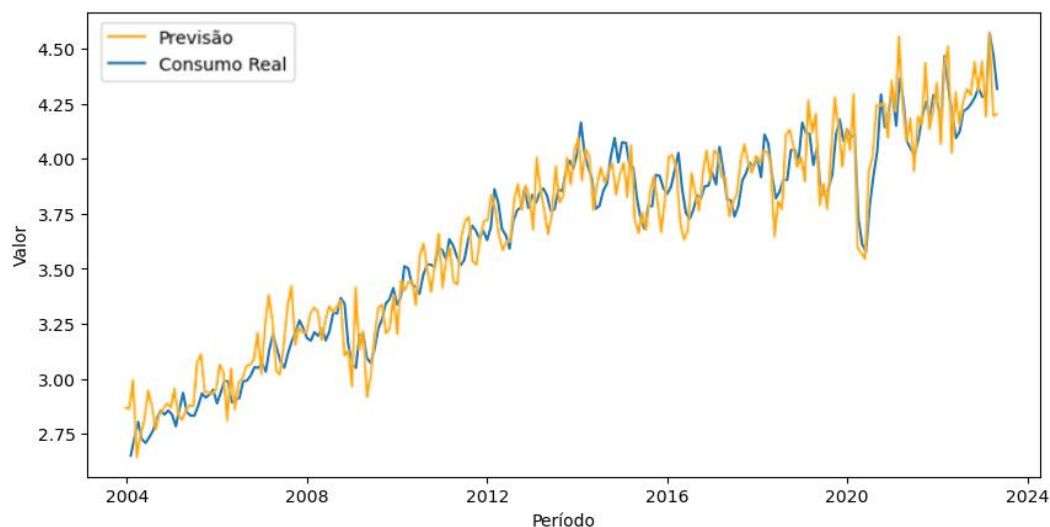
Fonte: elaborada pela própria autora

A metodologia ARIMA combinou componentes autoregressivos, integrados e de média móvel para modelar padrões temporais complexos presentes nos dados de consumo de energia elétrica. A etapa de diferenciação integrada permitiu transformar a série temporal original em uma série estacionária, tornando-a mais adequada para a modelagem ARIMA. A identificação dos parâmetros p , d e q do modelo ARIMA foi realizada de forma automatizada, simplificando o processo de seleção e resultando em uma abordagem mais eficiente em comparação com métodos manuais.

A representação gráfica das estatísticas de média móvel e desvio padrão móvel, juntamente com a série temporal original, ofereceu informações visuais valiosas sobre tendências, padrões e variações ao longo do tempo. A aplicação do teste Dickey-Fuller auxiliou na avaliação da estacionariedade da série, garantindo que o modelo ARIMA fosse aplicado a dados apropriados.

O gráfico apresentado na Figura 14 ilustra a previsão (em laranja) em comparação com o consumo real (em azul). Uma análise visual rápida revela uma notável proximidade entre os valores previstos e reais. Esse alinhamento mais estreito entre as previsões e a realidade desempenha um papel crucial na otimização da alocação de recursos e na tomada de decisões informadas em relação à gestão energética.

Figura 14 - Comparação entre Previsão (laranja) e Consumo Real (azul) de energia Elétrica pelo método ARIMA



Fonte: elaborada pela própria autora

A abordagem ARIMA demonstrou ser robusta e eficaz na previsão do consumo de energia elétrica no Brasil. Suas conquistas superiores em comparação com o modelo de Suavização Exponencial ressaltaram sua notável aplicabilidade e potencial para aprimorar significativamente a precisão das previsões.

5.3 MÁQUINAS DE VETORES DE SUPORTE

Durante a avaliação do modelo de Máquinas de Vetores de Suporte para previsão de consumo de energia elétrica, dedicamos uma fase crucial à otimização dos parâmetros essenciais, visando alcançar um desempenho maximizado. Nesse processo de aprimoramento, seguimos uma abordagem sequencial, priorizando as configurações que resultassem em erros de previsão minimizados, conforme as métricas de avaliação analisadas.

A primeira etapa desse processo envolveu a seleção da função de *kernel* mais apropriada. Diversas funções de *kernel* foram testadas, e a função *RBF* se destacou ao proporcionar os melhores resultados. A função *RBF* demonstrou habilidade para acompanhar de maneira eficaz a variação do consumo de energia elétrica observada durante o período da pandemia. Os resultados promissores obtidos por meio da aplicação da técnica de Máquinas de Vetores

de Suporte para a previsão de consumo de energia elétrica são apresentados na Tabela 7.

Tabela 7– Métricas de Avaliação MVS

| Métrica | Valor |
|--|-----------|
| <i>Mean Absolute Error</i> (MAE) | 0,0367781 |
| <i>Mean Squared Error</i> (MSE) | 0,0046713 |
| <i>Root Mean-Squared Error</i> (RMSE) | 0,0465637 |

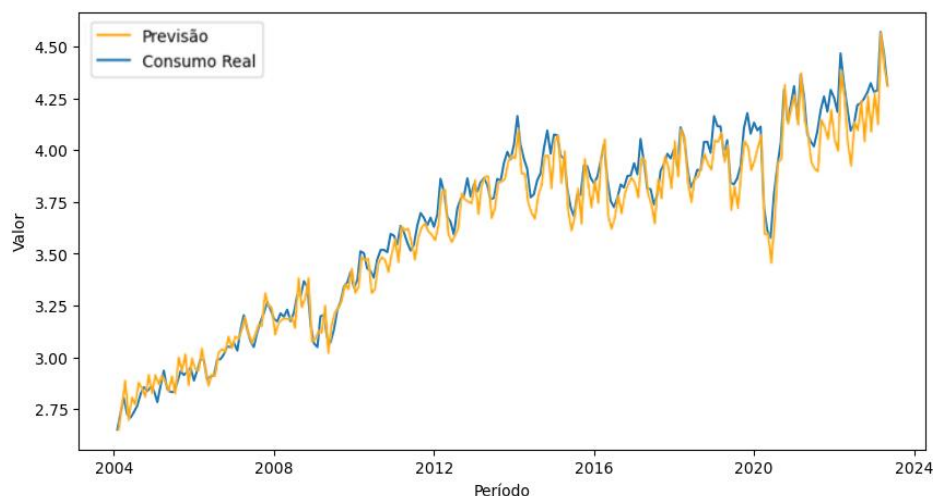
Fonte: elaborada pela própria autora

A interpretação dessas métricas ressalta o desempenho promissor do modelo MVS na previsão de consumo de energia elétrica. Valores baixos em todas as métricas indicam uma precisão substancialmente maior das previsões em relação aos valores reais.

Essa análise revela um progresso notável em comparação com as técnicas de Suavização Exponencial e ARIMA, anteriormente exploradas. Ao demonstrar um desempenho superior, o MVS evidencia seu potencial para fazer contribuições substanciais no âmbito das tomadas de decisões estratégicas e do planejamento otimizado do fornecimento de energia elétrica.

Na Figura 15, podemos ver o gráfico representativo da previsão e do consumo real de energia elétrica. Esse visual ilustrativo reforça de maneira concreta a precisão da previsão do modelo MVS, enfatizando como suas estimativas estão alinhadas de forma notável com os valores reais de consumo. Esse aspecto visual não apenas corrobora a validade do modelo, mas também torna tangível seu papel na otimização das operações e na gestão eficaz dos recursos energéticos.

Figura 15 - Previsão utilizando o método MVS



Fonte: elaborada pela própria autora

Comprovando sua capacidade, o MVS não apenas oferece uma previsão confiável, mas também abre portas para aprimorar significativamente a forma como abordamos a gestão e a alocação dos recursos energéticos. A habilidade do MVS em capturar padrões complexos e sutilezas nos dados de consumo elétrico resulta em uma ferramenta confiável para informar as estratégias de suprimento energético.

5.4 REDE NEURAL ARTIFICIAL

Em nossa busca contínua por aprimorar a precisão das previsões de consumo de energia elétrica no contexto específico do Brasil, exploramos o potencial das Redes Neurais. Inicialmente, direcionamos nossos esforços para a implementação de um MLP, uma abordagem padrão em aprendizado de máquina. Entretanto, ao aplicarmos o MLP à nossa série temporal de consumo de energia elétrica no Brasil, os resultados alcançados não atenderam às nossas expectativas de precisão.

Diante desse desafio, redirecionamos nossa abordagem e voltamos nossa atenção para as Redes Neurais de Longa Memória de Curto Prazo, uma variação avançada das redes neurais recorrentes. A escolha da RNA se mostrou estratégica, uma vez que essa arquitetura é especialmente adequada para

capturar dependências temporais de longo alcance (Li *et al.*, 2020), o que é crucial para a previsão precisa de séries temporais complexas, como o consumo de energia elétrica.

Ao considerar a natureza peculiar da série temporal de consumo de energia elétrica no Brasil, caracterizada por variações sazonais, tendências de longo prazo e fatores imprevisíveis, reconhecemos que uma abordagem mais sofisticada era necessária. A escolha da RNA não apenas respeitou essa complexidade, mas também se alinhou com nossa meta de obter resultados mais promissores. Avaliar a performance de modelos de previsão é fundamental para entender a confiabilidade de suas estimativas. A Tabela 8 apresenta os resultados detalhados da previsão utilizando a RNA.

Tabela 8 - Métricas de Desempenho da Previsão com RNA

| Métrica | Valor |
|---------------------------------------|-----------|
| <i>Mean Absolute Error (MAE)</i> | 0,0230906 |
| <i>Mean Squared Error (MSE)</i> | 0,0038744 |
| <i>Root Mean-Squared Error (RMSE)</i> | 0,0263905 |

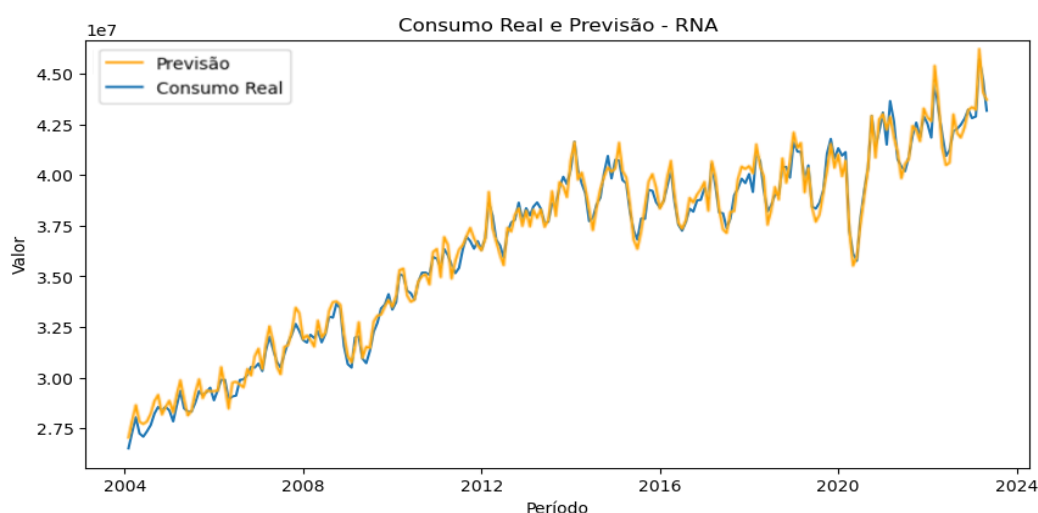
Fonte: elaborada pela própria autora

Os resultados claramente evidenciam a notável superioridade da abordagem RNA em relação a outras metodologias, tanto em termos de precisão quanto de confiabilidade, no contexto da previsão de consumo de energia elétrica. O valor do MAE é de 0,0230906, indicando que, em média, as previsões geradas pela RNA desviam cerca de 0,0230906 unidades dos valores reais. Esse desvio médio atesta a impressionante capacidade da RNA em capturar os padrões subjacentes presentes na série temporal. Ao considerarmos o MSE de 0,0038744 e a RMSE de 0,0263905, é possível observar um nível consistentemente baixo de erro quadrático médio entre as previsões e os valores reais.

A proeminente superioridade da abordagem RNA deriva de sua habilidade ímpar em capturar relações complexas e não lineares presentes nos dados, bem como de sua competência em aprender sequências de padrões de longo alcance. Cabe ressaltar, entretanto, que a RNA possui uma complexidade computacional mais elevada em comparação com abordagens mais simples (Li *et al.*, 2020).

Com uma abordagem visual, a Figura 16 traz uma representação gráfica que ressalta as previsões em laranja e os valores reais em azul, permitindo uma compreensão imediata das semelhanças entre essas duas séries. Essa visualização oferece uma perspectiva tangível das tendências, flutuações e desvios presentes ao longo do período analisado. A clareza proporcionada pela Figura 16 contribui para uma análise mais profunda do desempenho da abordagem RNA na previsão do consumo de energia elétrica, permitindo identificar os momentos de maior alinhamento ou divergência entre as previsões e os dados reais. Os resultados enfatizam a eficácia da abordagem RNA na previsão do consumo de energia elétrica, destacando sua habilidade singular em apreender padrões complexos e a sequencialidade intrínseca dos dados.

Figura 16 - Comparação entre Previsões e Valores Reais de Consumo de energia Elétrica



Fonte: elaborada pela própria autora

5.5 COMPARAÇÃO ENTRE OS MÉTODOS

A análise comparativa dos métodos de previsão de consumo de energia elétrica no Brasil - Suavização Exponencial, ARIMA, MVS e RNA - revela informações valiosas sobre suas capacidades e limitações.

A técnica de Suavização Exponencial é amplamente utilizada para modelar séries temporais e mostra eficácia na captura de padrões sazonais e de tendência. No entanto, suas limitações se tornam evidentes ao apresentar desvios significativos em relação aos valores reais, resultando em menor precisão em comparação com outros métodos. Os desvios em relação aos valores reais, como evidenciado pelo *Mean Absolute Error* de 0,1656, levantam preocupações sobre sua confiabilidade.

O modelo ARIMA, por sua vez, exibe uma boa precisão na previsão e é capaz de lidar com componentes autoregressivos, integrados e de média móvel, o que o torna eficaz na modelagem de padrões temporais lineares. No entanto, suas limitações residem na possível falta de eficácia na captura de padrões não lineares e complexos. As métricas, como o MSE de 0,0187, indicam uma melhoria em relação à Suavização Exponencial, mas há espaço para aprimoramento.

Os resultados evidenciaram um desempenho superior do modelo ARIMA em comparação com o modelo de Suavização Exponencial, anteriormente utilizado para prever o consumo de energia elétrica. Essas métricas de avaliação destacaram a maior precisão do ARIMA na previsão, crucial para tomar decisões informadas e otimizar a gestão de recursos.

As Máquinas de Vetores de Suporte se destacam ao proporcionar uma precisão substancialmente maior das previsões em relação aos valores reais. Essa abordagem é capaz de capturar relações complexas e minimizar erros. No entanto, a otimização de parâmetros é necessária, e foi exigente em termos computacionais. A abordagem MVS se destaca com métricas promissoras, como MAE de 0,0367, sinalizando uma precisão substancialmente maior em relação aos valores reais. No entanto, requer otimização de parâmetros e pode ser computacionalmente intensivo.

Essa análise revela um progresso notável em comparação com as técnicas de Suavização Exponencial e ARIMA, anteriormente exploradas. Ao

demonstrar um desempenho superior, o MVS evidencia o seu potencial para fazer contribuições substanciais no âmbito das tomadas de decisões estratégicas e do planejamento otimizado do fornecimento de energia elétrica.

A RNA emerge como a vencedora clara. Com um MAE de apenas 0,0231, demonstra notável superioridade em precisão e confiabilidade. A RNA capta relações complexas e sequenciais, sendo especialmente adequada para séries temporais complexas como o consumo de energia elétrica. Apesar da complexidade computacional mais alta, sua habilidade em modelar padrões de longo alcance a torna a opção preferida para decisões informadas e otimização da gestão de recursos.

A RNA se destaca como a melhor escolha para a previsão de consumo de energia elétrica no Brasil. Sua capacidade de previsão precisa e confiável, juntamente com a habilidade de modelar padrões complexos, a coloca à frente das outras abordagens consideradas, impulsionando a tomada de decisões informadas e otimização dos recursos energéticos.

No entanto, é importante destacar que a escolha do modelo final deve considerar não apenas as métricas de avaliação, mas também a adequação do modelo às características dos dados, a simplicidade do modelo, a interpretabilidade das previsões e a eficiência computacional. A avaliação completa deve considerar um equilíbrio entre esses fatores para selecionar a abordagem mais adequada para a previsão do consumo de energia elétrica no Brasil.

6 CONCLUSÃO

A análise realizada demonstra a importância e a viabilidade das abordagens avançadas de previsão, utilizando tanto métodos tradicionais quanto técnicas de Inteligência Artificial (IA) e Aprendizado de Máquina (ML), no contexto da previsão de consumo de energia elétrica no Brasil. O setor energético desempenha um papel fundamental na busca por soluções sustentáveis diante das preocupações globais com as mudanças climáticas, e a precisão das previsões é essencial para garantir um fornecimento eficiente e sustentável.

A comparação detalhada entre abordagens estatísticas e modelos de ML, forneceu informações valiosas sobre suas respectivas capacidades e limitações. A análise demonstra que a escolha da técnica de previsão deve considerar uma série de fatores, incluindo desempenho, adaptabilidade, complexidade computacional e a adequação aos dados específicos.

No contexto específico da previsão de consumo de energia elétrica no Brasil, a abordagem vencedora é claramente a RNA. Sua capacidade de modelar padrões complexos, sazonalidades e relações não lineares nos dados de consumo de energia elétrica a torna uma escolha superior. Embora possa requerer mais recursos computacionais, seus benefícios em termos de precisão e capacidade de previsão superam amplamente essa consideração.

No entanto, é fundamental reconhecer que cada técnica abordada tem seu próprio conjunto de vantagens e desafios, e a seleção da abordagem adequada dependerá das necessidades específicas da empresa do setor energético. As conclusões deste estudo fornecem uma base sólida para a tomada de decisões informadas, otimização de recursos e gestão eficaz do consumo de energia elétrica.

Em última análise, este trabalho contribui significativamente para a compreensão das melhores práticas na previsão de consumo de energia elétrica. À medida que o cenário global continua a evoluir, a aplicação de abordagens avançadas de previsão, como a Rede Neural Artificial RNA, permitirá uma

adaptação mais ágil e uma resposta mais eficaz aos desafios futuros, promovendo um futuro energético mais sustentável e confiável para o Brasil.

REFERÊNCIAS

- ABRAHÃO, Cristina de Freitas Jorge. **A influência da estrutura demográfica sobre o consumo residencial de energia elétrica**. 2020. Tese (Doutorado em Arquitetura e Urbanismo) - Escola de Arquitetura, Universidade Federal de Minas Gerais, Belo Horizonte, 10 dez. 2020. *E-book*. Disponível em: <http://hdl.handle.net/1843/37754>. Acesso em: 09 jun. 2023.
- AHMAD, A. S.; HASSAN, M. Y.; ABDULLAH, M. P.; RAHMAN, H. A.; HUSSIN, F.; ABDULLAH, H.; SAIDUR, R. A review on applications of ANN and SVM for building electrical energy consumption forecasting. **Renewable and Sustainable Energy Reviews**, [s. l.], v. 33, p. 102–109, 2014. DOI: <https://doi.org/10.1016/j.rser.2014.01.069>. Disponível em: <https://www.sciencedirect.com/science/article/abs/pii/S1364032114000914>. Acesso em: 14 mai.2023.
- ALTOÉ, Leandra; COSTA, José Márcio; OLIVEIRA FILHO, Delly; REY MARTINEZ, Francisco Javier; FERRAREZ, Adriano Henrique; VIANA, Lucas de Arruda. Políticas públicas de incentivo à eficiência energética. **Estudos Avançados**, [São Paulo] v. 31, p. 285-297, 2017. DOI: <https://doi.org/10.1590/s0103-40142017.31890022>. Disponível em: <https://www.scielo.br/j/ea/a/vPxbFKL9Jvwg559c6cgCZWp/?lang=pt>. Acesso em: 18 jun. 2023.
- ALVES, Marleide F.; LOTUFO, Anna Diva P.; LOPES, Mara Lúcia M. Seleção de variáveis stepwise aplicadas em redes neurais artificiais para previsão de demanda de cargas elétricas. **Proceeding Series of the Brazilian Society of Computational and Applied Mathematics**, [s. l.] v. 1, n. 1, 2013. DOI: <https://doi.org/10.5540/03.2013.001.01.0144>. Disponível em: <https://proceedings.sbmacc.org.br/sbmacc/article/view/144/144>. Acesso em: 9 jun. 2023.
- BARAK, S.; SADEGH, S. S. Forecasting energy consumption using ensemble ARIMA–ANFIS hybrid algorithm. **International Journal of Electrical Power & Energy Systems**, [s. l.], v. 82, p. 92–104, 2016. DOI: <https://doi.org/10.1016/j.ijepes.2016.03.012>. Disponível em: <https://www.sciencedirect.com/science/article/abs/pii/S0142061516303702>. Acesso em: 13 mai.2023.
- BATISTA, Gustavo Enrique de Almeida Prado Alves. **Pré-processamento de dados em aprendizado de máquina supervisionado**. 2003. Tese (Doutorado em Ciências de Computação e Matemática Computacional) - Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo, São Carlos, 2003. DOI10.11606/T.55.2003.tde-06102003-160219. Disponível em: <https://www.teses.usp.br/teses/disponiveis/55/55134/tde-06102003-160219/publico/TeseDoutorado.pdf>. Acesso em: 4 jul. 2023.
- BERWICK, R. **An Idiot's guide to Support vector machines (SVMs)**. [S. l.: s. n.], [200-?]. Disponível em: <https://web.mit.edu/6.034/wwwbob/svm-notes-long-08.pdf>. Acesso em: 20 jun. 2023.
- BOX, G. E. P.; JENKINS, G. M.; REINSEL, G. C.; LJUNG, G. M. **Time series analysis: forecasting and control**. 5th ed. [S. l.]: John Wiley & Sons, 2015. Disponível em: https://www.researchgate.net/publication/299459188_Time_Series_Analysis_Forecasting_and_Control5th_Edition_by_George_E_P_Box_Gwilym_M_Jenkins_Gregory_C_Reinsel_and_Greta_M_Ljung_2015_Published_by_John_Wiley_and_Sons_Inc_Hoboken_New_Jersey_pp_712_ISBN_. Acesso em: 10 jun. 2023.
- BRASIL, Millena. Maioria dos brasileiros planeja diminuir consumo de energia, diz pesquisa. **Agência Eixos**, [Rio de Janeiro], 06 fev. 2023. Disponível em: <https://epbr.com.br/maioria-dos-brasileiros-planeja-diminuir-consumo-de-energia-aponta-pesquisa/>. Acesso em: 15 jun. 2023.
- BRASIL. Ministério do Meio Ambiente. **Diretrizes para uma Estratégia Nacional para Neutralidade Climática**. Brasília, DF: Ministério do Meio Ambiente, 2020. *E-book*. Disponível em: <https://www.gov.br/mma/pt-br/assuntos/mudanca-do-clima/clima>. Acesso em: 2 maio. 2023.

BROCKWELL, Peter J.; DAVIS, Richard A. **Introduction to Time Series and Forecasting**, 2nd ed. New York: Springer, 2002. ISBN 0-387-95351-5. Disponível em: <http://home.iitj.ac.in/~parmod/document/introduction%20time%20series.pdf>. Acesso em: 15 jun. 2023.

BROWN, R. G. **Statistical forecasting for inventory control**. [S. l.]: McGraw/Hill, 1959. *E-book*.

CAI, M. Application research of artificial intelligence in electrical automation control. **Procedia Computer Science**, [s. l.], v. 166, p. 292-295, 2020. DOI: <https://doi.org/10.1016/j.procs.2020.02.097>. Disponível em: <https://www.sciencedirect.com/science/article/pii/S1877050920302192>. Acesso em: 16 mai. 2023.

CARVALHO, D. **Métodos de previsão de consumo de energia elétrica residencial em grande volume de dados**. 2019. Trabalho de Conclusão de Curso (Bacharelado em Sistemas de Informação) – Universidade Federal Rural de Pernambuco, Recife, 2019. Disponível em: https://www.repository.ufrpe.br/bitstream/123456789/2135/1/tcc_danieljos%C3%A9decarvalho.pdf. Acesso em: 2 mai. 2023.

CARVALHO, M. M. Q.; LA ROVERE, E. L.; GONÇALVES, A. C. M. Analysis of variables that influence electric energy consumption in commercial buildings in Brazil. **Renewable and Sustainable Energy Reviews**, [s. l.], v. 14, n. 9, p. 3199-3205, 2010. DOI: <https://doi.org/10.1016/j.rser.2010.07.009>. Disponível em: <https://www.sciencedirect.com/science/article/abs/pii/S1364032110001899>. Acesso em: 09 jun. 2023.

CHICCO, D.; WARRENS, M. J.; JURMAN, G. The coefficient of determination R-squared is more informative than SMAPE, MAE, MAPE, MSE and RMSE in regression analysis evaluation. **PeerJ Computer Science**, [s. l.], v. 7, e623, 5 jul. 2021. DOI 10.7717/peerj-cs.623. Disponível em: <https://pubmed.ncbi.nlm.nih.gov/34307865/>. Acesso em: 5 jul. 2023.

CORREIA-SILVA, David Costa; RODRIGUES, Marcos. Análise da eficiência no consumo de energia dos estados brasileiros. **Planejamento e Políticas Públicas**, Brasília, v. 46, 2015. Disponível em: <https://repositorio.ipea.gov.br/handle/11058/6547>. Acesso em: 12 jun. 2023.

DIAS, Danilo de Souza; MOREIRA, Hermom Leal; OLIVEIRA, Pedro André Braga de; SILVA, José Ricardo da. Eficiência energética na Indústria 4.0. **Revista Brasileira de Mecatrônica**, [S. l.], v. 5, n. 3, p. 44–61, 2023. Disponível em: <https://revistabrmecatronics.sp.senai.br/ojs/index.php/revistabrmecatronics/article/view/193>. Acesso em: 10 jun. 2023.

DONG, Jian-Xiong; KRZYZAK, A.; SUEN, C. Y. Fast SVM training algorithm with decomposition on very large data sets. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, [s. l.], v. 27, n. 4, p. 603–618, abr. 2005. DOI: 10.1109/TPAMI.2005.77. Disponível em: <https://ieeexplore.ieee.org/document/1401912>. Acesso em: 23 mai. 2023.

DU, Juan; MIN, Qilong; ZHANG, Penglin; GUO, Jinhui; YANG, Jun; YIN, Bangsheng. Short-Term Solar Irradiance Forecasts Using Sky Images and Radiative Transfer Model. **Energies**, [S. l.], v. 11, n. 5, p. 1107, 2018. DOI: <https://doi.org/10.3390/en11051107>. Disponível em: <https://www.mdpi.com/1996-1073/11/5/1107>. Acesso em: 5 jun. 2023.

EMPRESA DE PESQUISA ENERGÉTICA (EPE). **Consumo de energia Elétrica: Quem Somos**. [S. l.]: EPE, [202-?]. Disponível em: <https://www.epe.gov.br/pt/a-epe/quem-somos>. Acesso em: 8 jul. 2023.

EMPRESA DE PESQUISA ENERGÉTICA (EPE). **Consumo Mensal de energia Elétrica por Classe (regiões e subsistemas)**. [S. l.]: EPE, [202-?]. Disponível em: <https://www.epe.gov.br/pt/publicacoes-dados-abertos/publicacoes/consumo-de-energia-eletrica>. Acesso em: 15 jun. 2023.

EMPRESA DE PESQUISA ENERGÉTICA (EPE). **MATRIZ ENERGÉTICA**. [S. l.]: EPE, [202-?]. Disponível em: <https://www.epe.gov.br/pt/abcdenergia/matriz-energetica-e-eletrica>. Acesso em: 16 jun. 2023.

FERREIRA, B.; BORELLI, R. **Eficiência energética**: técnicas de aproveitamento, gestão de recursos e fundamentos. [S. l.]: Saraiva Educação S.A., 2015. *E-book*.

FLECK, Leandro; TAVARES, Maria Hermínia Ferreira; EYNG, Eduardo; HELMANN, Andrieli Cristina; ANDRADE, Minéia Aparecida de Moares. Redes neurais artificiais: Princípios básicos. **Revista Eletrônica Científica Inovação e Tecnologia**, Medianeira, v. 1, n. 13, p. 47-57, 2016. Disponível em: <https://periodicos.utfpr.edu.br/recit/article/viewFile/4330/Leandro>. Acesso em: 30 jun. 2023.

FRANCISCO, Eduardo de Rezende. **Indicadores de renda baseados em consumo de energia elétrica**: abordagens domiciliar e regional na perspectiva da estatística espacial. 2010. Dissertação (Mestrado em Estatística) – Universidade Federal de Minas Gerais, Belo Horizonte, 2010. *E-book*. Disponível em: <https://repositorio.fgv.br/items/c8de5d31-a952-41fd-b1d6-52218cc474a1>. Acesso em: 16 jun. 2023.

GONÇALVES, Luiz Fernando. **Contribuições para o estudo teórico e experimental de sistemas de geração distribuída**. 2004. Dissertação (Mestrado em Engenharia Elétrica) — Universidade Federal do Rio Grande do Sul, Escola de Engenharia, Programa de Pós-Graduação em Engenharia Elétrica, Porto Alegre, 2004. Disponível em: <http://hdl.handle.net/10183/6778>. Acesso em: 9 jun. 2023.

HAMILTON, James D. **Time series analysis**. Princeton: Princeton University Press, 1994. *E-book*.

HANSEN, Alice Maria Dreher. **Padrões de consumo de energia elétrica em diferentes tipologias de edificações residenciais, em Porto Alegre**. 2000. Dissertação (Mestrado em Engenharia Civil) — Universidade Federal do Rio Grande do Sul, Escola de Engenharia, Programa de Pós-Graduação em Engenharia Civil, Porto Alegre, 2000. Disponível em: <http://hdl.handle.net/10183/3546>. Acesso em: 5 mai. 2023.

HAYKIN, S. **Neural networks**: a comprehensive foundation. 2nd ed. New York: Macmillan, 1998. *E-book*.

HAYKIN, S. S. **Neural networks and learning machines**. Hamilton: Pearson, 2009. *E-book*.
HO, Sui-Lau; XIE, Min; GOH, Thong Ngee. A comparative study of neural network and Box-Jenkins ARIMA modeling in time series prediction. **Computers & Industrial Engineering**, [s. l.], v. 42, n. 2-4, p. 371-375, 2002. DOI: [https://doi.org/10.1016/S0360-8352\(02\)00036-0](https://doi.org/10.1016/S0360-8352(02)00036-0). Disponível em: <https://www.sciencedirect.com/science/article/abs/pii/S0360835202000360>. Acesso em: 28 jun. 2023.

HOBBY, J. D.; TUCCI, G. H. Analysis of the residential, commercial and industrial electricity consumption. In: IEEE PES INNOVATIVE SMART GRID TECHNOLOGIES, 2011, Perth. **Anais [...]**. Perth: IEEE, 2011. p. 1-7. DOI 10.1109/ISGT-Asia.2011.6167087. Disponível em: <https://ieeexplore.ieee.org/document/6167087>.

HOLT, C. C. Forecasting seasonals and trends by exponentially weighted averages. **International Journal of Forecasting**, [Pittsburgh], v. 20, p. 5-10, 2004. DOI: <https://doi.org/10.1016/j.ijforecast.2003.09.015>. *E-book*. Disponível em: <https://www.jstor.org/stable/2627346>. Acesso em: 20 jun. 2023.

HYNDMAN, R. J.; ATHANASOPOULOS, G. **Forecasting**: Principles and Practice. Melbourne: OTexts, 2021. Disponível em: <https://otexts.com/fpp3>. Acesso em: 6 mai. 2023.

HYNDMAN, Rob J. Forecasting: An Overview. **International Encyclopedia of Statistical Science**, [Berlin], v. 1, p. 536-539, 2010. DOI 10.1007/978-3-642-04898-2_256. Disponível em: <https://robjhyndman.com/papers/forecastingoverview.pdf>. Acesso em: 26 jun. 2023.

INTERNATIONAL ORGANIZATION FOR STANDARDIZATION. ISO 50001:2018 – **Sistemas de gestão de energia**: Requisitos e orientações para uso. Genebra: ISO, 2018. Disponível em: <https://www.iso.org/files/live/sites/isoorg/files/store/en/PUB100400.pdf>. Acesso em: 20 jul. 2023.

JAKKULA, Vikramaditya. **Tutorial on support vector machine (SVM)**. Washington State University, [Pullman], 2011. Disponível em: <https://course.ccs.neu.edu/cs5100f11/resources/jakkula.pdf>. Acesso em: 28 jun. 2023.

KIRCHGÄSSNER, Gebhard; WOLTERS, Jürgen; HASSLER, Uwe. **Introduction to modern time series analysis**. Berlin: Springer, 2013. *E-book*. Disponível em: <https://content.E-bookshelf.de/media/reading/L-12229-7ecd026c65.pdf>. Acesso em: 21 jun. 2023.

KLUYVER, Thomas; RAGAN-KELLEY, Benjamin; PÉREZ, Fernando; GRANGER, Brian; BUSSONNIER, Matthias; FREDERIC, Jonathan; KELLEY, Kyle; HAMRICK, Jessica; GROUT, Jason; CORLAY, Sylvain; IVANOV, Paul; AVILA, Damián; ABDALLA, Safia; WILLING, Carol; Jupyter Development Team. Jupyter Notebooks-a publishing format for reproducible computational workflows. **International Conference on Electronic Publishing**, [s. l.], 2016. p. 87-90. DOI 10.3233/978-1-61499-649-1-87. Disponível em: <https://eprints.soton.ac.uk/403913/1/STAL9781614996491-0087.pdf>. Acesso em: 4 jul. 2023.

KLYUEV, Roman Vladimirovich; MORGIOEV, Irbek Dzhabrailovich; MORGIOVA, Angelika; GAVRINA, Oksana Aleksandrovna; MARTYUSHEV, Nikita V.; EFREMENKOV, Egor A.; QI, Mengxu. Methods of Forecasting Electric Energy Consumption: A Literature Review. **Energies**, [s. l.], v. 15, n. 23, p. 8919, 25 nov. 2022. Disponível em: <https://www.semanticscholar.org/reader/fba5553ec41367ba1c63cd1b90a1996d329d71af>. Acesso em: 22 jun. 2023.

LI, Hongze; GUO, Sen; ZHAO, Huiru; SU, Chenbo; WANG, Bao.. Annual Electric Load Forecasting by a Least Squares Support Vector Machine with a Fruit Fly Optimization Algorithm. **Energies**, [s. l.], v. 5, n. 11, p. 4430-4445, 2012. DOI: <https://doi.org/10.3390/en5114430>. Disponível em: <https://www.mdpi.com/1996-1073/5/11/4430>. Acesso em: 15 jul. 2023.

LI, Jian; ZHAO, Junbo; ZHANG, Zhenyuan; HU, Weihao; HUANG, Qi; DENG, Daiyu. A novel hybrid short-term load forecasting method of smart grid using MLR and RNA neural network. **IEEE Transactions on Industrial Informatics**, [s. l.], 2020. DOI [10.1109/TII.2020.3000184](https://doi.org/10.1109/TII.2020.3000184). *E-book*. Disponível em: https://www.researchgate.net/publication/341815013_A_Novel_Hybrid_Short-Term_Load_Forecasting_Method_of_Smart_Grid_Using_MLR_and_LSTM_Neural_Network. Acesso em: 15 jun. 2023.

LIN, Kuo-Huang; LIU, Bin-Da. A gray system modeling approach to the prediction of calibration intervals. **IEEE Transactions on Instrumentation and Measurement**, [s. l.], v. 54, n. 1, p. 297-304, 2005. DOI 10.1109/TIM.2004.840234. *E-book*. Disponível em: <https://ieeexplore.ieee.org/document/1381830>. Acesso em: 15 jun. 2023.

LORENA, Ana Carolina; CARVALHO, André C. P. L. F. de. **Introdução às Máquinas de Vetores Suporte (Support Vector Machines)**. São Carlos: Instituto de Ciências Matemáticas e de Computação – Universidade de São Paulo, 2003. (Relatórios Técnicos do ICMC, n. 192). Disponível em: https://repositorio.usp.br/directbitstream/a7ed198b-f6a3-4cec-b132-7e113bd51424/BIBLIOTECA_113_RT_192.pdf. Acesso em: 28 jun. 2023.

MACHADO, N. Eólica chega a 21 GW de capacidade instalada no Brasil. **Agência Eixos**, [Rio de Janeiro], 21 jan. 2022. Disponível em: <https://epbr.com.br/eolica-chega-a-21-gw-de-capacidade-instalada-no-brasil/>. Acesso em: 10 abr. 2023.

MAKRIDAKIS, Spyros; HIBON, Michele. ARMA models and the Box–Jenkins methodology. **Journal of forecasting**, Fontainebleau, v. 16, n. 3, p. 147-163, 1997. Disponível em: https://flora.insead.edu/fichiersti_wp/Inseadwp1995/95-33.pdf. Acesso em: 28 jun. 2023.

MAT DAUT, Mohammad Azhar; HASSAN, Mohammad Yusri; ABDULLAH, Hayati; RAHMAN, Hasimah Abdul; ABDULLAH, Md Pauzi; HUSSIN, Faridah. Building electrical energy consumption forecasting analysis using conventional and artificial intelligence methods: A review. **Renewable and Sustainable Energy Reviews**, [s. l.], v. 70, p. 1108–1118, 2017. DOI: <https://doi.org/10.1016/j.rser.2016.12.015>. Disponível em: <https://www.sciencedirect.com/science/article/abs/pii/S1364032116310619>. Acesso em: 14 mai.2023.

McCULLOCH, Warren S.; PITTS, Walter. A logical calculus of the ideas immanent in nervous activity. **The bulletin of mathematical biophysics**, [s. l.], v. 5, p. 115-133, 1943. DOI: <https://doi.org/10.1007/BF02478259>. Disponível em: <https://link.springer.com/article/10.1007/BF02478259>. Acesso em: 28 jun. 2023

MELONI, Raphael. **Classificação Supervisionada**. In: CLASSIFICAÇÃO DE IMAGENS DE SENSORIAMENTO REMOTO USANDO SVM. PONTIFÍCIA UNIVERSIDADE CATÓLICA DO RIO DE JANEIRO - PUC-RIO, [Rio de Janeiro], 2017. DOI: <https://doi.org/10.17771/PUCRio.acad.31439>. Disponível em: <https://www.maxwell.vrac.puc-rio.br/colecao.php?strSecao=resultado&nrSeq=31439&idi=1>. Acesso em: 28 jun. 2023.

MOHAMED, Z.; BODGER, P. Forecasting electricity consumption in New Zealand using economic and demographic variables. **Energy**, [s. l.], v. 30, n. 10, p. 1833–1843, jul. 2005. DOI: <https://doi.org/10.1016/j.energy.2004.08.012>. Disponível em: <https://www.sciencedirect.com/science/article/abs/pii/S0360544204003639>. Acesso em: 2 mai.. 2023.

MORETTIN, P. A.; TOLOI, C. M. C. **Análise de Séries Temporais**. 1. ed. São Paulo: Blucher, 2020. *E-book*. (Coleção Projeto Fisher). ISBN: 9786555060041. Disponível em: https://www.blucher.com.br/analise-de-series-temporais-volume-2_9786555060041. Acesso em: 13 jun. 2023.

NOGALES, Francisco Javier; CONTRERAS, Javier; CONEJO, Antonio J.; ESPÍNOLA, Rosario. Forecasting next-day electricity prices by time series models. **IEEE Transactions on power systems**, [s. l.] v. 17, n. 2, p. 342-348, 2002. Disponível em: <https://halweb.uc3m.es/fjnm/esp/papers/TransferPrices.pdf>. Acesso em: 22 jun. 2023.

OLIVEIRA, Adilson de; SILVEIRA, Getulio Borges da; BRAGA, Julia de Medeiros. Diversidade sazonal do consumo de energia elétrica no Brasil. **Pesquisa e Planejamento Econômico** (PPE), Rio de Janeiro, v. 30, n. 2, p. 211-258, ago. 2000. Disponível em: https://repositorio.ipea.gov.br/bitstream/11058/3904/1/PPE_v30_n02_Diversidade.pdf. Acesso em: 11 jun. 2023.

SCHMIDT, Cristiane Alkmin Junqueira; LIMA, Marcos AM. A demanda por energia elétrica no Brasil. **Revista Brasileira de Economia**, [s. l.], v. 58, p. 68-98, 2004. DOI: <https://doi.org/10.1590/S0034-71402004000100004>. Disponível em: <https://www.scielo.br/j/rbe/a/Bnvng4fCCnhvns9cRBYTSHB/?format=pdf&lang=pt>. Acesso em: 30 mai. 2023.

SEABOLD, Skipper; PERKTOLD, Josef. Statsmodels: Econometric and statistical modeling with Python. In: PYTHON IN SCIENCE CONFERENCE, 9., 2010, Austin. **Anais [...]**. Austin: SciPy, 2010. Disponível em: <https://pub.curvenote.com/01908379-2a1a-7a74-a157-6a0df64b92f2/public/seabold-34d6671a7bae7c2c09a284f57c0422d9.pdf>. Acesso em: 2 jul. 2023.

STANČIN, Igor; JOVIĆ, Alan. An overview and comparison of free Python libraries for data mining and big data analysis. In: INTERNATIONAL CONVENTION ON INFORMATION AND COMMUNICATION TECHNOLOGY, ELECTRONICS AND MICROELECTRONICS – MIPRO,

42., 2019, Opatija. **Anais [...]**. Opatija: IEEE, 2019. DOI 10.23919/MIPRO.2019.8757088. Disponível em: http://161.53.22.65/datoteka/1002799.Stancin_Jovic.pdf. Acesso em: 1 jul. 2023.

TSO, Geoffrey KF; YAU, Kelvin KW. Predicting electricity energy consumption: A comparison of regression analysis, decision tree and neural networks. **Energy**, [s. l.] v. 32, n. 9, p. 1761-1768, 2007. DOI: <https://doi.org/10.1016/j.energy.2006.11.010>. Disponível em: <https://www.sciencedirect.com/science/article/abs/pii/S0360544206003288>. Acesso em: 17 jun. 2023.

WEEDY, B. M. **Electric power systems**. Chichester: Wiley, 2012. *E-book*.

WERNER, Liane; RIBEIRO, José Luis Duarte. Previsão de demanda: uma aplicação dos modelos Box-Jenkins na área de assistência técnica de computadores pessoais. **Gestão & Produção**, [São carlos], v. 10, p. 47-67, 2003. DOI: <https://doi.org/10.1590/S0104-530X2003000100005>. *E-book*. Disponível em: <https://www.scielo.br/j/gp/a/XkNgqJXD9xCx8mkyHTk9xQQ/?format=pdf&lang=pt>. Acesso em: 24 jun. 2023.

WILAMOWSKI, Bogdan M. Neural network architectures and learning algorithms. **Industrial Electronics Magazine**, [s. l.], v. 3, n. 4, p. 56-63, 2009. DOI 10.1109/MIE.2009.934790. Disponível em: <https://www.eng.auburn.edu/~wilambm/pap/2009/Neural%20network%20architectures%20and%20learning%20algorithms.pdf>. Acesso em: 30 jun. 2023.

WINTERS, P. R. Forecasting sales by exponentially weighted moving averages. **Management Science**, [s. l.], v. 6, n. 3, p. 324–342, 1960. *E-book*. Disponível em: <http://www.jstor.org/stable/2627346>. Acesso em: 24 jun. 2023.

XIAO, J.; LI, Y.; XIE, L.; LIU, D.; HUANG, J. A hybrid model based on selective ensemble for energy consumption forecasting in China. **Energy**, [s. l.], v. 159, p. 534-546, 2018. DOI: <https://doi.org/10.1016/j.energy.2018.06.161>. *E-book*. Disponível em: <https://www.sciencedirect.com/science/article/abs/pii/S036054421831226X>. Acesso em: 25 jun. 2023.