

PREVISÃO DE CONSUMO DE ENERGIA ELÉTRICA NO ESTADO DO PIAUÍ: UMA ABORDAGEM BASEADA EM MACHINE LEARNING

Antonio Rawan Carvalho dos Santos¹
Jaciara Carvalho de Sousa Oliveira²

RESUMO: O consumo de energia elétrica no estado do Piauí apresenta variações relevantes ao longo do tempo, influenciadas por fatores econômicos, climáticos e comportamentais. Nesse contexto, métodos de previsão tornam-se essenciais para apoiar o planejamento energético e orientar decisões estratégicas de concessionárias e órgãos governamentais. Este trabalho aplica técnicas de *Machine Learning* para prever o consumo mensal de energia elétrica no Piauí utilizando dados reais referentes ao período de 2020 a 2024. O modelo desenvolvido foi baseado no algoritmo *Random Forest*, após etapas de pré-processamento, criação de variáveis defasadas e separação dos dados em treino e teste. A previsão para o ano de 2024 apresentou desempenho satisfatório, alcançando um Erro Percentual Absoluto Médio (MAPE) de 3,33%, evidenciando a alta capacidade do modelo em capturar a tendência de crescimento e a sazonalidade da série temporal. A análise de importância das variáveis confirmou que a tendência histórica e o mês do ano são os fatores mais determinantes para o consumo no estado. Os resultados demonstram que o uso de *Machine Learning* é uma alternativa eficaz para previsão de carga elétrica e pode contribuir para o planejamento energético no estado.

2590

Palavras-chave: Previsão de consumo. Energia elétrica. *Machine Learning*. *Random Forest*.

ABSTRACT: The electricity consumption in the state of Piauí shows significant variations over time, influenced by economic, climatic, and behavioral factors. In this context, forecasting methods become essential to support energy planning and guide strategic decisions made by utility companies and governmental agencies. This work applies Machine Learning techniques to predict the monthly electricity consumption in Piauí using real data from the period between 2020 and 2024. The developed model was based on the Random Forest algorithm, following preprocessing steps, creation of lagged variables, and separation of the data into training and testing sets. The forecast for the year 2024 showed satisfactory performance, achieving a Mean Absolute Percentage Error (MAPE) of 3.33%, demonstrating the model's strong ability to capture both the growth trend and the seasonality of the time series. The variable importance analysis confirmed that the historical trend and the month of the year are the most determining factors for consumption in the state. The results demonstrate that the use of Machine Learning is an effective alternative for electrical load forecasting and can contribute to the state's energy planning.

Keywords: Consumption forecasting. Electric energy. *Machine Learning*. *Random Forest*.

¹Acadêmico de Graduação do curso de Bacharelado em Engenharia Elétrica na UNISA - Centro Universitário Santo Agostinho. Teresina – PI.

²Mestre em Engenharia de Produção (UNIP), Engenheira Eletricista (UESPI).Universidade Paulista, Atualmente professora e orientadora na UNIFSA - Centro Universitário Santo.

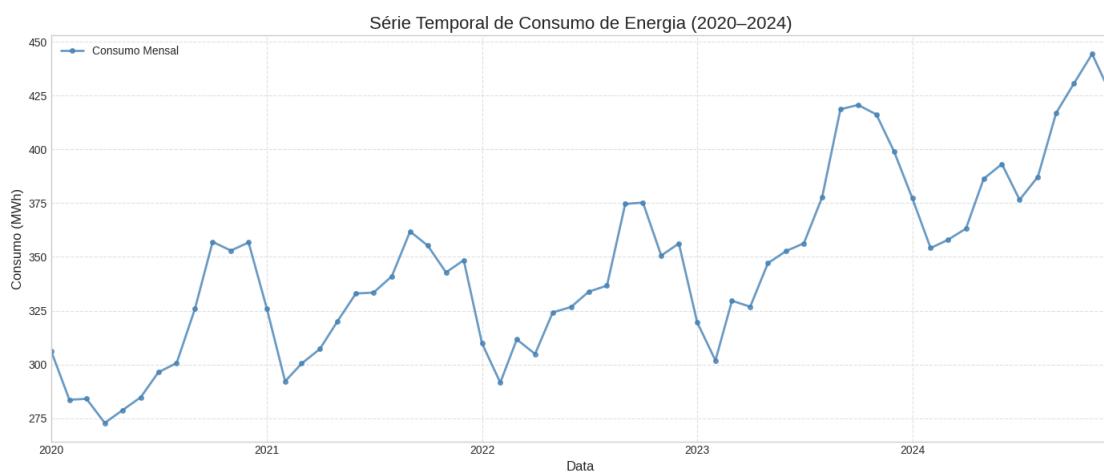
I. INTRODUÇÃO

O consumo de energia elétrica é reconhecido como um dos indicadores essenciais para o planejamento econômico e para o desenvolvimento social, uma vez que está diretamente relacionado ao desempenho das atividades produtivas e à qualidade de vida da população (EPE, 2023). No Brasil, devido às dimensões territoriais e à diversidade socioeconômica entre seus estados, o comportamento da demanda elétrica apresenta características distintas conforme a região analisada, exigindo estudos específicos que considerem suas particularidades estruturais (ONS, 2022). No contexto do Nordeste brasileiro, o estado do Piauí destaca-se pela expansão acelerada de sua matriz energética renovável, ao mesmo tempo em que apresenta padrões de consumo influenciados por fatores climáticos, populacionais e produtivos, o que justifica a necessidade de análises direcionadas ao seu comportamento energético recente.

A Figura 01 apresenta a série temporal de consumo total de energia elétrica no estado do Piauí entre os anos de 2020 e 2024, intervalo que corresponde ao período analisado neste estudo. A análise visual dos dados revela um comportamento caracterizado por oscilações sazonais bem definidas e uma tendência geral de crescimento, evidenciando a dinâmica de expansão contínua do sistema elétrico local. Esse panorama inicial reforça a complexidade da série e justifica a aplicação de técnicas robustas de previsão para auxiliar o planejamento energético estadual.

2591

Figura 01 – Série temporal do consumo total de energia elétrica no estado do Piauí entre 2020 e 2024.



Fonte: Autoria própria (2025)

Essa trajetória ascendente reflete fatores estruturais e conjunturais ocorridos no período, impulsionados pela retomada das atividades econômicas pós-pandemia, pelo avanço dos setores

industriais e comerciais e pelas mudanças nos hábitos de uso residencial da energia. Diante desse cenário, prever a demanda elétrica tornou-se uma etapa fundamental para o funcionamento eficiente do sistema energético, garantindo segurança operacional, planejamento adequado da infraestrutura e otimização dos recursos disponíveis. Estudos como os de Di Piazza et al. (2020) destacam que estimativas confiáveis permitem não apenas minimizar custos operacionais, mas também evitar investimentos desnecessários, garantindo maior estabilidade e qualidade no fornecimento de energia. Por outro lado, previsões imprecisas — tanto superestimadas quanto subestimadas — podem acarretar desperdícios financeiros, falhas no abastecimento e redução na confiabilidade do sistema, como discutido por Kaytez et al. (2015).

Historicamente, técnicas estatísticas tradicionais como Modelo Autorregressivo Integrado de Média Móvel (ARIMA), *Holt-Winters* e modelos de regressão têm sido amplamente empregadas em estimativas de curto e médio prazo (Hyndman & Athanasopoulos, 2018). Entretanto, o avanço da Inteligência Artificial e o amadurecimento das técnicas de *Machine Learning* (ML) vêm impulsionando o uso de algoritmos capazes de modelar relações complexas e não lineares, superando limitações de métodos puramente estatísticos (Goodfellow, Bengio & Courville, 2016). Pesquisas comparativas demonstram que algoritmos como *Random Forest*, *Gradient Boosting* e redes neurais frequentemente apresentam desempenho competitivo ou superior na previsão de demandas energéticas, destacando-se pela robustez ao lidar com séries temporais sujeitas a variações abruptas e padrões sazonais irregulares (Amasyali & El-Gohary, 2018).

2592

Nesse contexto, torna-se pertinente investigar a aplicação dessas técnicas ao consumo de energia elétrica do Piauí, considerando dados recentes compreendidos entre 2020 e 2024. A seleção desse período fundamenta-se na disponibilidade de dados consolidados pela Empresa de Pesquisa Energética (EPE) e no fato de que o intervalo abrange mudanças estruturais e alta volatilidade na série temporal, características que impõem desafios significativos ao desempenho dos modelos preditivos convencionais. Além disso, a incorporação de métodos modernos de aprendizado de máquina possibilita a elaboração de previsões mais acuradas, oferecendo suporte estratégico para o planejamento energético estadual.

A relevância deste estudo justifica-se também pela escassez de pesquisas focadas especificamente no estado do Piauí, uma vez que grande parte da literatura se concentra em análises nacionais ou em regiões com maior densidade industrial. Assim, ao propor um modelo

de previsão dedicado ao contexto piauiense, este trabalho contribui tanto para o avanço científico quanto para a geração de subsídios que apoiam concessionárias de energia, gestores públicos e agentes dependentes de análises do comportamento elétrico. Dessa forma, espera-se que os resultados fundamentem decisões estruturais no setor, contribuindo para uma gestão mais eficiente e sustentável da demanda energética regional.

2. REFERENCIAL TEÓRICO

2.1 O Setor Elétrico Brasileiro

O setor elétrico brasileiro é caracterizado por sua estrutura complexa, composta por agentes responsáveis pela geração, transmissão, distribuição e comercialização de energia elétrica, todos regulados por diretrizes federais que asseguram a continuidade e qualidade do serviço. De acordo com a Empresa de Pesquisa Energética (EPE, 2024), a matriz elétrica do Brasil é predominantemente renovável, com destaque para a geração hidrelétrica, que historicamente responde por grande parte da capacidade instalada do país. Nos últimos anos, entretanto, observa-se um avanço expressivo das fontes eólica e solar, impulsionado por políticas de incentivo e pela crescente competitividade tecnológica. Essas transformações ampliam a diversidade da matriz, mas também aumentam a complexidade operacional do sistema, especialmente em relação ao equilíbrio entre oferta e demanda.

2593

A operação do Sistema Interligado Nacional (SIN), composto pelas principais regiões do país, é coordenada pelo Operador Nacional do Sistema (ONS), responsável por garantir o suprimento de energia com segurança, eficiência e custo otimizado. Segundo o ONS (2023), a integração das diferentes fontes e a gestão das variações naturais — como períodos de estiagem que afetam a geração hídrica — tornam a previsão de demanda um elemento estratégico para o planejamento energético. A Agência Nacional de Energia Elétrica (ANEEL), por sua vez, exerce função regulatória, monitorando tarifas, contratos e padrões de desempenho das concessionárias, de modo a assegurar equilíbrio entre sustentabilidade econômica do setor e modicidade tarifária para os consumidores (ANEEL, 2022).

Além desses fatores, a expansão urbana, o aumento da eletrificação residencial e a maior dependência de tecnologias digitais intensificam a demanda nacional de energia. Dados do Balanço Energético Nacional (EPE, 2024) indicam que, mesmo com oscilações econômicas, o consumo elétrico mantém tendência de crescimento estrutural. Em regiões como o Nordeste, onde se encontra o estado do Piauí, essa elevação é ainda mais perceptível devido à

diversificação econômica e ao avanço de setores comerciais e industriais de médio porte. Assim, compreender o funcionamento do setor elétrico e seus desafios estruturais é essencial para contextualizar a necessidade de modelos preditivos capazes de apoiar decisões relacionadas ao planejamento e operação do sistema elétrico brasileiro.

2.2 Previsão de Demanda e sua Importância

A previsão de demanda de energia elétrica é uma atividade central no planejamento e na operação dos sistemas de potência, pois permite antecipar o comportamento futuro do consumo e orientar decisões estratégicas relacionadas à expansão da geração, reforço das redes de transmissão e distribuição, gestão de reservatórios e definição de tarifas. Segundo Bunn e Farmer (1985), a acurácia dessas projeções é um fator determinante para a eficiência econômica do setor, visto que o aumento no erro de previsão acarreta elevação desproporcional nos custos operacionais, seja pela necessidade de despacho de unidades geradoras mais onerosas no curto prazo, seja pelo desperdício de capital em infraestrutura ociosa no longo prazo.

De acordo com Weron (2014), o processo de previsão pode ser classificado em três horizontes temporais distintos: curto prazo (horas a dias), focado na operação diária e despacho de carga; médio prazo (semanas a meses), essencial para a programação energética e manutenção do parque gerador; e longo prazo (anos), direcionado ao planejamento da expansão do sistema. Para cada horizonte, modelos específicos são empregados, variando desde métodos estatísticos tradicionais até algoritmos de aprendizado de máquina capazes de capturar padrões não lineares nos dados.

A relevância da previsão torna-se ainda mais crítica em cenários de crescente integração de fontes renováveis variáveis, como eólica e solar. Conforme apontam Suganhi e Samuel (2012), a intermitência dessas fontes eleva a incerteza operacional, exigindo modelos preditivos robustos para lidar com volatilidades rápidas no perfil de consumo e geração. Simultaneamente, o avanço das redes inteligentes (*smart grids*) tem ampliado a disponibilidade de dados, fortalecendo o papel de métodos computacionais modernos na busca por estimativas mais detalhadas.

No contexto brasileiro, estudos da EPE (2023) evidenciam que a previsão de demanda é um pilar para a segurança energética nacional e a modicidade tarifária. A correta antecipação do crescimento do consumo é determinante para o planejamento de investimentos, evitando desequilíbrios entre a oferta disponível e a capacidade instalada necessária.

Dante desse cenário, a aplicação de modelos de previsão ajustados à realidade regional torna-se indispensável. Em estados como o Piauí, onde o consumo de energia tem crescido de forma expressiva, estimativas confiáveis são ferramentas vitais para auxiliar gestores públicos e concessionárias na tomada de decisões eficientes, sustentáveis e alinhadas ao desenvolvimento econômico local.

2.3 Métodos de Previsão de Demanda (Quantitativos e Qualitativos)

Os métodos de previsão de demanda podem ser classificados, tradicionalmente, em duas grandes categorias: qualitativos e quantitativos. Cada abordagem possui características próprias, sendo escolhida de acordo com a natureza dos dados disponíveis, o horizonte de previsão e o nível de precisão desejado. De acordo com Hyndman e Athanasopoulos (2018), métodos qualitativos são mais adequados quando há escassez de dados históricos ou quando se deseja obter estimativas baseadas em julgamentos especializados. Já técnicas quantitativas são preferidas em cenários nos quais existe uma base histórica consistente, permitindo a modelagem matemática da demanda.

Os métodos qualitativos incluem opiniões de especialistas, método *Delphi*, pesquisa de mercado, cenários prospectivos e análises estruturadas baseadas em visão gerencial (Armstrong, 2001). Embora sejam úteis em ambientes incertos e em situações nas quais a evolução da demanda depende de fatores subjetivos, suas estimativas podem apresentar elevado grau de variabilidade e menor rigor estatístico. Por essa razão, são geralmente utilizados como apoio ao processo decisório, complementando previsões quantitativas.

Já os métodos quantitativos podem ser subdivididos em modelos causais e modelos de séries temporais. Modelos causais — como regressão linear, modelos econométricos e redes neurais multivariadas — procuram relacionar a demanda a variáveis explicativas, como temperatura, renda ou indicadores socioeconômicos (Suganthy & Samuel, 2012). Em contrapartida, os modelos de séries temporais concentram-se no comportamento histórico da própria demanda, analisando tendências e sazonalidades (Box, Jenkins & Reinsel, 2015). Entre esses, destacam-se desde técnicas clássicas, como ARIMA e suavização exponencial, até algoritmos de *Machine Learning* como *Random Forest* e Redes Neurais. Segundo Zhang, Patuwo e Hu (1998), a grande vantagem desses métodos computacionais reside justamente na capacidade de capturar relações não lineares e complexas de forma flexível, sem a necessidade de assumir hipóteses rígidas sobre a estrutura dos dados.

A escolha entre métodos qualitativos e quantitativos depende, portanto, do objetivo do estudo e do tipo de dados disponíveis. Em contextos com histórico robusto e regular — como é o caso da série mensal de consumo de energia do Piauí — abordagens quantitativas tendem a oferecer maior precisão. Assim, técnicas baseadas em aprendizado de máquina tornam-se particularmente adequadas, especialmente quando combinadas com engenharia de atributos e validação rigorosa, conforme demonstrado por estudos recentes no setor energético (Kaytez et al., 2015; Weron, 2014).

2.4 Séries Temporais

Séries temporais são conjuntos de observações registradas ao longo do tempo, geralmente em intervalos regulares, como dias, meses ou anos. A principal característica desse tipo de dado é a dependência temporal: valores futuros tendem a ser influenciados pelo comportamento passado, conforme destacado por Box, Jenkins e Reinsel (2015), que apontam essa relação como essencial para a formulação de modelos preditivos. No contexto energético, essa dependência se torna particularmente relevante, pois o consumo de eletricidade reflete tanto padrões socioeconômicos quanto variações climáticas e hábitos de uso da população.

Segundo Hyndman e Athanasopoulos (2018), séries temporais podem apresentar quatro componentes estruturais: tendência, sazonalidade, ciclos e ruído. A tendência representa movimentos persistentes de crescimento ou declínio; a sazonalidade corresponde a padrões que se repetem em intervalos regulares, como meses ou estações do ano; os ciclos refletem flutuações econômicas de longo prazo; e o ruído consiste em variações aleatórias que não podem ser explicadas por outros componentes. A análise adequada desses elementos é essencial para a construção de previsões robustas.

Em séries com janela curta, como a analisada neste estudo, é natural que componentes cíclicos não se manifestem de forma estatisticamente significativa. Isso não compromete a modelagem, uma vez que os modelos de previsão capturam adequadamente os componentes realmente presentes e relevantes — tendência e sazonalidade. No consumo energético do Piauí, por exemplo, observa-se variações sazonais associadas a períodos mais quentes do ano, além de um padrão geral de crescimento relacionado ao aumento populacional e expansão do comércio.

A literatura destaca que a correta identificação desses componentes é fundamental para o desempenho de qualquer método preditivo, seja ele estatístico ou baseado em aprendizado de máquina. Hyndman e Athanasopoulos (2018) reforçam que a compreensão estrutural da série

permite selecionar técnicas adequadas de modelagem e auxilia na interpretação dos resultados. No caso deste trabalho, essa análise preliminar embasa o uso do algoritmo *Random Forest* (*RF*) aplicado à previsão de consumo mensal, pois essa técnica é capaz de explorar relações não lineares entre variáveis temporais derivadas, como mês, ano e defasagens, beneficiando-se diretamente da estrutura observada na série.

Assim, o estudo das séries temporais não apenas fundamenta teoricamente a modelagem adotada neste trabalho, mas também orienta a preparação dos dados e a definição das variáveis utilizadas no modelo de machine learning, assegurando maior precisão e consistência na previsão do consumo elétrico no estado do Piauí.

2.5 Machine Learning Aplicado à Previsão de Consumo de Energia

O uso de técnicas de *ML* na previsão de séries temporais tem se consolidado como uma abordagem eficaz para modelar relações complexas e não lineares, especialmente em cenários onde variáveis explicativas desempenham papel relevante no comportamento futuro da série. Segundo Makridakis et al. (2020), métodos de *ML* apresentam vantagens importantes sobre técnicas estatísticas tradicionais, sobretudo pela capacidade de capturar padrões intrincados sem a necessidade de pressupor uma estrutura paramétrica específica. Diferentemente de modelos como ARIMA, que dependem de estacionariedade, ordem de diferenciação e parâmetros autorregressivos, algoritmos de *ML* adaptam-se a diferentes formatos de dados e conseguem incorporar múltiplas variáveis derivadas, tornando-se particularmente adequados para a dinâmica sazonal e estrutural do consumo de energia.

2597

O processo de modelagem baseado em no método supracitado geralmente envolve etapas fundamentais como preparação dos dados, criação de atributos (*feature engineering*), ajuste de hiperparâmetros e avaliação do modelo. Conforme Géron (2019), a engenharia de atributos é uma das etapas mais determinantes para o desempenho final, pois permite ao algoritmo extrair padrões relevantes a partir da série original. No presente estudo, foram utilizados atributos como mês, ano e valores defasados (*lags*), que auxiliam o modelo a compreender a dependência temporal e a sazonalidade característica do consumo energético no Piauí.

Entre os diversos algoritmos disponíveis, o *Random Forest* tem se destacado por sua robustez, interpretabilidade relativa e boa capacidade de generalização. De acordo com Breiman (2001), o método constrói múltiplas árvores de decisão treinadas de forma independente, combinando-as por meio de votação ou média, o que reduz o risco de sobreajuste (*overfitting*) e

melhora a estabilidade das previsões. Sua estrutura baseada em árvores permite capturar relações não lineares entre variáveis, além de lidar bem com conjuntos de dados relativamente pequenos — característica presente na base analisada neste trabalho, composta por observações mensais entre 2020 e 2024.

Conforme aponta Zhang et al. (2019), modelos de ML são particularmente eficazes quando a série temporal possui comportamento influenciado por múltiplos fatores temporais ou quando há necessidade de flexibilidade para capturar padrões sazonais e tendências variáveis ao longo do tempo. Essa abordagem se mostrou adequada para o caso do consumo de energia no Piauí, que segue uma dinâmica impactada tanto por variações climáticas quanto pelo crescimento socioeconômico regional. A combinação entre engenharia de atributos e o algoritmo RF permitiu construir um modelo capaz de representar tais padrões de forma consistente, resultando em previsões alinhadas ao comportamento real observado.

Assim, o uso de ML representa uma evolução significativa nas técnicas de previsão de carga, oferecendo maior adaptabilidade, capacidade de modelagem e desempenho preditivo. A adoção dessa abordagem neste estudo fundamenta-se em evidências científicas consolidadas e atende às necessidades específicas da análise do consumo energético estadual, contribuindo para maior precisão e confiabilidade no planejamento energético.

2598

2.6 Random Forest Aplicado a Séries Temporais

A aplicação de técnicas de aprendizado de máquina em séries temporais tem ganhado destaque nos últimos anos, principalmente devido à capacidade desses métodos de capturar padrões complexos que vão além da linearidade tradicionalmente assumida por modelos estatísticos. Dentre essas técnicas, o algoritmo RF é amplamente utilizado por sua robustez, capacidade de generalização e eficiência no tratamento de variáveis preditoras diversas. De acordo com Breiman (2001), criador do método, o *Random Forest* baseia-se na construção de múltiplas árvores de decisão que, ao serem combinadas, produzem previsões mais estáveis e menos suscetíveis a variações aleatórias dos dados. Embora inicialmente desenvolvido para tarefas de classificação e regressão em dados independentes, sua aplicação em séries temporais torna-se possível mediante a transformação da série em um conjunto de atributos que capturam dependências temporais relevantes.

No contexto de séries temporais, a adaptação fundamental para utilizar o RF consiste na criação de variáveis derivadas da própria série, como defasagens (*lags*), variações percentuais,

médias móveis e indicadores sazonais. Esse procedimento, conhecido como engenharia de atributos temporais, permite ao modelo aprender relações entre valores passados e o comportamento futuro da variável-alvo. Ao incorporar esses atributos, o RF consegue capturar padrões complexos, irregularidades e mudanças estruturais presentes nos dados — características que frequentemente desafiam métodos estatísticos tradicionais, os quais dependem de pressupostos paramétricos mais rígidos e apresentam menor flexibilidade frente a dinâmicas não lineares (Makridakis et al., 2020).

É importante ressaltar que, conforme discutido no capítulo 2.4, séries temporais são compostas por elementos estruturais como tendência, sazonalidade, ciclos e ruído (Hyndman & Athanasopoulos, 2018). Entretanto, em séries com janela temporal limitada, como a analisada neste trabalho, é comum que ciclos de longo prazo não apareçam de forma estatisticamente evidente. Em séries com janela curta, como a analisada neste estudo, é natural que componentes cíclicos não se manifestem de forma estatisticamente significativa. Isso não compromete a modelagem, uma vez que os modelos de previsão capturam adequadamente os componentes realmente presentes e relevantes — tendência e sazonalidade. Dessa forma, ao transformar a série em atributos, o *Random Forest* consegue absorver os padrões essenciais necessários para realizar previsões confiáveis mesmo sem explorar toda a estrutura teórica possível.

2599

Além disso, o RF apresenta vantagens práticas relevantes para o planejamento energético. Por não exigir suposições rígidas sobre normalidade, estacionariedade ou linearidade, ele se adapta bem a dados reais, frequentemente marcados por variações bruscas e comportamento irregular — características comuns no consumo de energia elétrica em estados em desenvolvimento, como o Piauí. Kaytez et al. (2015) destacam que técnicas de aprendizado de máquina podem oferecer previsões mais precisas em cenários complexos, especialmente quando associadas a séries temporais influenciadas por fatores climáticos, socioeconômicos e comportamentais. Isso reforça a pertinência do uso do RF no presente estudo, dado que o consumo energético regional é afetado por fatores variados, incluindo aumento populacional, expansão comercial e variações sazonais.

Por fim, a literatura aponta que a engenharia de atributos temporais, em combinação com modelos flexíveis como o RF, forma uma abordagem eficiente e validada para a previsão de séries temporais. Géron (2019) destaca a engenharia de atributos como fundamental para o desempenho preditivo, permitindo que modelos baseados em árvores explorem relações complexas de forma não paramétrica. No contexto deste trabalho, o RF foi selecionado não

apenas por sua relevância acadêmica, mas também por sua adequação às características da série histórica analisada. Sua capacidade de lidar com dependências temporais sem exigir modelagem paramétrica detalhada torna-o particularmente útil para estudos energéticos regionais, nos quais a disponibilidade de dados pode ser limitada e a dinâmica da série pode apresentar não linearidades significativas. Dessa forma, o uso do modelo mencionado acima se consolida como uma escolha metodológica alinhada ao objetivo do estudo: gerar previsões precisas e aplicáveis ao planejamento do consumo de energia elétrica no estado do Piauí.

2.7 Métricas de Avaliação de Modelos Preditivos

A avaliação de modelos de previsão é uma etapa essencial em qualquer estudo que envolve séries temporais, pois permite medir objetivamente o desempenho das técnicas aplicadas e determinar se as estimativas produzidas são adequadas para fins de planejamento e tomada de decisão. No contexto da previsão de consumo de energia elétrica, métricas bem definidas são fundamentais para garantir que os modelos adotados forneçam previsões consistentes e confiáveis. Segundo Makridakis et al. (2020), o uso de métricas quantitativas é vital para comparar o desempenho de diferentes abordagens, identificar a acurácia dos modelos e selecionar a técnica mais eficiente para um determinado conjunto de dados. Entre as métricas mais utilizadas na literatura e aplicadas em estudos energéticos destacam-se o Erro Médio Absoluto (MAE), a Raiz do Erro Quadrático Médio (RMSE) e o Erro Percentual Absoluto Médio (MAPE).

O Erro Médio Absoluto (MAE) é uma métrica que calcula a média das diferenças absolutas entre os valores previstos e os valores observados. No contexto da avaliação de modelos, Géron (2019) enfatiza que o MAE é frequentemente empregado por ser uma medida de erro simples e intuitiva. Ele expressa, em unidades da própria variável-alvo, o desvio médio cometido pelo modelo, permitindo interpretar diretamente o nível de erro na escala do consumo de energia elétrica. Por essa razão, é útil em análises que buscam avaliar o desempenho geral do modelo sem penalizar de forma excessiva erros maiores.

A Raiz do Erro Quadrático Médio (RMSE) é uma das métricas mais tradicionais na avaliação de modelos preditivos e é particularmente sensível a grandes erros. Em séries temporais com variações significativas e comportamento influenciado por múltiplos fatores (características do cenário analisado por Zhang et al., 2019), o RMSE é relevante por penalizar desvios maiores de forma mais intensa devido à etapa de quadratização. Essa sensibilidade a

grandes discrepâncias torna a métrica importante em cenários onde erros elevados podem acarretar impactos significativos, como decisões operacionais no setor energético. Por expressar o erro na mesma unidade da variável prevista, o RMSE também facilita a interpretação prática do desempenho do modelo.

O Erro Percentual Absoluto Médio (MAPE) é uma métrica que calcula o erro médio em termos percentuais, permitindo avaliar o desempenho do modelo independentemente da escala da série. Makridakis et al. (2020), ao analisarem a performance de diversos métodos em competições de previsão, demonstram a importância de métricas de erro relativas para possibilitar a comparação de resultados entre séries com escalas ou variações sazonais distintas. Este indicador é, portanto, amplamente difundido em pesquisas de previsão energética por expressar o erro em um formato percentual mais comprehensível para gestores e tomadores de decisão. No entanto, é crucial notar que essa métrica pode ser sensível a valores muito baixos na série, situação que exige cautela na sua interpretação.

A escolha combinada das métricas MAE, RMSE e MAPE está alinhada às recomendações de Makridakis et al. (2020), que destacam a importância de utilizar múltiplos indicadores para obter uma avaliação mais completa e imparcial do desempenho do modelo. Dessa forma, este estudo adota as três métricas para garantir uma análise equilibrada, 2601 capturando diferentes aspectos do erro — magnitude média, penalização de grandes discrepâncias e interpretação percentual. Essa abordagem assegura maior robustez na avaliação das previsões e permite selecionar o modelo mais adequado às necessidades da análise energética no estado do Piauí.

3. METODOLOGIA

3.1 Caracterização da Pesquisa

Este estudo configura-se como uma pesquisa aplicada, pois busca gerar conhecimento voltado à solução de um problema real: a previsão do consumo de energia elétrica no estado do Piauí. A abordagem é quantitativa, fundamentada na análise numérica de séries temporais e na utilização de métodos de *Machine Learning* para construção dos modelos de previsão. Quanto aos objetivos, trata-se de uma pesquisa descritiva, uma vez que se propõe a analisar padrões históricos, identificar sazonalidades, investigar tendências e descrever o comportamento da demanda energética ao longo do tempo com base em técnicas preditivas. Essa classificação se

alinha ao que Gil (2008) define como pesquisas destinadas a observar e descrever fenômenos sem manipulá-los diretamente.

3.2 Dados Utilizados

Os dados empregados neste trabalho foram obtidos por meio da Empresa de Pesquisa Energética (EPE), órgão vinculado ao Ministério de Minas e Energia e responsável pela consolidação das estatísticas energéticas nacionais, contidas na base de Consumo Mensal de Energia Elétrica por Classe (Regiões e Subsistemas).

Para a construção do modelo, foi utilizada a série histórica do consumo total de energia elétrica referente ao estado do Piauí, abrangendo as modalidades Cativo e Livre, no período de janeiro de 2020 a dezembro de 2024. Todos os dados e resultados apresentados neste trabalho estão expressos na unidade de Megawatts-hora (MWh).

As séries mensais, disponibilizadas publicamente em formato de planilha, foram importadas para o ambiente *Python* e, em seguida, convertidas para o tipo *datetime*, permitindo a ordenação cronológica e a manipulação adequada para análises de séries temporais. A inspeção inicial não revelou valores ausentes, duplicados ou inconsistentes, indicando que a base fornecia condições adequadas para aplicação dos métodos propostos.

2602

3.3 Procedimentos Metodológicos

O tratamento dos dados envolveu a criação de variáveis derivadas da própria série temporal, necessárias para que o modelo *RF* conseguisse capturar padrões temporais, já que este não é um modelo intrinsecamente sequencial. Foram construídas variáveis defasadas (*lags*) para representar o consumo do mês anterior, dos dois, três, seis e doze meses anteriores, permitindo capturar dependências de curto e longo prazo. Além disso, foram incluídas variáveis de calendário, como mês e ano, além de uma variável de tendência construída como uma sequência crescente, útil para representar o comportamento global da série, conforme recomenda Hyndman e Athanasopoulos (2018). Após a criação dessas features, a base final foi dividida temporalmente: os dados de 2020 a 2023 foram utilizados no treinamento do modelo, enquanto os dados de 2024 foram reservados para teste, assegurando que a avaliação respeitasse a lógica temporal da série.

O modelo escolhido para previsão foi o *Random Forest Regressor*, proposto por Breiman (2001), devido à sua robustez, capacidade de lidar com relações não lineares e eficiência

comprovada em previsões energéticas em estudos recentes (Makridakis et al., 2020). Para maximizar o desempenho do modelo, realizou-se a otimização dos hiperparâmetros utilizando o método RandomizedSearchCV com validação cruzada ($k\text{-fold} = 3$). Foram testadas 50 combinações aleatórias de parâmetros, buscando minimizar o erro quadrático médio negativo. A grade de busca incluiu variações no número de estimadores (100 a 1000), profundidade máxima das árvores, critérios de divisão e número mínimo de amostras por folha. O algoritmo selecionou automaticamente a melhor configuração para a generalização dos dados. A validação do modelo foi realizada por meio das métricas MAE, RMSE e MAPE, garantindo uma análise completa sobre o erro absoluto, quadrático e percentual entre as previsões e os valores reais, conforme recomendação de Willmott e Matsuura (2005).

3.4 Ambiente de Desenvolvimento e Ferramentas Utilizadas

Todas as etapas de preparação, processamento, modelagem e avaliação foram realizadas utilizando a linguagem *Python*, executada no ambiente *Google Colab*, que oferece acesso a recursos computacionais em nuvem e integração direta com o *Google Drive*. O *Colab* foi escolhido por permitir a organização do código em células, execução incremental e reproduzibilidade do projeto, além de dispensar instalação local de ferramentas.

2603

A biblioteca (*pandas*) foi o principal instrumento para manipulação e estruturação dos dados. A base de consumo, previamente tratada para conter exclusivamente os valores referentes ao estado do Piauí, foi carregada utilizando o comando *pd.read_csv()*. Em seguida, procedimentos essenciais de formatação foram executados, incluindo a renomeação das colunas (*df.columns = [...]*), conversão da variável temporal para o tipo *datetime* (*pd.to_datetime()*), ordenação cronológica (*df.sort_values()*), remoção de valores iniciais inválidos decorrentes das defasagens (*df.dropna()*) e definição das datas como índice da série (*df.set_index()*). Ademais, *pandas* foi fundamental para a engenharia de atributos, permitindo a criação das variáveis defasadas com *shift()* — como *lag_1*, *lag_2*, *lag_3*, *lag_6* e *lag_12* — e a extração de componentes sazonais diretos, como o mês (*df.index.month*) e o ano (*df.index.year*), utilizados para separar treino e teste.

A biblioteca (*NumPy*) desempenhou papel complementar nas operações matemáticas necessárias ao modelo. Sua principal contribuição foi a criação da variável de tendência, construída por meio do comando *np.arange(len(df))*, que gera uma sequência crescente ao longo do tempo. Essa variável adiciona ao modelo informação sobre padrões de crescimento ou

declínio persistentes na série, sendo frequentemente utilizada em problemas de previsão energética para representar comportamentos estruturais de longo prazo.

Para o processo de modelagem, otimização e avaliação, foi empregada a biblioteca (*scikit-learn*), que forneceu todos os recursos necessários ao treinamento do *RF*. O algoritmo principal foi importado pelo módulo *ensemble*, por meio de *RandomForestRegressor()*, responsável por construir uma coleção de árvores de decisão capazes de capturar relações não lineares entre o consumo, as defasagens e as variáveis sazonais. A otimização dos hiperparâmetros foi realizada com *RandomizedSearchCV()*, fornecido pelo módulo (*model_selection*), que permitiu testar combinações de profundidade das árvores, número de estimadores, critérios de divisão, limites de amostras e quantidade de atributos considerados a cada divisão. Após a busca, o comando *random_search.best_estimator()* selecionou automaticamente o modelo com melhor desempenho. Para avaliação, o módulo *metrics* disponibilizou funções específicas: *mean_absolute_error()*, *mean_squared_error()* e *mean_absolute_percentage_error()*, utilizadas para calcular MAE, RMSE e MAPE — métricas amplamente consolidadas em estudos de previsão de demanda energética.

A etapa final, de visualização dos resultados, foi desenvolvida com (*Matplotlib*), que permitiu comparar graficamente os valores reais de 2024 com as estimativas do modelo otimizado. Comandos como *plt.figure()*, *plt.plot()*, *plt.xlabel()*, *plt.ylabel()* e *plt.title()* foram responsáveis pela construção do gráfico final, enquanto *plt.grid()*, *plt.legend()* e *plt.xticks(rotation=45)* aprimoraram a legibilidade, enfatizando alinhamentos e discrepâncias entre a série prevista e a série observada. Essa representação visual é essencial para avaliar qualitativamente o desempenho do modelo, sobretudo em séries com forte sazonalidade mensal. 2604

Assim, o ambiente de desenvolvimento integrou ferramentas robustas e amplamente validadas na literatura científica, permitindo a construção de um fluxo completo que abrangeu carregamento, manipulação, transformação, modelagem, otimização e visualização dos dados, assegurando rigor metodológico e consistência nas análises realizadas.

4. RESULTADO E DISCUSSÕES

4.1 Análise Descritiva da Série

A série temporal de consumo de energia elétrica no estado do Piauí, de janeiro de 2020 a dezembro de 2024, apresentou um padrão consistente de tendência crescente e sazonalidade mensal. Os gráficos iniciais evidenciam que o consumo aumenta gradualmente ao longo do tempo, refletindo a expansão da demanda residencial e industrial, enquanto oscilações mensais

sugerem influência de fatores sazonais, como variações climáticas e períodos de maior atividade econômica. A série se mostrou estável, sem lacunas ou rupturas significativas, o que permitiu aplicar o modelo *RF* com confiabilidade. A inclusão de variáveis de tendência e *lags* (1, 2, 3, 6 e 12 meses) possibilitou capturar os ciclos mensais e padrões de curto e longo prazo, essenciais para a previsão de séries temporais.

4.2 Resultados do Modelo

O modelo *Random Forest* otimizado apresentou a seguinte configuração de hiperparâmetros, selecionada como a mais eficiente durante a etapa de validação:

Número de estimadores (árvore): 700

Profundidade máxima: *None* (ilimitada, permitindo expansão total das árvores)

Mínimo de amostras para divisão: 5

Features máximas: *None* (considera todas as variáveis em cada divisão)

A aplicação deste modelo aos dados de teste (ano de 2024) gerou as previsões detalhadas na Tabela 01.

Tabela 01 – Comparativo Real vs. Previsto (2024)

Mês/2024	Consumo Real (MWh)	Previsão RF (MWh)	Erro Absoluto (MWh)	2605
Janeiro	377,33	369,32	8,01	
Fevereiro	354,12	366,87	12,75	
Março	357,88	359,19	1,31	
Abril	363,19	361,32	1,87	
Maio	386,31	372,74	13,57	
Junho	393,22	389,35	3,87	
Julho	376,58	396,3	19,72	
Agosto	387,29	402,23	14,95	
Setembro	416,97	405,7	11,27	
Outubro	430,61	408,71	21,9	
Novembro	444,42	409,22	35,2	
Dezembro	426,79	409,35	17,44	

Fonte: Autoria própria (2025)

As métricas de avaliação consolidam o desempenho satisfatório do modelo, com redução dos erros em comparação aos testes preliminares:

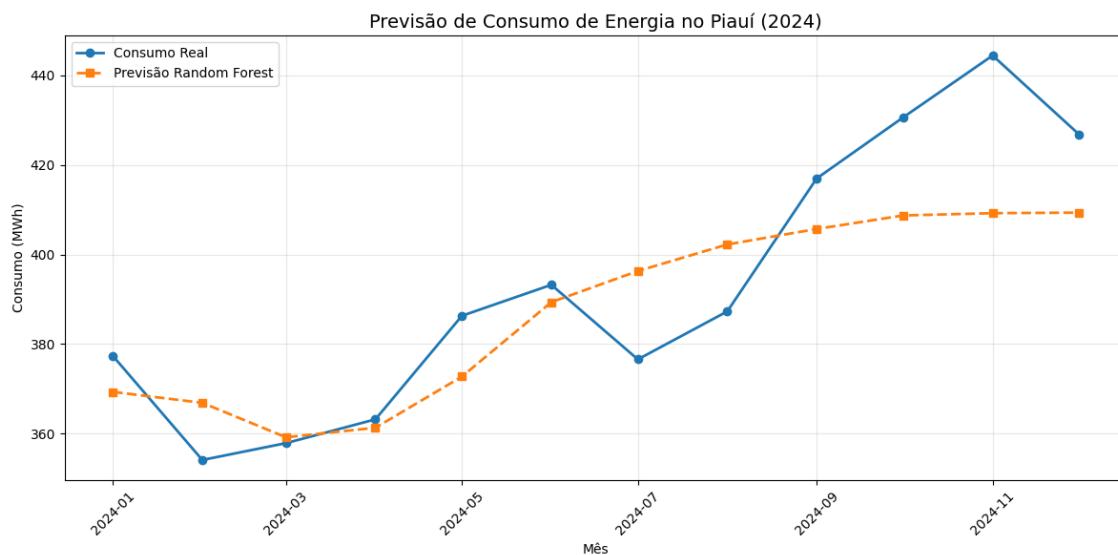
MAE (Erro Médio Absoluto): 13,49 MWh

RMSE (Raiz do Erro Quadrático Médio): 16,31 MWh

MAPE (Erro Percentual Absoluto Médio): 3,33%

Para complementar a análise numérica, a Figura 02 apresenta a comparação visual entre a série de consumo realizada (linha azul contínua) e as estimativas geradas pelo modelo RF (linha laranja tracejada) ao longo dos 12 meses de 2024.

Figura 02 - Comparativo entre Consumo Real e Previsão do Modelo Random Forest (2024).



Fonte: Autoria própria (2025)

2606

A análise gráfica revela que o modelo conseguiu capturar com êxito a dinâmica estacional da série, reproduzindo corretamente a queda de consumo no primeiro trimestre (fevereiro e março) e a retomada de crescimento a partir de abril. Observa-se uma aderência expressiva no primeiro semestre, onde as curvas praticamente se sobrepõem, o que justifica o baixo erro percentual médio obtido.

No entanto, nota-se um comportamento de suavização (*smoothing*) nos meses finais do ano. Enquanto o consumo real apresentou um pico acentuado em outubro e novembro — atingindo 444,42 MWh —, o modelo projetou uma curva de crescimento mais conservadora, estabilizando em torno de 409 MWh. Esse fenômeno sugere que, embora a tendência de alta tenha sido identificada, a magnitude exata dos picos de fim de ano pode ter sido influenciada por fatores exógenos atípicos (como ondas de calor extremas) que extrapolam o padrão histórico aprendido pelo algoritmo.

Observa-se que o modelo consegue reproduzir de forma consistente a tendência geral de crescimento e os padrões sazonais mensais, apresentando alta aderência na maior parte do ano. Os desvios mais expressivos ocorreram nos meses de novembro (35,20 MWh) e outubro (21,90

MWh), indicando que o modelo foi conservador ao estimar os picos de consumo do último trimestre, possivelmente devido a fatores climáticos extremos ou dinâmicas de curto prazo não capturadas apenas pelas defasagens históricas. Em contrapartida, nos meses de março (1,31 MWh), abril (1,87 MWh) e junho (3,87 MWh), as previsões apresentaram discrepâncias mínimas, evidenciando a excelente capacidade do algoritmo em representar o comportamento médio e os períodos de transição da série.

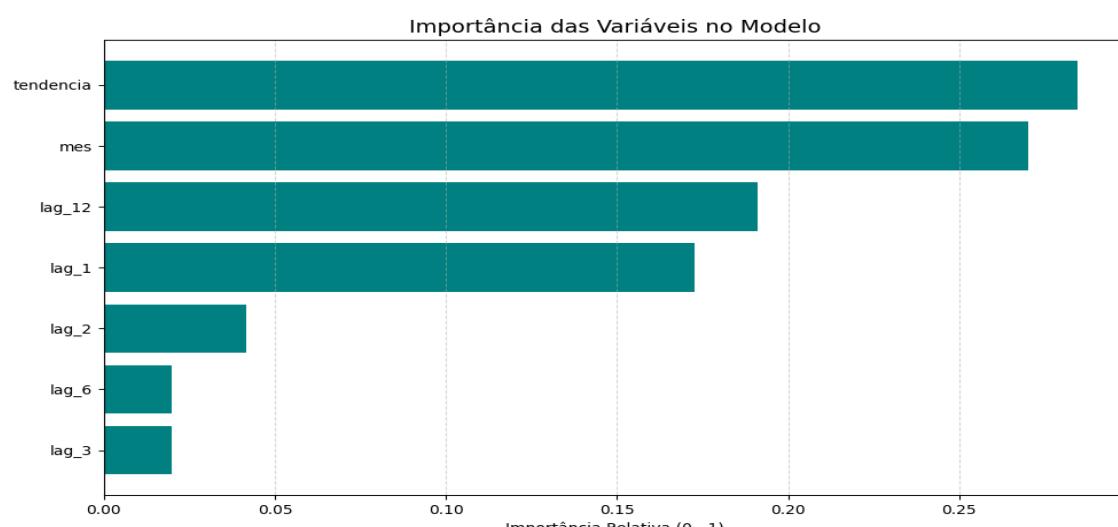
Visualmente, os gráficos comparativos corroboram esses resultados, mostrando que a curva prevista acompanha a série real com grande proximidade, especialmente no primeiro semestre. As divergências concentram-se pontualmente nos picos de demanda de fim de ano, onde o modelo apresentou um comportamento de suavização, mas sem comprometer a identificação da tendência estrutural de alta, reforçando a robustez do RF como ferramenta de planejamento.

4.3 DISCUSSÕES

A análise dos resultados indica que o modelo *Random Forest* apresenta desempenho satisfatório para previsão de consumo de energia elétrica no Piauí, alcançando um erro percentual médio (MAPE) de apenas 3,33%. A capacidade do algoritmo de lidar com relações não lineares e incorporar variáveis derivadas, como *lags* e sazonalidade, permitiu reproduzir fielmente grande parte da série histórica. 2607

Um avanço significativo nesta análise foi a quantificação dos fatores que mais influenciam a demanda, conforme apresentado na Figura 03.

Figura 03 - Importância Relativa das Variáveis no Modelo *Random Forest*.



Fonte: Autoria própria (2025)

A análise de importância das variáveis revelou que a "Tendência" (peso de 0,28) e o "Mês" (peso de 0,27) são os atributos mais determinantes. Isso confirma estatisticamente que o consumo no Piauí é regido primordialmente por um crescimento vegetativo estrutural e por uma forte sazonalidade climática, sendo secundado pelo histórico de consumo do mesmo período no ano anterior (*lag_12*).

Apesar da alta precisão global, os maiores erros concentraram-se em períodos com aumentos acentuados de consumo, principalmente em outubro e novembro. Essa subestimação dos picos sugere que a inclusão de variáveis exógenas, como temperatura média, ondas de calor ou indicadores econômicos, poderia refinar a sensibilidade do modelo a eventos extremos que fogem ao padrão histórico médio.

Do ponto de vista prático, os resultados fornecem suporte valioso para o planejamento energético e a operação do sistema, permitindo antecipar a demanda, otimizar a alocação de recursos e subsidiar decisões estratégicas. O estudo demonstra que, com um conjunto bem estruturado de *features* temporais e o uso de *Random Forest*, é possível obter previsões confiáveis para sistemas elétricos regionais com forte sazonalidade mensal, como o Piauí.

5. CONCLUSÃO

2608

O presente estudo teve como objetivo desenvolver um modelo preditivo capaz de estimar o consumo de energia elétrica no estado do Piauí, utilizando técnicas de *Machine Learning* aplicadas a séries temporais mensais. O método adotado, baseado no algoritmo *Random Forest*, mostrou-se eficaz na captura das tendências de longo prazo e das flutuações sazonais presentes na série histórica. Os resultados obtidos para o ano de 2024 indicam uma precisão elevada, evidenciada pelas métricas de desempenho finais: MAE de 13,49 MWh, RMSE de 16,31 MWh e um erro percentual (MAPE) de 3,33%. A análise de importância das variáveis corroborou a hipótese de que o consumo no estado é regido primordialmente por uma tendência de crescimento estrutural e por forte sazonalidade mensal.

Além disso, a abordagem utilizada evidencia o potencial de aplicação do modelo no setor elétrico estadual, permitindo às concessionárias e planejadores energéticos antecipar demandas, otimizar alocação de recursos e tomar decisões estratégicas com maior confiabilidade. A utilização de variáveis derivadas da própria série, como defasagens e componentes sazonais, mostrou-se fundamental para o sucesso do modelo, destacando a importância da engenharia de atributos em tarefas de previsão energética.

Para trabalhos futuros, recomenda-se explorar técnicas avançadas que possam capturar ainda mais nuances da série temporal e possíveis padrões complexos não lineares. Entre as alternativas sugeridas estão redes neurais recorrentes, como LSTM, capazes de lidar com dependências temporais mais longas; algoritmos de *boosting*, como XGBoost, que podem melhorar a performance em cenários com alta variabilidade; análises utilizando séries de maior granularidade, como dados diários ou horários; e a investigação de modelos híbridos, que combinem abordagens estatísticas e de machine learning para potencializar a acurácia das previsões.

Em suma, o estudo comprova que técnicas de *Machine Learning*, quando aplicadas de maneira estruturada e acompanhadas de pré-processamento adequado, são ferramentas robustas e confiáveis para previsão de demanda elétrica, oferecendo suporte estratégico para a gestão energética e planejamento operacional no estado do Piauí.

REFERÊNCIAS

AMASYALI, Kadir; EL-GOHARY, Nora M. A review of data-driven building energy consumption prediction studies. *Renewable and Sustainable Energy Reviews*, v. 81, n. 1, p. 1192–1205, jan. 2018. Disponível em:

<https://www.sciencedirect.com/science/article/abs/pii/S1364032117306093>. Acesso em: 30 set. 2025. 2609

ANEEL – AGÊNCIA NACIONAL DE ENERGIA ELÉTRICA. Relatório Anual 2022. Brasília: ANEEL, 2022. Disponível em: <https://www.aneel.gov.br>. Acesso em: 20 set. 2025.

ARMSTRONG, J. S. *Principles of Forecasting: A Handbook for Researchers and Practitioners*. Boston: Kluwer Academic, 2001. Disponível em: <https://gwmn.net/doc/statistics/prediction/2001-armstrong-principlesforecasting.pdf>. Acesso em: 30 set. 2025.

BOX, G. E. P.; JENKINS, G. M.; REINSEL, G. C.; LJUNG, G. M. *Time Series Analysis: Forecasting and Control*. 5. ed. Hoboken: John Wiley & Sons, 2015. 720 p.

BREIMAN, Leo. Random Forests. *Machine Learning*, v. 45, n. 1, p. 5–32, 2001. Disponível em: <https://mfm.uchicago.edu/wp-content/uploads/2020/06/Breiman-Random-Forests.pdf>. Acesso em: 28 set. 2025.

BUNN, D. W.; FARMER, E. D. *Comparative Models for Electrical Load Forecasting*. New York: John Wiley & Sons, 1985.

DI PIAZZA, A.; DI SANTO, S.; LO BELLO, L.; LUCIANO, C.; RANDO, F.; SFERLAZZA, A. An artificial neural network-based forecasting model of energy-related time series for electrical grid management. *Mathematics and Computers in Simulation*, v. 177, p. 369–383, 2020. Disponível em:

<https://www.sciencedirect.com/science/article/abs/pii/S0378475420301695>. Acesso em: 28 set. 2025.

EPE – EMPRESA DE PESQUISA ENERGÉTICA. Balanço Energético Nacional (BEN) 2024. Rio de Janeiro: EPE, 2024. Disponível em: <https://www.epe.gov.br/pt/publicacoes-dados-abertos/publicacoes/balanco-energetico-nacional-2024>. Acesso em: 23 set. 2025.

EPE – EMPRESA DE PESQUISA ENERGÉTICA. Anuário Estatístico de Energia Elétrica 2023. Rio de Janeiro: EPE, 2023. Disponível em: <https://www.epe.gov.br/pt/publicacoes-dados-abertos/publicacoes/anuario-estatistico-de-energia-eletrica>. Acesso em: 20 set. 2025.

EPE – EMPRESA DE PESQUISA ENERGÉTICA. Consumo de Energia Elétrica. Rio de Janeiro, 2025. Disponível em: <https://www.epe.gov.br/pt/publicacoes-dados-abertos/publicacoes/consumo-de-energia-eletrica>. Acesso em: 20 set. 2025.

GÉRON, Aurélien. Hands-On Machine Learning with Scikit-Learn, Keras & TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems. 2. ed. Sebastopol, CA: O'Reilly Media, 2019.

GIL, Antonio Carlos. Métodos e Técnicas de Pesquisa Social. 6. ed. São Paulo: Atlas, 2008.

GOODFELLOW, Ian; BENGIO, Yoshua; COURVILLE, Aaron. Deep Learning. Cambridge: MIT Press, 2016. Disponível em: <https://www.deeplearningbook.org>. Acesso em: 23 set. 2025.

HYNDMAN, Rob J.; ATHANASOPOULOS, George. Forecasting: Principles and Practice. 2. ed. Melbourne: OTexts, 2018. Disponível em: <https://otexts.com/fpp2/index.html>. Acesso em: 30 set. 2025. 2610

KAYTEZ, F.; TAPLAMACIOĞLU, M. C.; ÇAM, E.; HARDALAC, F. Forecasting electricity consumption: A comparison of regression analysis, neural networks and least squares support vector machines. International Journal of Electrical Power & Energy Systems, v. 67, p. 431–438, 2015. Disponível em: https://www.researchgate.net/publication/270006584_Forecasting_electricity_consumption_A_comparison_of_regression_analysis_neural_networks_and_least_squares_support_vector_machines. Acesso em: 20 set. 2025.

MAKRIDAKIS, Spyros; SPILLOTIS, Evangelos; ASSIMAKOPOULOS, Vassilios. The M4 Competition: 100,000 time series and 61 forecasting methods. International Journal of Forecasting, v. 36, n. 1, p. 54–74, 2020. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0169207019301128>. Acesso em: 15 out. 2025.

ONS – OPERADOR NACIONAL DO SISTEMA ELÉTRICO. Relatório Anual 2022. Brasília: ONS, 2022. Disponível em: https://www.ons.org.br/relatorio_anual/index2022.html. Acesso em: 20 set. 2025.

ONS – OPERADOR NACIONAL DO SISTEMA ELÉTRICO. Nota Técnica DPL 0008/2023: Previsão de carga para o Planejamento Anual da Operação Energética (PLAN) 2023–2027. Rio de Janeiro: ONS, 2023. Disponível em:

https://www.ons.org.br/AcervoDigitalDocumentosEPublicacoes/NT_PLAN_2023_2027_EPE_ONS_CCEE_Jan2023.pdf. Acesso em: 23 set. 2025.

SUGANTHI, L.; SAMUEL, A. A. Energy models for demand forecasting—A review. *Renewable and Sustainable Energy Reviews*, v. 16, n. 2, p. 1223–1240, 2012. Disponível em: <https://www.sciencedirect.com/science/article/abs/pii/S1364032111004242>. Acesso em: 28 set. 2025.

WERON, Rafal. Electricity price forecasting: A review of the state-of-the-art with a look into the future. *International Journal of Forecasting*, v. 30, n. 4, p. 1030–1081, out. 2014. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0169207014001083>. Acesso em: 28 set. 2025.

WILLMOTT, Cort J.; MATSUURA, Kenji. Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance. *Climate Research*, v. 30, n. 1, p. 79–82, 2005. Disponível em: https://www.int-res.com/articles/cr_2005/30/c03op079.pdf. Acesso em: 15 out. 2025.

ZHANG, G. Peter; PATUWO, B. Eddy; HU, Michael Y. Forecasting with artificial neural networks: The state of the art. *International Journal of Forecasting*, v. 14, n. 1, p. 35–62, mar. 1998. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0169207097000447>. Acesso em: 30 set. 2025.