

Previsão do consumo de energia elétrica na região Sudeste: aplicação de modelos ARIMA e LSTM

Iago Gomes Gonçalves¹ 

¹ Doutorando em Economia Aplicada pela Universidade Federal de Viçosa. E-mail: iago.g.goncalves@ufv.br

RESUMO

O objetivo deste estudo é prever o consumo de energia elétrica na região Sudeste do Brasil, de abril de 2023 a março de 2024, utilizando modelos ARIMA e redes neurais LSTM. Utilizando dados mensais de 2002 a 2023, a pesquisa compara os modelos com base nas métricas de erro RMSE, EAM e MAPE. O modelo ARIMA captura padrões sazonais e lineares no curto prazo, enquanto o modelo LSTM se destaca na previsão de tendências não lineares e de longo prazo. A combinação das duas abordagens se mostrou promissora para aumentar a precisão das previsões, indicando que os formuladores de políticas podem criar expectativas razoáveis quanto às projeções futuras. Esta pesquisa contribui metodologicamente ao explorar abordagens complementares e, em termos práticos, contribui para um planejamento energético eficiente, baseado em previsões de curto prazo mais assertivas, que permitem a operação segura do sistema elétrico, reduzindo o risco de sobrecargas e interrupções no fornecimento de energia.

PALAVRAS-CHAVE

Economia da energia, ARIMA, LSTM

Forecasting electricity consumption in the southeast region: Application of ARIMA and LSTM models

ABSTRACT

The objective of this study is to forecast electricity consumption in the Southeast region of Brazil from April 2023 to March 2024 using ARIMA models and LSTM neural networks. Using monthly data from 2002 to 2023, the research compares the models based on the error metrics RMSE, EAM, and MAPE. The ARIMA model captures seasonal and linear patterns in the short term, while the LSTM model excels in predicting nonlinear and long-term trends. The combination of the two approaches has shown promise in improving forecasting accuracy, suggesting that policymakers can have reasonable expectations about future projections. This research contributes methodologically by exploring complementary approaches, and a practical contribution to efficient energy planning, based on more assertive short-term forecasts, which allow for the safe operation of the electricity system, reducing the risk of overloads and interruptions in energy supply.

KEYWORDS

Energy economics, ARIMA, LSTM

CLASSIFICAÇÃO JEL

C22, O13, Q47

1. Introdução

A demanda por energia é um fenômeno atrelado às necessidades básicas da humanidade. Essa conexão destaca a importância da energia no apoio às atividades humanas e à sobrevivência ao longo da história (Barkhordar e Saboohi, 2014). IEA (2024) aponta que a eletricidade é considerada um elemento-chave do desenvolvimento moderno, impactando vários aspectos da vida, do crescimento econômico ao bem-estar social; de modo que, conforme Adeosun et al. (2023), o acesso confiável e sustentável à eletricidade se tornou um indicativo importante para o desenvolvimento socioeconômico dos países. Além disso, o crescimento econômico, aliado ao avanço da urbanização, continuará impulsionando a demanda energética, exigindo sistemas modernos e eficientes de fornecimento (Sambodo et al., 2021). Assim, a energia se consolida como um elemento indissociável da atividade econômica.

Para Ritchie et al. (2024), o aumento na demanda energética reflete a dinâmica das economias emergentes, que impulsionaram esse aumento nas últimas décadas, sobretudo China, Índia e Brasil. Isso ocorre, em grande medida, devido à expectativa de um maior crescimento econômico dessas economias (Ahmad e Zhang, 2020). IEA (2024) denota que, atualmente, os mercados emergentes e em desenvolvimento respondem por cerca de 80% do crescimento global de procura por eletricidade, enquanto IEA (2023) estima que, até 2026, cerca de 85% da eletricidade adicional proverá da China, da Índia e do Sudeste Asiático, em que a recuperação industrial, eletrificação contínua dos setores residencial e o setor de transportes serão os principais responsáveis por esse aumento.

Billewar et al. (2023) apontam que, até 2035, o consumo de energia nos países da Organização para a Cooperação e Desenvolvimento Econômico (OCDE) crescerá aproximadamente 14%, enquanto, nos países não pertencentes à OCDE, esse aumento será de 84%. Além disso, os autores destacam que China e Índia apresentarão a primeira e a segunda maior demanda de energia do mundo entre os países em desenvolvimento, de modo que esses dois países responderão por metade do crescimento global na demanda por energia nos próximos 15 anos. Por sua vez, os Estados Unidos permanecerão como o segundo maior consumidor global de eletricidade, com um crescimento significativo da participação de energia solar e eólica na matriz elétrica (IEA, 2023).

Rahman (2020) enfatiza que todos os setores de um sistema econômico consomem energia e sua demanda é substancialmente significativa, superando a demanda global de energia, evidenciando que o crescimento econômico de uma nação está intrinsecamente relacionado ao consumo de eletricidade. Sachs (2015) se vale do conceito de heterogeneidade estrutural para descrever as variações produtivas, ocupacionais e regionais presentes nas economias. Nesse aspecto, as políticas públicas precisam considerar as diversas heterogeneidades existentes, sejam elas geográficas, climáticas ou sociais, à medida que as políticas implementadas ao longo do tempo moldam a

trajetória de desenvolvimento de uma região, o que pode resultar em melhorias ou retrocessos. Por essa razão, é um desafio para o Brasil formular políticas que atendam a toda a população.

A Agenda 2030 para o Desenvolvimento Sustentável juntamente com o Acordo de Paris constituem uma resposta multilateral dos países aos crescentes desafios ambientais, econômicos e sociais que comprometem o bem-estar das futuras gerações. Entre os Objetivos de Desenvolvimento Sustentável (ODS), para fins deste estudo, destaca-se o ODS 7, que visa garantir o acesso à energia acessível, confiável, sustentável e moderna para todos (UN, 2016). No caso brasileiro, o Programa Luz Para Todos, instituído em 2003, foi desenhado para democratizar e viabilizar o acesso e o uso da energia elétrica pela população ainda não atendida (MME, 2023). Vinte anos após sua implementação, a energia elétrica fornecida pela rede geral ou fonte alternativa alcançou 99,8% dos domicílios do país (IBGE, 2023).

Após alcançar grande parte dos domicílios brasileiros, o próximo desafio em termos energéticos consiste na garantia de uma matriz energética sustentável e ambientalmente responsável. Atualmente, cerca de 49% da energia ofertada advém de fontes renováveis, o que posiciona a matriz energética brasileira entre as mais limpas do mundo (EPE, 2024a). Os dados também apontam que o Brasil possui uma matriz elétrica majoritariamente renovável, com a fonte hídrica representando 59% da oferta interna de energia elétrica, e que 89% da eletricidade gerada no país provém de fontes renováveis.

De acordo com EPE (2024a), os setores industrial e residencial se destacam como os maiores consumidores de energia elétrica no Brasil, representando 36,4% e 27,5% do consumo total de eletricidade, respectivamente. Cabreira et al. (2024) mostram a trajetória do consumo total de eletricidade no Brasil, destacando o peso exercido pelo setor industrial na região Sudeste em relação ao consumo total do país, respondendo por cerca de 50% do consumo total de eletricidade. Freitas et al. (2019) destacam que a preocupação do Brasil quanto à política de segurança energética está relacionada ao crescimento econômico nas últimas duas décadas que, conseqüentemente, levou a um aumento no consumo de energia primária do país em mais de um terço.

Os autores apontam que as políticas energéticas brasileiras estabelecidas nesse período foram capazes de garantir segurança energética ao país, frente aos principais desafios ambientais emergentes. Por sua vez, as mudanças climáticas têm desencadeado fenômenos extremos, intensificando a demanda por eletricidade, sobretudo em áreas densamente povoadas, como a região Sudeste do Brasil (Zuin et al., 2022; Simões et al., 2025).

O aumento das temperaturas gera uma demanda crescente por sistemas de refrigeração e ar-condicionado, levando a picos inesperados no consumo de energia elétrica durante períodos de calor elevado. De acordo com EPE (2024b), as ondas de calor têm impulsionado o consumo de eletricidade no Brasil. Além disso, ao projetar a trajetória

futura da demanda brasileira de energia, é preciso considerar as incertezas a que o setor está sujeito (EPE, 2020). Isso ressalta a necessidade de expansão do sistema energético para atender à demanda crescente por eletricidade, com implementação de políticas e investimentos adequados para garantir o suprimento de energia no longo prazo. Paralelo a isso, Silva (2022) destaca que o desenvolvimento infraestrutural das redes hidrelétricas brasileiras não seguiu o crescimento populacional, o que gerou um desencontro entre os níveis de oferta e demanda de eletricidade.

Tendo em vista que os cenários de consumo de energia elétrica para as próximas décadas indicam tendências de aumento e, ao mesmo tempo, sugerem incertezas que não permitem estabelecer definitivamente os padrões de consumo pela qual essas mudanças se darão, a questão energética se mantém relevante (Gonçalves, 2022). A região Sudeste do Brasil, apesar da universalização do acesso à eletricidade, segue como a maior consumidora de energia elétrica do país, exigindo um planejamento energético alinhado ao crescimento populacional e à demanda por serviços (EPE, 2020).

Amaral et al. (2022) ressaltam que prever a demanda de energia é importante para garantir o bom funcionamento do sistema, à medida que se observam mudanças no sistema elétrico. Desse modo, modelos de previsão permitem identificar padrões nos dados observados, tendências e sazonalidades, além da comparação entre diferentes abordagens, visando aprimorar a precisão das projeções futuras. Em termos práticos, essas previsões são utilizadas para melhorar o planejamento de geração e distribuição de energia, principalmente diante das mudanças nos padrões de consumo ocasionadas por fatores climáticos e econômicos.

Billewar et al. (2023) afirmam que diversos órgãos internacionais analisam as tendências mundiais na adoção de energia elétrica, para avaliar quanto capital será necessário para atender a essa demanda no futuro. Contudo, ainda há necessidade de explorar soluções alternativas e eficazes para os problemas existentes. Nesse contexto, os resultados obtidos por meio dos modelos de previsão são ferramentas estratégicas para antecipar flutuações na demanda, permitindo o planejamento eficiente das cargas nos sistemas de energia das regiões e garantindo maior confiabilidade no abastecimento.

Portanto, compreender os padrões de consumo de eletricidade e antecipar as tendências futuras são passos necessários para o planejamento e a gestão adequados dos recursos energéticos, sobretudo eletricidade. Isso se deve ao fato de que a oferta de energia elétrica deve ocorrer de forma praticamente simultânea à sua demanda, exigindo uma coordenação precisa entre produção, transmissão e distribuição (Velasquez et al., 2022). Desse modo, a fim de contribuir com novas estimativas sobre modelos de previsão, o objetivo principal deste trabalho é buscar um modelo que melhor se ajusta aos dados e projetar o consumo de energia elétrica na região Sudeste por 12 meses à frente do final da amostra selecionada, de abril de 2023 a março de 2024, utilizando modelos de séries temporais, do tipo *Autoregressive Integrated Moving Average* (ARIMA), e redes neurais recorrentes, do tipo *Long Short-Term Memory*

(LSTM).

Embora existam estudos que realizem análise semelhante, o presente estudo avança em alguns pontos. O primeiro ponto diz respeito à discussão prévia que ressalta a importância de se investigar a demanda por eletricidade, neste caso para a região Sudeste, e não apenas à aplicação de métodos econométricos pela sua ampla utilização e facilidade de implementação, sem a devida contextualização teórica ou motivação empírica. Portanto, a ideia central é demonstrar que a previsão não é um fim em si mesma, mas uma ferramenta para entender e gerenciar o suprimento de energia elétrica na região de interesse.

Em segundo lugar, ao contrário de estudos anteriores, como o de Santos e Chaukoski (2020), que utilizaram dados até maio de 2020, o presente trabalho expande a base de dados, abrangendo o período de janeiro de 2002 a março de 2023 e realiza previsões para um horizonte de 12 meses à frente. Metodologicamente, há avanços na aplicação tanto do modelo ARIMA quanto do LSTM. Santos e Chaukoski (2020) não discutem as etapas subsequentes para estimar o modelo, recorrendo a *softwares* que selecionam modelos automaticamente com base no critério de informação de Akaike (AIC) e apenas mencionam a aplicação do LSTM, ao contrário do presente estudo, que o faz. Ademais, este estudo explora diferentes configurações do modelo LSTM, comparando abordagens com e sem reposição do estado interno da rede ao longo das previsões.

Hyndman e Khandakar (2008) argumentam que o maior espaço de seleção dos modelos ARIMA pode, na verdade, prejudicar o desempenho da previsão por introduzir incerteza adicional, sugerindo que confiar unicamente na previsão automática utilizando critérios de informação como o AIC, em um grande conjunto de modelos, pode não ser a abordagem mais eficaz. Nesse sentido, existem outras maneiras de seleção de ordens, utilizadas neste estudo, como a Função de Autocorrelação (FAC) e a Função de Autocorrelação Parcial (FACP), que não devem ser usadas isoladamente, mas em conjunto com instrumentos adicionais (Svetunkov, 2023) e com a abordagem descrita por Lütkepohl e Krätzig (2004), que consiste em ajustar modelos de ordem baixa para encontrar a especificação mais adequada.

A decisão de utilizar o modelo ARIMA, mesmo diante da sazonalidade, foi embasada nos resultados das estimativas do teste Box-Pierce, os quais indicaram que a diferenciação sazonal levaria a resíduos autocorrelacionados. Esse ponto se alinha com Svetunkov (2022), à medida que, independentemente da hipótese formulada ou da abordagem adotada, é preciso ter uma compreensão prévia do problema. Nesse sentido, mesmo que a sazonalidade seja observável a partir da visualização sequencial dos dados e confirmada por algum teste empírico, não é trivial concluir que se as diferenças sazonais são necessárias para otimizar os resultados (Svetunkov, 2023). Assim, na presença de resíduos autocorrelacionados, o modelo pode produzir previsões imprecisas e intervalos de previsão com largura incorreta (Svetunkov, 2022).

Portanto, essa escolha metodológica não é arbitrária, mas uma resposta aos possíveis resultados imprecisos e aos erros na previsão da série temporal analisada, visando obter melhor ajuste e, conseqüentemente, previsões mais confiáveis para o problema em questão. Dessa forma, uma vez realizadas as projeções futuras, torna-se possível comparar a capacidade preditiva de cada um dos modelos estimados através das estatísticas de erros de previsão.

Este estudo contribui para a literatura ao detalhar as metodologias de previsão e ao analisar as implicações de projetar o consumo de energia elétrica no curto prazo, em um horizonte de 12 meses à frente na região Sudeste, com reflexos diretos em políticas públicas e no planejamento estratégico do setor elétrico. Para o governo, estimativas precisas da demanda orientam investimentos em geração, transmissão e distribuição, garantindo a eficiência e a segurança do sistema durante variações sazonais. Para as empresas, essas previsões possibilitam uma alocação mais eficaz de recursos, um planejamento operacional aprimorado e decisões estratégicas assertivas, assegurando a estabilidade da capacidade instalada e o atendimento adequado à demanda. Isso é particularmente relevante em um contexto climático global, em que a eficiência no uso da energia e a mitigação de impactos ambientais são prioridades para o desenvolvimento econômico sustentável, além da busca recorrente por atingir os ODS estabelecidos na Agenda 2030 para o Desenvolvimento Sustentável.

Além desta introdução, o artigo está organizado em mais quatro seções. Na seção dois, destacam-se os aspectos teóricos referentes aos modelos de previsão do consumo de energia elétrica, bem como uma discussão sobre a literatura do tema e seu respectivo estado atual. Na seção seguinte, estão detalhados os dados utilizados e suas respectivas fontes, além da descrição detalhada dos modelos de previsão utilizados no trabalho. A seção quatro contém a apresentação e a discussão dos resultados e, por último, são realizadas as considerações finais do estudo.

2. Revisão de literatura

Realizar previsões de demanda envolve identificar padrões em séries históricas para antecipar comportamentos futuros ou, alternativamente, analisar fatores causais que influenciam o comportamento e projetar suas tendências. Nas últimas décadas, surgiram inúmeras técnicas capazes de realizar previsões de dados, incluindo previsão da carga demandada de energia elétrica. Tais técnicas variam tanto em termos metodológicos quanto em horizonte de previsão, de curto, médio ou longo prazo. Nesse aspecto, Rahman e Hazim (1993) dividem as técnicas de previsão de curto prazo em duas: técnicas estatísticas e técnicas baseadas em sistemas inteligentes.

A técnica estatística mais adotada para previsões se refere aos modelos ARIMA, de forma seminal abordado por Box e Jenkins (1976), que é uma abordagem clássica e reconhecida pela capacidade de lidar com padrões lineares e dados estacionários, sendo amplamente utilizada na previsão estatística. Por sua vez, dentre as técnicas

baseadas em sistemas inteligentes, inclui-se a lógica Fuzzy e Redes Neurais Artificiais (RNA), a qual é capaz de lidar com padrões de dados complexos e não lineares, que modelos lineares tradicionais como o ARIMA podem ignorar, o que pode levar, de certo modo, a uma previsão aprimorada das RNA em ambientes não estacionários e dinâmicos. Ainda, conforme Zhang (2003), a combinação dos dois modelos pode ser uma forma eficaz de melhorar a precisão da previsão alcançada, em comparação ao uso de qualquer um dos modelos individualmente.

Alsheheri (2025) ressalta que, embora sejam duas abordagens distintas, elas são, ao mesmo tempo, complementares, pois enquanto o ARIMA oferece transparência e facilidade interpretativa, uma RNA demonstra flexibilidade e desempenho superior na captura de padrões não lineares. Por esse motivo, há pelo menos duas décadas, é comum encontrar na literatura estudos que realizam a comparação das técnicas de previsão para determinar qual o modelo ou técnica é ideal para o tipo de dado escolhido ou objetivo específico da previsão (Zhang, 2003; Kardakos et al., 2013; Barros, 2014; Pierre et al., 2023; Alsheheri, 2025).

Ackermann e Sellitto (2022) apontam que a análise de séries temporais é uma das maneiras mais adequadas de realizar previsões de curto prazo. Dessa maneira, a utilização de modelos de previsão univariados oferece uma abordagem relevante para compreender as dinâmicas do consumo de eletricidade, à medida que tais modelos geram previsões precisas no curto prazo, mesmo diante de fatores que possam impactar o consumo recorrente de energia elétrica. Kaur et al. (2023) denotam que ARIMA e meio ambiente estão altamente correlacionados, pois a previsão é necessária em todas as nuances do meio ambiente, como poluição atmosférica e sonora, combustíveis fósseis, dados pluviométricos e águas subterrâneas. A aplicabilidade se expande dentro da área de meio ambiente, a exemplo de estudos sobre desmatamento e modelos de previsão (da Silva Costa et al., 2021; Gonçalves e Ferreira, 2023).

Por sua vez, o uso crescente da abordagem de RNA vem sendo proeminente para tarefas de aprendizado de máquina, impulsionado pela disponibilidade cada vez maior de dados. Zhang et al. (1998) apontam algumas razões que tornaram essa modelagem atraente para boas previsões, como, poucas suposições iniciais do modelo, generalizações para dados não observados e capacidade de modelar dados não lineares.

Hochreiter e Schmidhuber (1997) introduziram um tipo de Rede Neural Recorrente, denominadas Redes de Memória de Longo Prazo ou LSTM, um tipo de rede neural desenvolvida para lidar com dados sequenciais, à medida que havia um desafio enfrentado pelas redes neurais tradicionais para capturar dependências temporais em grandes sequências de dados. Uma característica das LSTM é a sua estrutura de células de memória, que atuam como unidades de armazenamento de informações. Hewamalage et al. (2021) destacam a natureza competitiva dos modelos de redes neurais recorrentes em vários cenários de previsão, enfatizando seu potencial como alternativa aos métodos estatísticos tradicionais, como suavização exponencial e ARIMA.

A literatura de séries temporais tem realizado previsões dos modelos ARIMA e LSTM para fins de comparação e robustez. Além do mais, a previsão do consumo de energia elétrica utilizando modelos ARIMA é objeto de estudo há pelo menos uma década (Campos, 2008; Chujai et al., 2013; Pires, 2016; de Oliveira e Oliveira, 2018). Contudo, com o avanço dos algoritmos de previsão, os modelos LSTM ganharam espaço, de modo que a literatura sobre previsão energética se ampliou, a exemplo de Vassalli (2018); Sachin et al. (2020); Pooniwala e Sutar (2021); Pierre et al. (2023).

Os estudos sobre previsões podem ser encontrados tanto em termos agregados quanto considerando diferentes regiões. Amaro et al. (2017), por exemplo, estimam a capacidade dos modelos ARIMA em prever a demanda de energia elétrica no Brasil de maio de 2011 a junho de 2015. Os resultados sugerem que o modelo ARIMA foi capaz de se ajustar bem aos dados da amostra. Por sua vez, Marcos e Júnior (2021) estimam o consumo futuro de energia elétrica da região Nordeste entre os anos de 2004 e 2019. Da mesma maneira, Belém (2024) projeta a demanda de energia elétrica na região Nordeste do Brasil, identificando que a curva de previsão se ajusta aos valores reais de forma satisfatória. Tidre et al. (2013) utilizam dados de consumo de eletricidade da região Norte de 1979 a 2009 e realizam a previsão 12 meses à frente, de janeiro a dezembro de 2010, de modo que os resultados apresentam bom ajuste ao modelo SARIMA.

Para a região Sul, Nunes (2019) avalia o consumo de eletricidade dos setores residencial, comercial e industrial do Rio Grande do Sul. Os resultados mostram que o modelo ARIMA se ajustou melhor aos dados dos setores comercial e industrial, enquanto para o setor residencial o modelo VAR se ajustou melhor. Santos (2019) projeta a demanda de médio prazo por energia do setor industrial nas regiões Sul e Sudeste do Brasil, a partir de um modelo ETS e ARIMA. Embora ambos os modelos tenham apresentado resultados satisfatórios, o modelo ARIMA obteve ajuste inferior ao modelo ETS.

Matos (2022) estima os modelos de previsão de Holt-Winters e Box e Jenkins de janeiro a dezembro de 2020, valendo-se de dados entre janeiro de 2011 a dezembro de 2019, para as regiões Centro-Oeste e Sudeste. Os resultados encontrados sugerem que o modelo SARIMA é o modelo mais adequado para a série de demanda energética. Já Feitosa (2023) compara quatro modelos preditivos (Holt-Winters, SARIMA, SARIMAX e Prophet) para o consumo de energia elétrica do Brasil, do subsistema Sudeste/Centro-Oeste e do estado de São Paulo. Os resultados mostram que o modelo Holt-Winters se adequou melhor aos dados.

Por sua vez, Santos e Chaucoski (2020) utilizam os modelos SARIMA e LSTM para prever o consumo de energia elétrica para a região Sudeste. Os autores se valem de dados de 2001 a 2020 e realizam a projeção de seis meses após o fim da amostra. Silva (2022) aplica os métodos de suavização exponencial e ARIMA aos dados de consumo de energia elétrica também para a região Sudeste e realiza previsões de janeiro de 2022 a dezembro de 2025. Os dois métodos aplicados cumpriram o propósito de realizar

previsões razoáveis no curto prazo, mas o modelo ARIMA apresentou desempenho ligeiramente inferior ao método de suavização exponencial.

Embora os resultados da maioria dos estudos discutidos nesta seção tenham demonstrado boa capacidade de generalização no curto prazo, eles possuem uma característica em comum, não há uma discussão metodológica detalhada das estimativas, ou seja, apresentam descrições superficiais dos procedimentos metodológicos empregados e a mesma situação ocorre para a discussão das implicações políticas dos resultados. Por exemplo, é comum que a seleção dos modelos ocorra por meio de *softwares* que aplicam, automaticamente, critérios de informação (como o AIC), sem que sejam explicitadas as etapas de diagnóstico, a validação dos parâmetros e a análise dos resíduos.

Mesmo que funções automatizadas sejam úteis e ofereçam praticidade, qualquer processo automático pode implicar em riscos, de modo que vale a pena entender o comportamento dos modelos, mesmo quando algum procedimento automático indique o modelo ideal (Hyndman e Athanasopoulos, 2025). Essa abordagem automatizada tende a negligenciar alguns aspectos que podem influenciar a robustez das previsões, como a verificação das condições de estacionariedade e a adequação dos modelos aos padrões específicos da série de consumo de energia elétrica.

Nesse aspecto, este trabalho se diferencia ao detalhar as etapas de estimação do modelo ARIMA e ao explorar diferentes configurações do modelo LSTM, comparando abordagens com e sem reposição do estado interno da rede neural ao longo das previsões. Enquanto Santos e Chaucoski (2020) apenas mencionam a implementação do modelo LSTM, este estudo testa diferentes combinações de neurônios (15, 30, 50 e 100) e épocas de treinamento (100, 200, 300 e 500), totalizando 32 variações de modelos LSTM testadas.

Portanto, todas as fases do processo de modelagem são descritas de forma clara e sistemática, desde o pré-processamento dos dados até a avaliação e validação dos modelos, tornando a metodologia mais transparente e passível de reprodução. Diante disso, o presente estudo busca superar as limitações observadas em outros estudos ao oferecer uma abordagem metodológica detalhada para as etapas de estimação, contribuindo para o avanço da literatura sobre previsão do consumo de energia elétrica, assim como para o entendimento das dinâmicas do setor energético em uma região que concentra grande parte da produção industrial do país.

3. Metodologia

3.1 Fonte de dados

O conjunto de dados utilizados neste estudo se refere à série temporal do consumo de energia elétrica mensal (em GWh) da região Sudeste. Os dados são disponibilizados

pelo Sistema Gerenciador de Séries Temporais do Banco Central do Brasil ¹. O período de análise compreende janeiro de 2002 a março de 2023, totalizando 255 observações. A fim de comparar a adequação dos modelos e verificar a qualidade de ajuste das previsões realizadas, 12 observações foram deixadas de fora da amostra, de abril de 2023 a março de 2024. A finalidade dessa separação é comparar os valores previstos pelo modelo com os valores deixados de fora da amostra.

3.2 Estacionariedade de uma série temporal

A metodologia Box-Jenkins é uma importante ferramenta de séries temporais, uma vez que possui ampla aplicabilidade em séries univariadas (Box e Jenkins, 1976). Essa metodologia explora a relação intrínseca dos dados ao longo do tempo e a aplicabilidade dessa modelagem se divide em quatro etapas: identificação, estimação, verificação e previsão (Bueno, 2018; Morettin e Tolo, 2022). A etapa de identificação busca verificar qual modelo melhor descreve o comportamento da série.

Nesse contexto, a análise gráfica, embora seja um procedimento informal, fornece indícios sobre a estacionariedade dos dados, se há presença ou não de tendência, sazonalidade e possíveis quebras estruturais (Enders, 2015). Além do mais, a etapa de identificação passa pela análise da Função de Autocorrelação (FAC) e da Função de Autocorrelação Parcial (FACP) para identificar séries não estacionárias e modelos candidatos a gerador da série.

Hyndman e Athanasopoulos (2025) denotam que a etapa de estimação consiste na obtenção dos parâmetros do modelo identificado estimados por máxima verossimilhança. Na etapa de verificação, avalia-se a adequação do modelo estimado através da análise dos resíduos, verificando se eles se comportam como ruído branco. Por fim, a etapa de previsão utiliza o modelo validado para projetar valores futuros da série, considerando que os intervalos de previsão para modelos ARIMA são baseados nas condições de que os resíduos não são autocorrelacionados e normalmente distribuídos.

Em uma série temporal estacionária, a autocorrelação parcial cairá para zero relativamente rápido, enquanto a autocorrelação parcial de dados não estacionários diminuirá lentamente. Uma abordagem mais objetiva para determinar a necessidade de diferenciação é empregar um teste de raiz unitária (Hyndman e Athanasopoulos, 2025). Para isso, são aplicados os testes Augmented Dickey-Fuller (ADF) e Philips-Perron (PP). A hipótese nula de ambos os testes considera a presença de raiz unitária, isto é, que a série é não estacionária em nível. Já a hipótese alternativa considera a ausência de raiz unitária, ou seja, que a série é estacionária em nível. Os critérios de informação de Akaike (AIC), Schwarz (SC) e Hannan-Quinn (HQ) são empregados para identificar o número ótimo de defasagens. Essa seleção considera o número de defasagens necessário para mitigar a autocorrelação dos resíduos.

¹Disponível em: <https://www3.bcb.gov.br/sgpspub/consultarvalores/telaCvsSelecionarSeries.paint>.

Além do mais, a sazonalidade é uma característica singular das séries temporais e, quando está presente nos dados, é importante compreender como é possível encontrar um modelo apropriado para os dados e como realizar inferência sobre o modelo escolhido (Choi, 2015). Um modelo popular para séries temporais sazonais é o modelo ARIMA sazonal ou SARIMA, que tem como elemento-chave a diferenciação sazonal, que é necessária quando estão presentes raízes unitárias sazonais.

Algumas séries temporais são derivadas de dados com frequência muito baixa, os quais geralmente exibem múltiplas sazonalidades. No caso da energia elétrica, em dados horários (a cada uma hora) nota-se a presença de sazonalidade, pois durante as noites espera-se uma queda no consumo de eletricidade, fazendo com que as séries apresentem padrões de crescimento e decrescimento ao longo do dia. Mesmo agregando esses valores diariamente, semanalmente ou mensalmente, esses padrões se mantêm. Nesse aspecto, ao lidar com séries temporais que exibem tais padrões, é necessário empregar modelos capazes de identificá-los com precisão. Para isso, procede-se com a aplicação do Teste de HEGY².

3.3 ARIMA e SARIMA

O modelo ARIMA é empregado para séries não estacionárias e combina três componentes principais: termos autorregressivos AR(p), que capturam a dependência linear em p defasagens, médias móveis MA(q), que modelam o efeito de choques passados em q períodos e diferenciações consecutivas de ordem d , indicadas por Δ^d , para tornar a série estacionária. Então, uma série temporal Z_t considerada um processo ARIMA (p,d,q) pode ser expressa por:

$$\phi(B)\Delta^d Z_t = \Theta(B)a_t \quad (1)$$

Em que:

- $\phi(B) = (1 - \phi_1 B - \dots - \phi_p B^p)$ é o operador autorregressivo de ordem p ;
- $\Theta(B) = (1 - \theta_1 B - \dots - \theta_q B^q)$ é o operador de médias móveis de ordem q ;
- a_t é um ruído branco com média zero e variância constante.

Por sua vez, o modelo autorregressivo integrado de médias móveis sazonal (SARIMA) é desenvolvido a partir do modelo ARIMA, aplicado a séries temporais sazonais. Além dos componentes AR(p), MA(q) e (d), os modelos SARIMA também incluem componentes autorregressivos sazonais AR(P), média móvel sazonal MA(Q) e diferenciações sazonais (D) necessárias caso a série apresente raiz unitária sazonal. Desse modo, a notação para essa classe de modelos geralmente é dada por SARIMA (p,d,q) (P,D,Q)_s e definida pela seguinte equação:

²No contexto de séries mensais, os testes de raiz unitária sazonal são abordados por Franses (1990).

$$\Phi(B^S) \phi(B) \Delta_S^D \Delta^d Z_t = \Theta(B^S) \theta(B) a_t \quad (2)$$

Além dos elementos descritos na Equação (1), tem-se:

- $\Phi(B^S) = (1 - \Phi_1 B^S - \dots - \Phi_P B^{SP})$ é o operador autorregressivo sazonal de ordem P ;
- $\Theta(B^S) = (1 - \Theta_1 B^S - \dots - \Theta_Q B^{SQ})$ é o operador de médias móveis sazonal de ordem Q ;
- $\Delta_S^D = (1 - B^S)^D$ indica o número de diferenças sazonais D .

Como os dados utilizados neste estudo são mensais, é adotada uma sazonalidade s igual a 12, correspondente ao número de meses de um ano. A exemplo da série de consumo de eletricidade utilizada neste estudo, os modelos da classe SARIMA se mostram adequados, à medida que cumprem a condição de modelar a correlação sazonal. Lütkepohl e Krätzig (2004) recomendam o ajuste de modelos de ordem baixa para encontrar o modelo mais ajustável, utilizando os critérios de informação de Akaike e Schwarz, em que a escolha se dá pela seleção do modelo que possui os menores valores.

Após selecionar o modelo adequado, o modelo passa a ter estimativas para os resíduos, representados por \hat{a}_t . Se o modelo for apropriado, espera-se que essas estimativas se aproximem dos verdadeiros resíduos a_t e, portanto, sejam não correlacionadas. Assim, ao plotar o correlograma dos resíduos do modelo, espera-se que todos os picos estejam dentro do limite de significância, sinalizando que os resíduos são do tipo ruído branco. Ainda, ao aplicar o teste Box-Pierce, é possível mostrar a adequação do modelo quanto à natureza dos resíduos, demonstrando que os resíduos não são autocorrelacionados.

Após encontrar um modelo adequado, a próxima etapa consiste na previsão. Neste estudo, a previsão do modelo é realizada 12 períodos à frente, levando em conta as doze observações que foram retiradas da amostra, a fim de comparar o desempenho do modelo com os dados observados.

Embora o modelo ARIMA seja amplamente utilizado e seja eficaz na captura de padrões sazonais e tendências em séries temporais, ele parte do pressuposto de que há relação linear entre o valor atual da série e seus valores passados. No entanto, ao avaliar o consumo de energia, podem-se encontrar padrões não lineares e interações mais complexas, que muitas vezes não são capturadas pelos modelos tradicionais. Para lidar com esse problema potencial e explorar dinâmicas temporais distintas, este estudo também aplica o modelo LSTM.

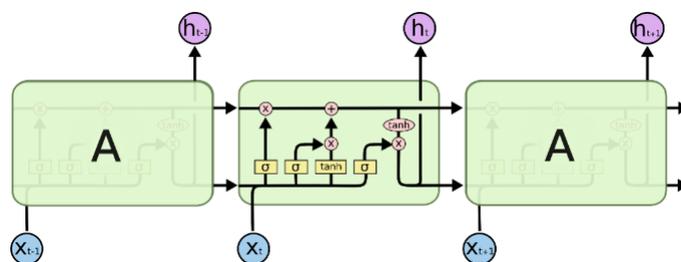
3.4 Redes Neurais Artificiais (RNA) e LSTM

Haykin (2001) define redes neurais como sistemas computacionais compostos por unidades de processamento simples, chamadas de neurônios artificiais. Essas redes são inspiradas na estrutura e no funcionamento do cérebro humano e são capazes de aprender e adaptar seu comportamento a partir de exemplos apresentados durante a fase de treinamento. Enders (2015) também denota que as Redes Neurais Artificiais (RNAs) aprendem por meio de exemplos e podem ser apropriadas para lidar com processos não lineares que possuem formas funcionais desconhecidas. Uma RNA é composta por uma camada inicial que recebe múltiplas entradas, uma ou mais camadas ocultas contendo neurônios e uma camada de saída, em que são apresentados os resultados.

Uma Rede Neural Recorrente (RNR) é um tipo de RNA utilizado para reconhecer padrões em dados sequenciais. O que diferencia as RNR's de outras redes é a forma como seus neurônios se conectam, de modo que as saídas de um neurônio podem ser usadas como entradas para neurônios em camadas subsequentes ou até mesmo nas camadas anteriores, criando um fluxo contínuo de informações. Nesse aspecto, é possível estabelecer que o modelo receberá a entrada de dados referente a 12 meses de consumo de energia elétrica e terá como saída o consumo de eletricidade do mês subsequente. Portanto, o objetivo é ajustar o modelo para prever o consumo de eletricidade de uma determinada entrada.

He et al. (2021) afirmam que o mecanismo de portões projetados para capturar tanto os estados de longo prazo da unidade inicial quanto os estados de curto prazo da unidade final é considerado um recurso particular nessa modelagem, o que tem levado ao amplo uso de LSTM em problemas relacionados à previsão de séries temporais. Por essa razão, diagramas são utilizados para representar visualmente o funcionamento de uma LSTM, uma vez que o primeiro passo é entender sua arquitetura. Olah (2015) apresenta o seguinte diagrama:

Figura 1. Arquitetura do modelo LSTM



Fonte: Extraído de Olah (2015).

Na Figura 1, cada linha transporta um vetor completo, conectando a saída de um nó às entradas de outros. Os círculos em rosa indicam operações elementares, como a soma de vetores, enquanto as caixas amarelas representam camadas de redes neurais treináveis. As linhas que se unem indicam a combinação de vetores, enquanto uma

linha que se divide indica que seu conteúdo foi duplicado e enviado para diferentes destinos.

Uma LSTM é composta por módulos repetitivos chamados células de memória, que possuem três portões principais: Portão de Esquecimento (f_t), Portão de Entrada (i_t) e Portão de Saída (o_t). Esses portões controlam o fluxo de informações ao longo da sequência de dados, permitindo que a rede aprenda quais informações devem ser mantidas ou descartadas em cada etapa temporal (Graves, 2013).

O portão do esquecimento é responsável por decidir quais informações do estado anterior (C_{t-1}) devem ser esquecidas. Ele recebe o estado oculto anterior (h_{t-1}) e a entrada atual (x_t), e então utiliza uma função de ativação sigmoide (σ) para calcular o valor de f_t , que varia entre 0 e 1. Valores próximos de 0 indicam que a informação será esquecida, e valores próximos de 1 indicam que a informação será mantida. Conforme a equação:

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \quad (2)$$

Em que W_f representa os pesos que conectam o estado oculto anterior e a entrada atual ao portão de esquecimento, e b_f é o viés. Por sua vez, o portão de entrada decide quais novas informações serão armazenadas na célula de memória. Também usa h_{t-1} e x_t para determinar quais informações devem ser atualizadas, conforme a seguinte expressão:

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \quad (3)$$

Ainda, é calculado um vetor de novos valores candidatos para a célula de memória, utilizando a função tangente hiperbólica (\tanh):

$$\tilde{C}_t = \tanh(W_C \cdot [h_{t-1}, x_t] + b_C) \quad (4)$$

Esses novos valores candidatos \tilde{C}_t são combinados com o resultado do portão de entrada para determinar o quanto de cada valor será armazenado. Com o portão de esquecimento e o portão de entrada definidos, o próximo passo é atualizar o estado da célula. Isso é feito combinando as informações que devem ser esquecidas ($f_t \cdot C_{t-1}$) e as novas informações que serão adicionadas ($i_t \cdot \tilde{C}_t$):

$$C_t = f_t \cdot C_{t-1} + i_t \cdot \tilde{C}_t \quad (5)$$

O portão de saída controla quais partes do estado da célula atual (C_t) serão utilizadas para gerar a saída. Ele também usa uma função sigmoide para calcular o_t , que

decide o quanto do estado da célula será considerado na saída:

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \quad (6)$$

A saída final da célula, h_t , é dada pela multiplicação do resultado do portão de saída (o_t) pelo estado atual da célula, que é transformado pela função tangente hiperbólica, ou seja, $h_t = o_t \cdot \tanh(C_t)$. Levando em conta a caracterização do modelo descrito acima, duas técnicas diferentes para a previsão serão consideradas, LSTM com reposição e sem reposição. Além disso, 50 observações mais recentes da série temporal são reservadas para o conjunto de treino, o que equivale a cerca de 20% dos dados.

No modelo LSTM sem reposição, o estado interno da rede neural é mantido ao longo das previsões. Isso permite que o modelo armazene informações sobre os períodos anteriores, o que é fundamental em séries temporais. A previsão é realizada de uma só vez para os 12 meses destinados para a previsão (abril de 2023 a março de 2024), utilizando como ponto de entrada os 12 meses anteriores ao período de previsão (lag = 12). Essa modelagem é similar ao modelo ARIMA, pois a previsão é gerada em um único passo para todos os 12 meses, sem a necessidade de iterar mês a mês.

Já no modelo LSTM com reposição, o estado interno da LSTM é reiniciado após cada lote de previsões. Nesse caso, a previsão segue uma lógica de recorrência mensal: para cada entrada de 12 meses anteriores, o modelo gera uma previsão para o mês subsequente. Isso implica que o modelo realiza 12 previsões de forma iterativa, utilizando tanto os dados reais quanto as previsões feitas para meses anteriores. Diferentemente do modelo sem reposição, as previsões anteriores podem ser usadas como entrada para ajustar o comportamento do modelo ao longo das iterações, o que pode corrigir as previsões com base no retorno obtido. No entanto, isso também pode aumentar o nível de erros, caso as previsões iniciais sejam imprecisas.

Além disso, é necessário definir a quantidade de épocas e neurônios que serão utilizados nos modelos. Dessa maneira, foram testadas diferentes combinações de neurônios e épocas de treinamento, com o intuito de obter a configuração que proporcionasse o melhor desempenho preditivo. Foram analisadas combinações de 15, 30, 50 e 100 neurônios, e épocas variando entre 100, 200, 300 e 500, totalizando 32 modelos (16 com reposição e 16 sem reposição).

3.5 Avaliação das previsões

Uma maneira de selecionar o melhor mecanismo de previsão é a comparação entre os valores previstos e observados, indicando a capacidade preditiva do modelo. Ao comparar metodologias de previsão aplicadas a uma única série temporal ou a várias séries temporais com as mesmas unidades, três métricas amplamente utilizadas para avaliar a qualidade das previsões são utilizadas, a Raiz Quadrada do Erro Quadrático Médio (RMSE), o Erro Absoluto Médio (EAM) e Erro Percentual Médio Absoluto (MAPE),

definidos como:

$$\text{RMSE} = \sqrt{\sum_{t=T+1}^{T+h} \frac{(Y_t - \hat{Y}_t)^2}{h}} \quad (7)$$

$$\text{EAM} = \sum_{t=T+1}^{T+h} \frac{|Y_t - \hat{Y}_t|}{h} \quad (8)$$

$$\text{MAPE} = \frac{1}{h} \sum_{t=1}^h \left| \frac{Y_t - \hat{Y}_t}{Y_t} \right| \cdot 100 \quad (9)$$

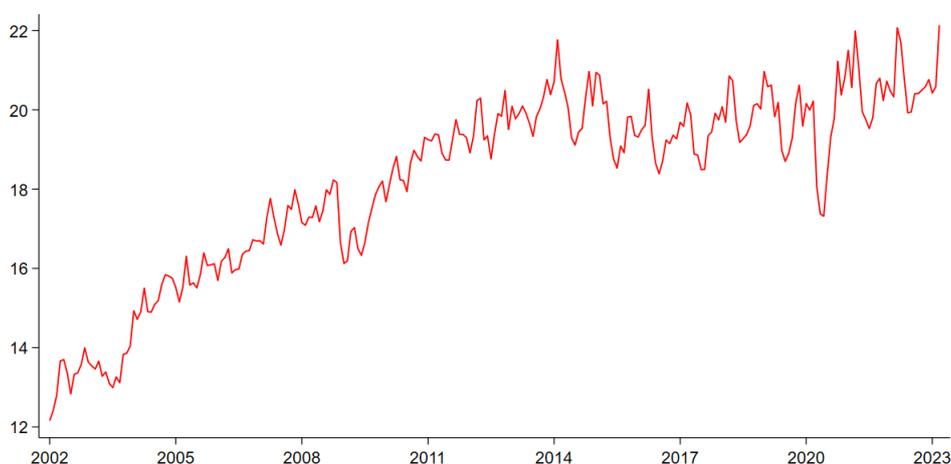
As três métricas são selecionadas a partir dos menores valores encontrados, indicando uma maior precisão das previsões dos modelos. A fim de comparar o desempenho preditivo dos modelos LSTM, o critério de escolha se dará com base no menor valor para o EAM. Será utilizado o MAPE para verificar a diferença percentual entre os valores observados e previstos.

4. Resultados e discussão

4.1 ARIMA

A Figura 2 apresenta o comportamento do consumo de energia elétrica total na região Sudeste entre janeiro de 2002 e março de 2023. Os dados fornecem indícios da não estacionariedade da série, uma vez que há uma tendência de crescimento ao longo dos anos, bem como sazonalidade.

Figura 2. Consumo de energia elétrica total na região Sudeste de 2002 a 2023

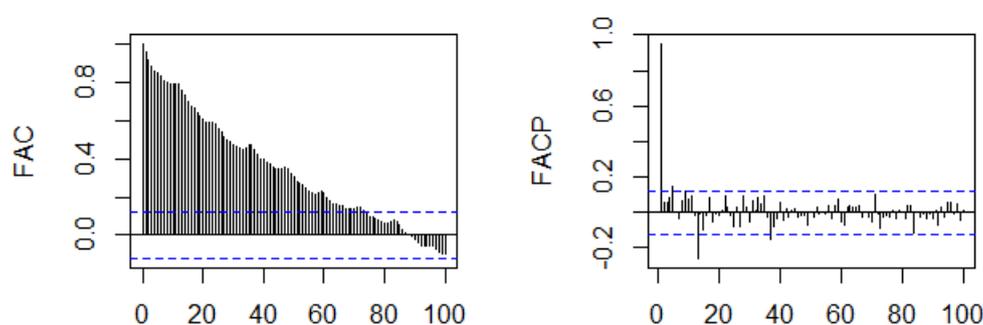


Fonte: Elaboração própria.

Ao examinar o correlograma da série, a FACP mostra um truncamento nas primeiras defasagens, indicando uma baixa ordem do componente autorregressivo (AR),

o que é típico de uma série estacionária. No entanto, a FAC exhibe um padrão de decaimento lento, sugerindo a presença de uma raiz unitária, indicando que a série possui um componente de tendência que persiste ao longo do tempo. Em séries estacionárias, as autocorrelações caem rapidamente para zero e o fato de a FAC decrescer lentamente e levar muitos *lags* para atingir zero indica que a série não é estacionária. A Figura 3 apresenta o correlograma da série.

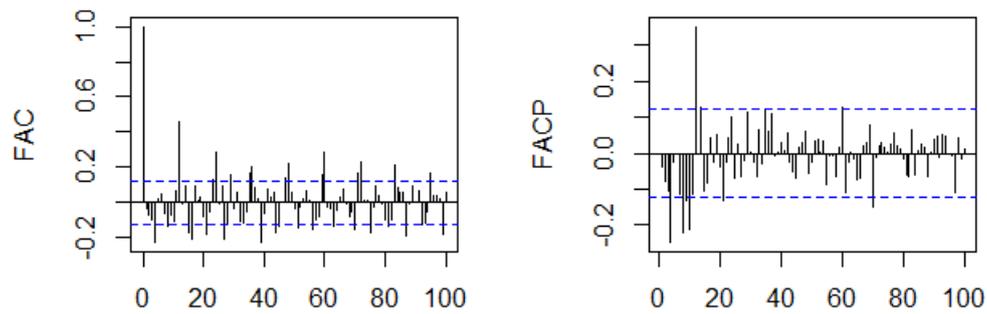
Figura 3. Função de Autocorrelação e Autocorrelação Parcial



Fonte: Elaboração própria.

Em seguida, foi aplicado o teste de Cox-Stuart para avaliar a presença de tendência na série de consumo de energia elétrica. O resultado indicou um valor-p igual a 0.000, permitindo rejeitar a hipótese nula de ausência de tendência ao nível de significância de 1%. Isso indica que a série apresenta uma tendência estatisticamente significativa ao longo do tempo. Esse resultado sugere que é necessário considerar essa tendência ao modelar a série temporal, aplicando uma diferenciação para remover a tendência e tornar a série estacionária antes de selecionar qual modelo ARIMA melhor se ajusta aos dados.

Desse modo, a série foi diferenciada para eliminar o componente de tendência e tornar o processo autorregressivo estacionário. Em seguida, foi realizado novamente o teste de Cox-Stuart e concluiu-se que a série do consumo de energia em primeira diferença não apresenta tendência ou um padrão sistemático significativo. A Figura 4 ilustra o correlograma da série diferenciada. Esse resultado mostra que a tendência foi removida, devido à convergência da FAC nas primeiras defasagens. No entanto, o comportamento senoidal ao longo das defasagens sugere que há um componente sazonal nos dados, que é esperado devido à natureza da série temporal utilizada.

Figura 4. Função de Autocorrelação e Autocorrelação Parcial diferenciada

Fonte: Elaboração própria.

Embora o correlograma seja um bom indicativo de estacionariedade, é necessário complementar a análise com testes formais de estacionariedade, como o teste ADF e PP. A Tabela 1 apresenta o teste de raiz unitária da série diferenciada.

Tabela 1. Testes de raiz unitária

	ADF	PP
Série	Valor do teste	Valor do teste
<i>consumo_t</i>	-12.13***	-214.77***

Fonte: Elaboração própria.

Nota: *** é significativa a 1%, ** a 5% e * a 10%.

A hipótese nula dos testes assume que a série possui uma raiz unitária, ou seja, que ela é não estacionária. Os resultados obtidos mostram que a hipótese nula de presença de raiz unitária pode ser rejeitada, indicando que a série diferenciada é estacionária em nível. O teste de HEGY indicou a presença de sazonalidade estocástica (p-valor igual a 0.000), o que normalmente implicaria a necessidade de diferenciação sazonal. Contudo, o modelo ARIMA busca otimizar o modelo se valendo dos critérios de informação AIC e BIC. Nesse caso, é possível selecionar um modelo sem diferenciação sazonal, mesmo que a sazonalidade tenha sido detectada pelo teste de HEGY, se o modelo resultante for considerado o de melhor ajuste, o que de fato ocorreu.

Ao forçar a inclusão de diferenciação sazonal e comparar o desempenho dos modelos sem diferenciação sazonal, encontra-se que os resíduos não são independentes e que há, portanto, autocorrelação significativa nos resíduos. Desse modo, os modelos sem diferenciação sazonal foram suficientes para capturar tanto a tendência quanto parte da sazonalidade da série. Assim, a Tabela 2 mostra onze especificações e os critérios para avaliar o melhor modelo de previsão.

Os resultados sugerem que o modelo ARIMA (1,1,0)(2,0,0)₁₂ se destaca com um RMSE de 0.8457 e um MAPE de 3.4216. O RMSE baixo mostra que as previsões têm um erro médio menor. Esses valores indicam que, em média, as previsões desse mo-

delo estão mais próximas dos valores reais em comparação aos outros modelos listados na Tabela 2 e, portanto, fornecem as melhores previsões para a série de consumo de energia elétrica. Além disso, o modelo selecionado possui um valor de Box-Pierce igual a 0.7744, o que indica uma boa adequação dos resíduos como ruído branco.

Como robustez, foi estimado um modelo SARIMA (3,1,2)(0,1,2)₁₂ para verificar se a inclusão explícita da componente sazonal poderia melhorar o desempenho preditivo. O teste de Box-Pierce para os resíduos do modelo SARIMA resultou em um valor-p igual a 0.0000, rejeitando a hipótese nula de ausência de autocorrelação dos resíduos a 1% de significância, evidenciando autocorrelação nos resíduos e indicando que o modelo não capturou adequadamente toda a estrutura da série.

Tabela 2. Modelos ARIMA estimados e critérios de qualidade

Especificação	RMSE	MAPE	AIC	BIC	Box-Pierce
ARIMA(1,1,1)(2,0,0) ₁₂	1.0190	3.8553	328.953	346.640	0.9081
ARIMA(1,1,0)(2,0,0) ₁₂	0.8487	3.4216	341.271	355.421	0.7744
ARIMA(2,1,0)(2,0,0) ₁₂	0.8935	3.4396	338.512	355.421	0.9743
ARIMA(2,1,0)(1,0,0) ₁₂	0.8754	3.5858	341.417	355.566	0.9217
ARIMA(2,1,1)(1,0,1) ₁₂	0.8765	3.1133	323.576	344.800	0.9332
ARIMA(1,1,1)(1,0,1) ₁₂	0.8798	3.1179	321.913	339.600	0.6801
ARIMA(0,1,1)(2,0,0) ₁₂	0.8638	3.4044	340.032	354.181	0.7807
ARIMA(1,1,2)(1,0,1) ₁₂	0.8764	3.1141	323.530	344.784	0.9697
ARIMA(0,1,2)(2,0,0) ₁₂	0.9397	3.5409	336.058	353.745	0.9638
ARIMA(2,1,2)(1,0,1) ₁₂	0.8847	3.1339	325.333	350.094	0.5107
SARIMA(3,1,2)(0,1,2) ₁₂	0.8811	3.1542	351.120	358.390	0.0000

Fonte: Elaboração própria.

Tabela 3. Valores previstos do modelo ARIMA (1,1,0)(2,0,0)₁₂

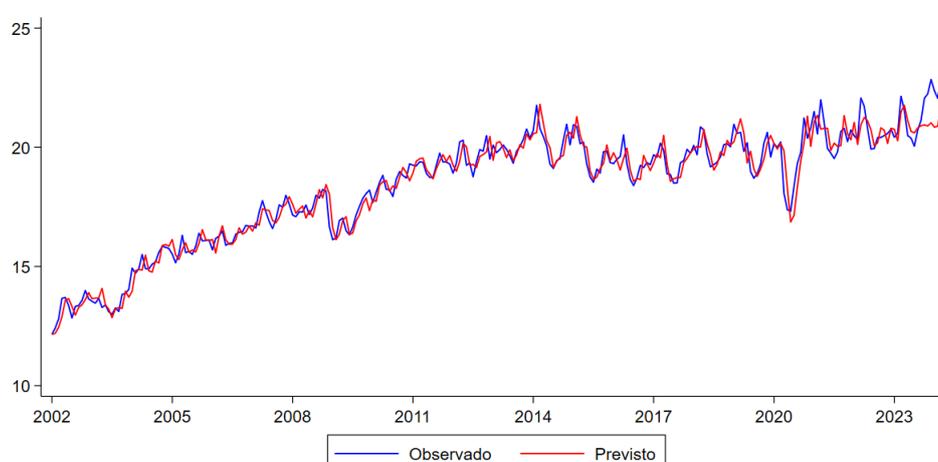
Mês	Observado	Previsto	Erro Absoluto	Erro Percentual
Abril/2023	21.55	21.76	0.21	0.96%
Mai/2023	20.49	21.13	0.64	3.11%
Junho/2023	20.39	20.67	0.28	1.37%
Julho/2023	20.05	20.60	0.55	2.79%
Agosto/2023	20.72	20.81	0.09	0.43%
Setembro/2023	21.13	20.90	0.23	1.08%
Outubro/2023	22.06	20.94	1.12	5.06%
Novembro/2023	22.24	20.89	1.35	6.06%
Dezembro/2023	22.85	21.03	1.82	7.96%
Janeiro/2024	22.37	20.84	1.53	6.84%
Fevereiro/2024	22.06	20.89	1.17	5.30%
Março/2024	23.02	21.79	1.23	5.34%

Fonte: Elaboração própria.

A Tabela 3 contém as previsões para o consumo de energia elétrica 12 períodos à frente, de abril de 2023 a março de 2024, utilizando a especificação ARIMA $(1,1,0)(2,0,0)_{12}$, bem como os valores observados deixados de fora da amostra.

A partir dos valores previstos, observa-se que as estimativas apresentam diferenças pequenas, em termos absolutos, em relação aos valores observados. Contudo, ao considerar a diferença percentual, notam-se valores maiores a partir de setembro de 2023. Graficamente, ao comparar o comportamento da série observada com a série estimada, obtém-se uma visão clara das projeções, que permite explorar possíveis razões para as diferenças entre os valores. A Figura 5 mostra o gráfico combinado das duas séries.

Figura 5. Consumo de energia elétrica observado e previsto.



Fonte: Elaboração própria.

Ao observar a Figura 5, nota-se que, ao longo dos anos, o modelo segue a tendência dos dados observados, capturando os picos e as quedas em geral. No entanto, há momentos em que o modelo subestima ou superestima os valores observados, especialmente em períodos de ocorrência de picos nas observações, sobretudo nos últimos 36 meses das observações. Portanto, algumas variações sazonais podem não ser bem capturadas pelo modelo, mas as métricas de erro encontradas foram baixas e sugerem que o modelo estimado se ajusta bem aos dados.

Embora a Figura 5 indique uma queda repentina no consumo de energia elétrica em 2020, esse comportamento não se manteve nos meses subsequentes, o que sugere a ocorrência de um ponto atípico e não de uma mudança estrutural no padrão da série. Para verificar formalmente essa possibilidade, foi aplicado o teste de Zivot-Andrews, que não identificou quebra estrutural no período da pandemia. Portanto, como o estudo se concentra na avaliação do desempenho preditivo dos modelos ao longo da série completa, optou-se por não segmentar a amostra, à medida que a presença de quebra estrutural no período da pandemia não foi estatisticamente significativa.

4.2 LSTM

Ao realizar a previsão do consumo de energia elétrica da região Sudeste utilizando modelos LSTM, observam-se diferenças significativas no desempenho entre as configurações dos modelos com e sem reposição. O modelo com reposição, ajustado com 15 neurônios e 500 épocas, apresentou o melhor resultado em termos de RMSE, atingindo um valor de 1.4975, conforme a Tabela 4.

Tabela 4. RMSE para os modelos LSTM com reposição

Neurônios	Épocas			
	100	200	300	500
15	3.8347	1.7795	3.2414	1.4975
30	2.0505	1.9535	1.8793	2.3656
50	1.9037	2.2387	6.5188	2.6669
100	2.3004	2.8601	7.5400	8.7188

Fonte: Elaboração própria.

Aumentar o número de neurônios nem sempre resulta em uma melhora de desempenho. Por exemplo, o modelo com 50 neurônios e 300 épocas teve um RMSE muito elevado (6.5188), indicando sobreajuste ou instabilidade no treinamento. A Tabela 5 mostra os valores previstos e os valores observados do modelo LSTM com reposição, bem como os respectivos erros absolutos de previsão.

Tabela 5. Valores previstos e observados do modelo LSTM com reposição

Mês	Observado	Previsto	Erro Absoluto	Erro Percentual
Abril/2023	21.55	22.42	0.87	4.04%
Maio/2023	20.49	22.47	1.98	9.66%
Junho/2023	20.38	22.26	1.88	9.23%
Julho/2023	20.04	21.31	1.27	6.34%
Agosto/2023	20.72	21.13	0.41	1.98%
Setembro/2023	21.13	21.29	0.16	0.76%
Outubro/2023	22.05	21.64	0.41	1.86%
Novembro/2023	22.24	21.96	0.28	1.26%
Dezembro/2023	22.84	22.53	0.31	1.36%
Janeiro/2024	22.36	22.59	0.23	1.03%
Fevereiro/2024	22.05	21.73	0.32	1.45%
Março/2024	23.02	22.35	0.67	2.91%

Fonte: Elaboração própria.

A Tabela 5 mostra que, para os 12 meses previstos, o modelo com reposição teve erros absolutos de previsão relativamente baixos, com o maior erro sendo 1.98 (maio/2023) e o menor erro sendo 0.16 (setembro/2023). Esse resultado sugere que, apesar de

algumas discrepâncias, o modelo capturou bem a tendência dos dados observados. Ademais, nota-se que, a partir de setembro de 2023, o modelo conseguiu captar mais de perto o comportamento dos dados observados.

O modelo LSTM sem reposição também apresentou uma variação no RMSE conforme o número de épocas e neurônios. O modelo sem reposição, ajustado com 100 neurônios e 200 épocas, apresentou o melhor resultado em termos de RMSE, atingindo um valor de 1.7750, conforme mostrado na Tabela 6. Esse resultado indica que ajustes intermediários podem ser mais adequados.

Tabela 6. RMSE para os modelos LSTM sem reposição

Neurônios	Épocas			
	100	200	300	500
15	1.8005	2.9129	2.1437	2.4852
30	2.2969	2.4205	4.0189	5.5798
50	3.0780	2.0505	4.1265	2.8720
100	2.6119	1.7750	2.2542	2.5522

Fonte: Elaboração própria.

Em seguida, a Tabela 7 revela que o modelo sem reposição também teve erros absolutos de previsão relativamente baixos, com o maior erro sendo 2.18 (julho/2023) e o menor erro sendo 0.06 (fevereiro/2024). Em comparação ao modelo com reposição, o erro médio é ligeiramente superior, mas ainda aceitável em termos gerais de previsão.

Tabela 7. Valores previstos e observados do modelo LSTM sem reposição

Mês	Observado	Previsto	Erro Absoluto	Erro Percentual
Abril/2023	21.55	22.60	1.05	4.87%
Maió/2023	20.49	22.27	1.78	8.69%
Junho/2023	20.38	21.98	1.60	7.85%
Julho/2023	20.04	22.22	2.18	10.88%
Agosto/2023	20.72	21.18	0.46	2.22%
Setembro/2023	21.13	21.25	0.12	0.57%
Outubro/2023	22.05	21.76	0.29	1.32%
Novembro/2023	22.24	21.98	0.26	1.17%
Dezembro/2023	22.84	22.92	0.08	0.35%
Janeiro/2024	22.36	22.54	0.18	0.80%
Fevereiro/2024	22.05	21.99	0.06	0.27%
Março/2024	23.02	22.25	0.77	3.34%

Fonte: Elaboração própria.

Assim como no modelo com reposição, o erro absoluto máximo foi encontrado nos primeiros meses da previsão (maio e julho de 2023), indicando dificuldades na captura

de certas flutuações sazonais ou variações fortes de consumo. Da mesma forma que o modelo anterior com reposição, a modelagem sem reposição mostra consistência de previsão a partir de setembro de 2023. Comparando os dois modelos, as estimativas baseadas no RMSE sugerem que o modelo com reposição foi mais eficiente em capturar as variações temporais do consumo de energia na região Sudeste.

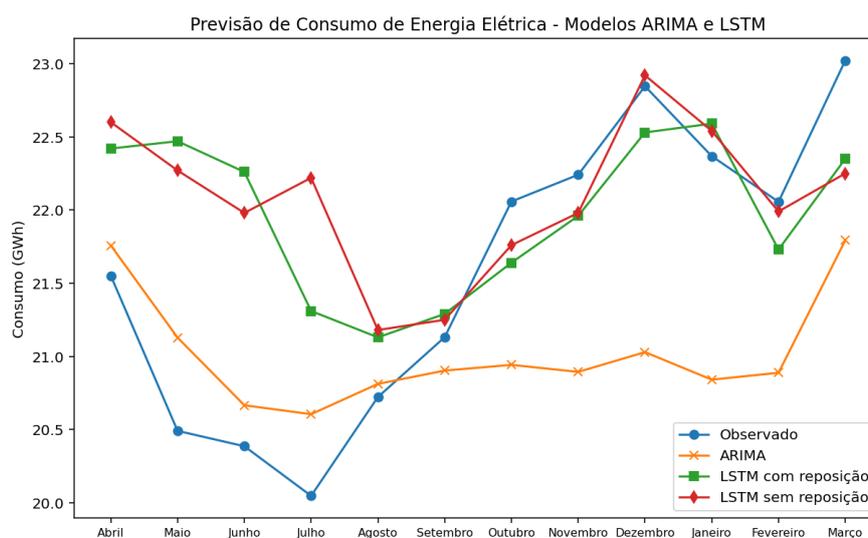
Vale ressaltar que o aumento do número de neurônios e épocas não resultou em melhorias de desempenho para ambos os modelos. Em configurações com muitos neurônios, como 50 e 100, observa-se um aumento nos valores do RMSE, especialmente no modelo com reposição, o que pode indicar um sobreajuste (*overfitting*). Nesses casos, o modelo se ajusta bem aos dados de treino, capturando ruídos que prejudicaram sua capacidade de generalização para os dados de teste. Por outro lado, configurações com poucas épocas ou neurônios, como as com 100 épocas, podem ter sofrido de subajuste (*underfitting*), resultando em previsões menos precisas.

Um aspecto importante que pode explicar os erros de previsão nos primeiros meses está relacionado às variações sazonais e às mudanças inesperadas no consumo de energia que não estavam refletidas nos dados utilizados para o treinamento. Embora o modelo LSTM seja capaz de capturar padrões temporais complexos, a dificuldade em prever esses picos sugere a necessidade de um maior refinamento na escolha de janela temporal e defasagens, que podem não ter captado adequadamente as relações de dependência temporal dos dados.

4.3 Comparando os três modelos

O modelo ARIMA obteve o menor RMSE geral, sugerindo maior precisão nas previsões. No entanto, sua performance apresentou falhas na identificação de mudanças estruturais e variações sazonais significativas. Por sua vez, o modelo LSTM com reposição obteve um desempenho superior em relação ao LSTM sem reposição, sobretudo nas especificações com menos neurônios e mais épocas. Além disso, o modelo ARIMA é eficaz em capturar tendências lineares e padrões sazonais, mas mostrou limitações diante de não linearidades e possíveis mudanças estruturais na série. Já os modelos LSTM, com e sem reposição, foram mais ajustáveis ao lidar com não linearidades, o que permitiu uma melhor captura de padrões mais complexos, embora o RMSE dos modelos LSTM selecionados tenha sido superior ao do ARIMA, corroborando a literatura de que os modelos LSTM apresentam melhores desempenhos com dados não lineares (Zhang, 2003; Kourentzes e Crone, 2011; Alsheheri, 2025).

A Figura 6 exibe a previsão de consumo de energia elétrica de abril de 2023 a março de 2024, com os três melhores modelos selecionados (ARIMA, LSTM com reposição e LSTM sem reposição) comparados aos dados observados.

Figura 6. Valores observados e previstos dos modelos estimados

Fonte: Elaboração própria.

O modelo ARIMA mostrou uma maior dificuldade em capturar variações sazonais expressivas, sobretudo após 2020, o que pode ser indicativo de uma mudança nos padrões de consumo. Já os modelos LSTM, em especial o modelo com reposição, mostraram maior capacidade de adaptação a mudanças sazonais, embora com maiores variações de erro. Em termos de desempenho geral, nenhum dos modelos captura perfeitamente a tendência dos dados observados. O modelo LSTM com reposição, apesar de apresentar erros crescentes nos meses finais, é o mais equilibrado ao longo do tempo, enquanto o ARIMA é o mais conservador, subestimando o consumo. Assim, a escolha do modelo ideal para previsões de consumo de energia deve levar em consideração tanto a estabilidade quanto a capacidade de capturar mudanças rápidas, dependendo do objetivo da previsão.

Os resultados obtidos neste estudo apresentam pontos de convergência, mas também de divergência com a literatura existente. Conforme discutido na seção 2, a aplicação de modelos ARIMA e LSTM para previsão de consumo de energia é uma prática recorrente e um assunto de interesse constante dos pesquisadores. Oliveira e Oliveira (2018) encontram estimativas consistentes para a previsão do consumo de energia elétrica em um grupo de 24 países, utilizando o modelo ARIMA. Os resultados buscam validar a utilização de metodologias univariadas para realizar previsões de médio e longo prazo, valendo-se da estatística MAPE para validar as estimativas.

No presente estudo, embora as estatísticas de ajuste tenham sido menores para o modelo ARIMA, o modelo LSTM teve melhor desempenho de previsão. Nesse sentido, os resultados obtidos corroboram os achados de Zheng et al. (2017), em que os experimentos de previsão conduzidos foram capazes de prever com precisão a demanda por energia elétrica no longo prazo, utilizando modelos LSTM, em comparação com outras modelagens, a exemplo do ARIMA.

Por sua vez, Santos e Chaucoski (2020) encontraram estimativas próximas aos valores observados da amostra ao utilizarem os modelos ARIMA e LSTM, mesmo que a previsão tenha sido feita apenas cinco períodos à frente. Marcos e Júnior (2021) fazem exercícios empíricos similares usando modelo ARIMA para a região Nordeste. Embora as discussões metodológica e empírica não sejam os pontos principais destes estudos, os resultados gerais apontam para bons ajustes matemáticos dos modelos selecionados, reforçando a eficácia de aplicar modelos univariados de séries temporais para previsão do consumo de energia elétrica.

Portanto, feitas as considerações a respeito dos resultados obtidos neste trabalho em consonância com os achados de outros autores, é preciso reforçar as implicações dos resultados encontrados e ressaltar a importância para a literatura. Nesse sentido, o presente estudo fornece robustez metodológica e empírica à literatura do tema, sobretudo em âmbito brasileiro, ao discutir a importância dessa categoria de estudos para a previsão do consumo de eletricidade da região Sudeste, bem como uma análise do estágio atual das pesquisas desenvolvidas no Brasil.

Em termos práticos, essas previsões são importantes para orientar o planejamento da geração e da distribuição de energia, especialmente em contextos de alteração nos padrões de consumo provocada por variações climáticas e econômicas. A boa performance dos modelos avaliados fornece aos gestores públicos uma base para planejar ações de curto prazo, como a antecipação de picos sazonais de consumo, a definição de medidas de eficiência energética e a alocação adequada de recursos em infraestrutura. No setor privado, as previsões geradas podem auxiliar empresas de energia na gestão operacional, no planejamento da oferta e demanda contratada, bem como na definição de estratégias comerciais e de precificação mais alinhadas ao comportamento do mercado.

Além disso, em um contexto de transição energética e compromissos internacionais de redução de emissões, previsões precisas de demanda elétrica são importantes para orientar políticas públicas voltadas à segurança do sistema energético elétrico e ao cumprimento das metas de sustentabilidade. Dessa forma, os modelos estimados contribuem não apenas para a melhoria da gestão do sistema elétrico, mas também para o alinhamento das decisões do setor aos objetivos mais amplos de desenvolvimento econômico e ambiental.

Os resultados corroboram a continuidade do uso de modelos univariados na captura de padrões sazonais e tendências em séries temporais de consumo de energia elétrica. Contudo, a presente análise aprofunda a discussão metodológica, uma vez que são exploradas diferentes configurações dos modelos LSTM, com e sem reposição, demonstrando como a escolha da arquitetura da rede neural pode impactar o desempenho preditivo. Em síntese, este estudo se alinha com a literatura ao utilizar os modelos ARIMA e LSTM para previsão de consumo de energia elétrica ao encontrar resultados consistentes com a proposta do estudo, mas contribui com uma análise mais detalhada dos resultados, com comparação entre diferentes arquiteturas de

LSTM e com aprofundamento metodológico.

5. Considerações finais

Neste estudo, foram explorados modelos ARIMA e LSTM com e sem reposição, para realizar previsões de curto prazo do consumo de energia elétrica, com o objetivo de identificar a melhor abordagem para lidar com essas séries temporais. A previsão mais assertiva possível do consumo de energia é estratégica para o planejamento de oferta e demanda de eletricidade. Previsões errôneas podem resultar em sobrecarga do sistema, desabastecimento ou subutilização da infraestrutura energética. Modelos como o LSTM, ao captar elementos não lineares nos dados, podem ajudar a ajustar os cronogramas de geração de energia de fontes renováveis e evitar o uso excessivo de fontes não renováveis durante os períodos de pico.

O estudo contribui para a literatura ao comparar a eficácia de modelos tradicionais, como ARIMA, com abordagens modernas de redes neurais, como LSTM, aplicadas ao contexto energético. Essa comparação sugere que, em contextos mais voláteis ou com variações não lineares, os modelos LSTM apresentam vantagens, mesmo que a um custo computacional mais elevado.

As previsões foram realizadas para um horizonte de 12 meses, o que limita a aplicação dos resultados para projeções de médio e longo prazo, mas não é um fator limitante para este estudo. No entanto, devido à natureza do modelo ARIMA em modelar relações de curto prazo, sua eficácia diminui em horizontes mais longos, principalmente na presença de não linearidades ou mudanças estruturais nos dados. Já o modelo LSTM, que é capaz de lidar com dependências de longo prazo e padrões não lineares, também foi utilizado para previsões no curto prazo neste estudo. Isso permitiu a comparação entre as duas abordagens, uma vez que o LSTM pode ser aplicado em horizontes tanto curtos quanto longos.

Estudos futuros podem explorar horizontes mais longos, integrando cenários econômicos e climáticos variáveis para ampliar a robustez das previsões. A abordagem LSTM com reposição, embora robusta, mostrou-se suscetível à propagação de erros, o que pode comprometer a precisão das previsões. Uma abordagem híbrida, que combine modelos ARIMA para capturar sazonalidades e LSTM para lidar com não linearidades, pode mitigar esse problema.

Referências

- Ackermann, A. E. e Sellitto, M. A. (2022). Métodos de previsão de demanda: uma revisão da literatura. *Innovar*, 32(85):83–99.
- Adeosun, T., Amosu, C., Omitogun, O., Amusa, O., e Morenikeji, B. (2023). Renewable energy for sustainable development in developing countries: benefits to the environment. *J. Energy Res. Rev*, 13(11).

- Ahmad, T. e Zhang, D. (2020). A critical review of comparative global historical energy consumption and future demand: The story told so far. *Energy Reports*, 6:1973–1991.
- Alsheheri, G. (2025). Comparative analysis of arima and nnar models for time series forecasting. *Journal of Applied Mathematics and Physics*, 13(1):267–280.
- Amaral, L. S., de Araújo, G. M., e de Moraes, R. A. R. (2022). Analysis of the factors that influence the performance of an energy demand forecasting model. *Advanced Notes in Information Science*, 2:92–102.
- Amaro, R. S., Ceretta, P. S., Coronel, D. A., Bender Filho, R., e de Oliveira, M. G. (2017). Comparação da capacidade preditiva de modelos arima para o consumo de energia elétrica no brasil. *Revista Unemat de Contabilidade*, 6(11).
- Barkhordar, Z. A. e Saboohi, Y. (2014). Modelling useful energy demand system as derived from basic needs in the household sector. *Energy Efficiency*, 7:903–921.
- Barros, T. M. B. (2014). Previsão de carga-comparação de técnicas. Dissertação de Mestrado, Universidade do Porto (Portugal).
- Belém, G. d. S. (2024). Previsão de curto prazo da demanda de energia elétrica no nordeste brasileiro utilizando séries temporais. Trabalho de Conclusão de Curso (Graduação em Engenharia Elétrica), Instituto Federal de Pernambuco, Garanhuns.
- Billewar, S. R., Londhe, G., e Mane, P. S. (2023). World energy demand. *Integrated Green Energy Solutions Volume 2*, Página 275–316.
- Box, G. E. P. e Jenkins, G. M. (1976). Time series analysis: Forecasting and control. *Holden-Day*, revised ed.
- Bueno, R. D. L. d. S. (2018). *Econometria de séries temporais*. São Paulo: Cengage Learning.
- Cabreira, M. M. L., Leite Coelho da Silva, F., da Silva Codeiro, J., Serrano Hernández, R. M., Canas Rodrigues, P., e López-Gonzales, J. L. (2024). A hybrid approach for hierarchical forecasting of industrial electricity consumption in brazil. *Energies*, 17(13):3200.
- Campos, R. J. (2008). Previsão de séries temporais com aplicações a séries de consumo de energia elétrica. Dissertação (mestrado em engenharia elétrica), Universidade Federal de Minas Gerais.
- Choi, I. (2015). *Almost all about unit roots: Foundations, developments, and applications*. Cambridge University Press.
- Chujai, P., Kerdprasop, N., e Kerdprasop, K. (2013). Time series analysis of household electric consumption with arima and arma models. In: *Proceedings of the international multiconference of engineers and computer scientists*, volume 1, Página 295–300.

- da Silva Costa, A., Queiroz, J. C. B., Chermont, L. S., Lameira, O. A., de Souza, E. B., Diniz, M. B., Moura, H. P., e Costa, D. L. C. (2021). Deforestation forecasts in the legal amazon using intervention models. *Research, Society and Development*, 10(4):e8710413787–e8710413787.
- de Oliveira, E. M. e Oliveira, F. L. C. (2018). Forecasting mid-long term electric energy consumption through bagging arima and exponential smoothing methods. *Energy*, 144:776–788.
- Enders, W. (2015). *Applied econometric time series fourth edition*. New York (US): University of Alabama, 4 edition.
- EPE (2020). *Plano Nacional de Energia 2050*. Brasília: Empresa de Pesquisa Energética, Ministério de Minas e Energia - MME.
- EPE (2024a). *Balanço Energético Nacional 2023: Ano Base 2022*. Rio de Janeiro: Empresa de Pesquisa Energética, Ministério de Minas e Energia - MME.
- EPE (2024b). *Resenha Mensal do Mercado de Energia Elétrica: Base Dezembro 2023*. Number 196. Brasília: Empresa de Pesquisa Energética, Ministério de Minas e Energia - MME.
- Feitosa, R. M. T. (2023). Comparação de métodos clássicos e avançados de previsão aplicados ao consumo de energia elétrica. Trabalho de Conclusão de Curso (Graduação em Engenharia Elétrica), Universidade Federal de São Carlos.
- Franses, P. H. (1990). Testing for seasonal unit roots in monthly data. *Econometric Institute Report*.
- Freitas, F., De Souza, S., Ferreira, L., Otto, R., Alessio, F., De Souza, S., Venturini, O., e Junior, O. A. (2019). The brazilian market of distributed biogas generation: Overview, technological development and case study. *Renewable and Sustainable Energy Reviews*, 101:146–157.
- Gonçalves, I. G. e Ferreira, M. D. P. (2023). Modelos para a previsão de alertas de desmatamento na amazônia legal. *Revista Brasileira de Estudos Regionais e Urbanos*, 17(4):537–560.
- Gonçalves, I. G. (2022). Emergência energética: impactos econômicos do aumento do consumo de energia pelo setor de energia. Dissertação (mestrado em economia aplicada), Universidade Federal de Viçosa.
- Graves, A. (2013). Generating sequences with recurrent neural networks. *arXiv preprint arXiv:1308.0850*.
- Haykin, S. (2001). *Redes neurais: princípios e prática*. Bookman Editora, 2 edition.
- He, K., Ji, L., Wu, C. W. D., e Tso, K. F. G. (2021). Using sarima–cnn–lstm approach to forecast daily tourism demand. *Journal of Hospitality and Tourism Management*, 49:25–33.

- Hewamalage, H., Bergmeir, C., e Bandara, K. (2021). Recurrent neural networks for time series forecasting: Current status and future directions. *International Journal of Forecasting*, 37(1):388–427.
- Hochreiter, S. e Schmidhuber, J. (1997). Long short-term memory. *Neural computation*, 9(8):1735–1780.
- Hyndman, R. J. e Athanasopoulos, G. (2025). *Forecasting: principles and practice*. OTexts: Melbourne, Australia, 3 edition.
- Hyndman, R. J. e Khandakar, Y. (2008). Automatic time series forecasting: the forecast package for r. *Journal of statistical software*, 27:1–22.
- IBGE (2023). *PNAD Contínua: Características Gerais dos Domicílios e dos Moradores 2022*. Rio de Janeiro: Instituto Brasileiro de Geografia e Estatística - Coordenação de Pesquisas por Amostra de Domicílios.
- IEA (2023). *World Energy Outlook 2023*. Paris: International Energy Agency (IEA).
- IEA (2024). *Electricity 2024 – Analysis and forecast to 2026*. Paris: International Energy Agency (IEA).
- Kardakos, E. G., Alexiadis, M. C., Vagropoulos, S. I., Simoglou, C. K., Biskas, P. N., e Bakirtzis, A. G. (2013). Application of time series and artificial neural network models in short-term forecasting of pv power generation. In: *2013 48th International Universities' Power Engineering Conference (UPEC)*, Página 1–6. IEEE.
- Kaur, J., Parmar, K. S., e Singh, S. (2023). Autoregressive models in environmental forecasting time series: a theoretical and application review. *Environmental Science and Pollution Research*, 30(8):19617–19641.
- Kourentzes, N. e Crone, S. F. (2011). Semi-supervised monitoring of electric load time series for unusual patterns. In: *The 2011 International Joint Conference on Neural Networks*, Página 2852–2859. IEEE.
- Lütkepohl, H. e Krätzig, M. (2004). *Applied time series econometrics*. Cambridge University Press.
- Marcos, I. P. e Júnior, A. P. P. (2021). Previsão do consumo de energia elétrica na região nordeste do brasil. *Revista de Engenharia e Pesquisa Aplicada*, 6(3):21–30.
- Matos, E. A. d. S. (2022). Análise e previsão de demanda de energia elétrica no sudeste e centro-oeste do brasil através de modelos de séries temporais. Trabalho de Conclusão de Curso (Graduação em Estatística), Universidade Federal Fluminense.
- MME (2023). *Programa Luz para Todos: um marco histórico 10 milhões de brasileiros saíram da escuridão*. Brasília: Ministério de Minas e Energia (MME).
- Morettin, P. A. e Toloi, C. M. d. C. (2022). *Análise de séries temporais*. Blucher, São Paulo.

- Nunes, G. d. S. (2019). O uso dos métodos arima e var-vec no estudo da demanda de energia elétrica no rio grande do sul. Dissertação (mestrado em modelagem computacional), Universidade Federal do Rio Grande.
- Olah, C. (2015). Understanding lstm networks. <https://colah.github.io/posts/2015-08-Understanding-LSTMs/>.
- Pierre, A. A., Akim, S. A., Semenyó, A. K., e Babiga, B. (2023). Peak electrical energy consumption prediction by arima, lstm, gru, arima-lstm and arima-gru approaches. *Energies*, 16(12):4739.
- Pires, C. L. (2016). Metodologia para previsão de carga e geração no horizonte de curtíssimo prazo. Dissertação (mestrado em engenharia elétrica), Universidade Federal de Santa Maria.
- Pooniwala, N. e Sutar, R. (2021). Forecasting short-term electric load with a hybrid of arima model and lstm network. In: *2021 International Conference on Computer Communication and Informatics (ICCCI)*, Página 1–6.
- Rahman, M. M. (2020). Environmental degradation: The role of electricity consumption, economic growth and globalisation. *Journal of Environmental Management*, 253:109742.
- Rahman, S. e Hazim, O. (1993). A generalized knowledge-based short-term load-forecasting technique. *IEEE Transactions on Power Systems*, 8(2):508–514.
- Ritchie, H., Rosado, P., e Roser, M. (2024). Access to energy. *Our World in Data*. <https://ourworldindata.org/energy-access>.
- Sachin, M. M., Baby, M. P., e Ponraj, A. S. (2020). Analysis of energy consumption using rnn-lstm and arima model. In: *Journal of Physics: Conference Series*, volume 1716, Página 012048.
- Sachs, J. D. (2015). *The age of sustainable development*. Columbia University Press, New York.
- Sambodo, M. T., Fuady, A. H., Negara, S. D., Handoyo, F. W., e Mychelisda, E. (2021). Introduction: Access to electricity and welfare. In: *Electricity Access and Community Welfare in Indonesia*, Página 1–10. Springer: Singapore.
- Santos, F. R. M. d. (2019). Previsão de médio prazo de demanda de energia elétrica do setor industrial na região sul e sudeste do brasil. Trabalho de Conclusão de Curso (Graduação em Engenharia de Produção), Universidade Federal do Paraná.
- Santos, J. A. A. e Chaucoski, Y. (2020). Previsão do consumo de energia elétrica na região sudeste: um estudo de caso usando sarima e lstm. *Revista Cereus*, 12(4):93–104.

- Silva, V. R. (2022). Modelos de séries temporais aplicados ao consumo de energia elétrica na região sudeste. Trabalho de Conclusão de Curso (Graduação em Engenharia Química), Universidade Estadual Paulista.
- Simões, J. V. O., Alves, L. M., Pedra, G. U., e Barreto, N. d. J. d. C. (2025). Spatio-temporal assessment of extreme precipitation events in southeastern brazil. *Modeling Earth Systems and Environment*, 11(3):212.
- Svetunkov, I. (2022). Statistics for business analytics. lecture notes. <https://openforecast.org>. OpenForecast, Lancaster, UK.
- Svetunkov, I. (2023). *Forecasting and Analytics with the Augmented Dynamic Adaptive Model (ADAM)*. CRC Press, Chapman and Hall Book, 1 edition.
- Tidre, P. V., Biase, N. G., e Silva, M. I. d. S. (2013). Utilização dos modelos de séries temporais na previsão do consumo mensal de energia elétrica da região norte do brasil. *Matemática e Estatística em Foco*, 1(1):57–66.
- UN (2016). *The sustainable development goals report 2016*. United Nations Department of Economic and Social Affairs, New York.
- Vassalli, L. C. (2018). Aplicação de redes neurais lstm para a previsão de curto prazo de vazão do rio paraíba do sul. Trabalho de Conclusão de Curso (Graduação em Engenharia Computacional), Universidade Federal de Juiz de Fora.
- Velasquez, C., Zocatelli, M., Estanislau, F. B., e Castro, V. F. (2022). Analysis of time series models for brazilian electricity demand forecasting. *Energy*, 247:123483.
- Zhang, G., Patuwo, E., e Hu, M. (1998). Forecasting with artificial neural networks: The state of the art. *International Journal of Forecasting*, 14(1):35–62.
- Zhang, P. (2003). Time series forecasting using a hybrid arima and neural network model. *Neurocomputing*, 50:159–175.
- Zheng, J., Xu, C., Zhang, Z., e Li, X. (2017). Electric load forecasting in smart grids using long-short-term-memory based recurrent neural network. In: *51st Annual Conference on Information Sciences and Systems*, Página 1–6.
- Zuin, G., Buechler, R., Sun, T., Zanocco, C., Castro, D., Veloso, A., e Rajagopal, R. (2022). Revealing the impact of extreme events on electricity consumption in brazil: A data-driven counterfactual approach. In: *2022 IEEE Power & Energy Society General Meeting (PESGM)*, Página 1–5. IEEE.