

MODELO HÍBRIDO PROPHET+LIGHTGBM PARA PREVISÃO DE DEMANDA ELÉTRICA: UMA ABORDAGEM EM DUAS CAMADAS PARA O SUBSISTEMA SUDESTE DO BRASIL

Elizeu Silva de Medeiros¹, Romário de Sousa Almeida^{1,2}, Ítalo Gabriel da Silva¹, Alexandre Bezerra de Souza¹, Eduarda da Silva Brito¹, José Wanderson Ferreira de Freitas¹

¹Universidade Estadual da Paraíba - UEPB, Campina Grande, Brasil
(elizeumedeiros@gmail.com)

²Universidade Federal de Lavras - UFLA, Campina Grande, Brasil

Resumo: Modelos de decomposição probabilística, como o Prophet, capturam tendência e sazonalidades com eficiência, porém deixam resíduos com estrutura não linear que algoritmos de *boosting* podem explorar. Este trabalho propõe e avalia um modelo híbrido para previsão diária da demanda elétrica no subsistema Sudeste/Centro-Oeste do Brasil, combinando o Prophet com o LightGBM treinado sobre os resíduos da primeira camada. Com dados horários do ONS (2000-2025) e período de teste de quatro anos (2022-2025), o híbrido reduz o RMSE em 42,1% e o MAPE em 46,7% em relação ao Prophet isolado. A superioridade é confirmada pelo teste de Diebold-Mariano (DM = 27,47; $p < 0,001$).

Palavras-chave: Previsão de carga elétrica; Modelagem híbrida; Aprendizado de máquina; Desempenho preditivo; Sistemas de energia.

INTRODUÇÃO

A previsão da demanda de energia elétrica ocupa papel central no planejamento e na operação de sistemas elétricos contemporâneos. Antecipar o comportamento futuro do consumo é fundamental para: manter o equilíbrio em tempo real entre geração e carga; assegurar a confiabilidade do fornecimento; promover eficiência econômica no despacho de energia. Nesse contexto, os erros de previsão dos fatores abordados geram impactos financeiros concretos e prejudiciais ao sistema.

A subestimação da demanda exige o acionamento emergencial de unidades geradoras menos eficientes, enquanto a superestimação resulta na manutenção de capacidade ociosa e na contratação de energia que não chega a ser utilizada, elevando os custos operacionais do sistema (LÚCIO *et al.*, 2018).

Em sistemas caracterizados por elevada complexidade estrutural, tais impactos tendem a se intensificar, tornando a qualidade das previsões um fator diretamente associado à segurança energética e à sustentabilidade econômica do setor.

O sistema elétrico brasileiro destaca-se como um dos mais complexos do mundo. Essa complexidade está ligada, sobretudo, à forte dependência da geração hidrelétrica e à necessidade de coordenar diferentes

fontes de geração distribuídas por um território de dimensões continentais. A operação desse sistema é conduzida pelo Operador Nacional do Sistema Elétrico (ONS), entidade responsável pela coordenação centralizada do Sistema Interligado Nacional (SIN).

Esse sistema integra a maior parte da infraestrutura de geração, transmissão e consumo de energia elétrica no país. Nesse arranjo, o atendimento à demanda depende de uma articulação constante entre os reservatórios hidrelétricos, que exercem papel fundamental na regulação da oferta de energia, e as usinas termelétricas, acionadas de forma complementar em períodos de escassez hídrica. Em um modelo hidrotérmico como o brasileiro, a variabilidade das condições hidrológicas acrescenta um nível adicional de incerteza ao planejamento da oferta de energia.

Diante disso, a definição do despacho ótimo, entendido como a alocação eficiente dos recursos de geração para atendimento da demanda, depende diretamente da confiabilidade das previsões de carga elétrica.

Quando há desvios sistemáticos entre os valores previstos e os efetivamente observados, surgem impactos que vão além do aumento de custos operacionais, afetando também os critérios de



segurança do sistema e os mecanismos de formação de preços no mercado de curto prazo (VELASQUEZ *et al.*, 2022).

Dentro do SIN, o subsistema Sudeste/Centro-Oeste responde por mais de metade do consumo nacional de eletricidade (ARIAS *et al.*, 2023), exercendo influência decisiva sobre o equilíbrio do sistema interligado.

O comportamento da demanda nessa região apresenta marcada sazonalidade intra-semanal, fortemente associada à estrutura industrial local. Em geral, observa-se redução significativa do consumo durante fins de semana e feriados em comparação aos dias úteis.

Além disso, há ciclos sazonais anuais bem definidos, com picos de consumo no verão austral, impulsionados pelo uso intensivo de sistemas de climatização, e níveis mais baixos durante o inverno. Tais características exigem que os modelos de previsão incorporem tanto os padrões intra-semanais quanto as variações sazonais de longo prazo para alcançar estimativas confiáveis (SOARES; MEDEIROS, 2008).

A complexidade desse cenário foi ampliada por eventos recentes que alteraram de maneira estrutural o padrão de demanda elétrica. A pandemia da COVID-19, em 2020, provocou uma ruptura abrupta na trajetória de crescimento do consumo, com impactos distintos entre setores econômicos e regiões do país. A partir de 2021, entretanto, observou-se uma recuperação econômica acelerada, que levou a níveis de demanda superiores aos máximos históricos registrados antes da pandemia, especialmente nos verões de 2023 e 2024.

Esses períodos foram acompanhados por ondas de calor de intensidade incomum na região Sudeste, ampliando significativamente o consumo associado à climatização e colocando à prova modelos calibrados sob regimes climáticos anteriores. A combinação entre mudanças estruturais pós-pandemia e choques climáticos de grande magnitude evidencia a necessidade de abordagens metodológicas capazes de lidar com não estacionaridades e padrões de demanda não lineares.

A literatura dedicada à previsão de carga elétrica evoluiu, de modo geral, a partir de duas vertentes metodológicas principais. A primeira reúne modelos estatísticos clássicos de séries temporais, como ARIMA, ETS e TBATS, reconhecidos por sua interpretabilidade, fundamentação teórica e bom desempenho em contextos com sazonalidades regulares.

Contudo, tais abordagens tendem a apresentar limitações na captura de relações não lineares e interações complexas presentes em séries de alta frequência com múltiplos padrões sazonais, limitação essa que se torna particularmente relevante em horizontes de previsão de curto prazo (TAYLOR *et al.*, 2006).

A segunda vertente baseia-se em algoritmos de aprendizado de máquina, entre os quais se destacam Random Forest, XGBoost e LightGBM. Esses métodos apresentam maior capacidade de modelar relações não lineares e interações entre variáveis, embora, em contrapartida, ofereçam menor interpretabilidade quanto aos componentes efetivamente capturados pelo modelo e maior dependência de etapas de engenharia de variáveis (VIVAS *et al.*, 2020).

Diante dessas limitações, a literatura recente tem proposto modelos híbridos que combinam a estrutura de decomposição dos modelos estatísticos com o poder preditivo de algoritmos de aprendizado de máquina, buscando ganhos simultâneos em acurácia e interpretabilidade.

Nesse cenário, destaca-se o modelo Prophet, proposto por Taylor e Letham (2018), o qual realiza a decomposição aditiva da série temporal em três componentes principais: (i) uma tendência por partes com pontos de mudança identificados automaticamente; (ii) sazonalidades periódicas representadas por séries de Fourier truncadas; (iii) efeitos explícitos de feriados e eventos especiais, incorporados como regressores adicionais.

Essa estrutura confere ao modelo capacidade de manter bom desempenho mesmo na presença de dados ausentes e mudanças de regime, razão pela qual tem sido amplamente empregado em aplicações de previsão de demanda elétrica (CHADALAVADA *et al.*, 2020). Entretanto, a hipótese implícita de que os resíduos resultantes dessa decomposição correspondem a ruído branco frequentemente não se sustenta em séries de demanda elétrica.

Dependências semanais persistentes, interações não lineares entre o nível de atividade econômica e os padrões de consumo, além de efeitos climáticos de alta frequência, como variações diárias de temperatura, tendem a deixar estruturas temporais remanescentes nos resíduos que os componentes determinísticos do Prophet não conseguem capturar (HYNDMAN; ATHANASOPOULOS, 2021). Essa estrutura residual representa, portanto, uma oportunidade adicional de melhoria preditiva, especialmente quando algoritmos de gradient boosting são treinados diretamente para modelar esses resíduos.

Motivado por esse cenário, objetivou-se com o presente trabalho desenvolver e avaliar uma arquitetura híbrida de duas camadas para a previsão da demanda elétrica no subsistema Sudeste/Centro-Oeste do Brasil.

MATERIAL E MÉTODOS

Fluxo metodológico

A Figura 1 apresenta o fluxo completo das etapas metodológicas, da aquisição dos dados na API (*Application Programming Interface*) pública do Operador Nacional do Sistema Elétrico (ONS), interface que permite o acesso automatizado às bases de dados do sistema, até a avaliação comparativa final. O trabalho organiza-se em três módulos principais: (i) pré-processamento da série; (ii) ajuste do modelo híbrido em duas camadas; (iii) avaliação fora da amostra com comparação estatística.

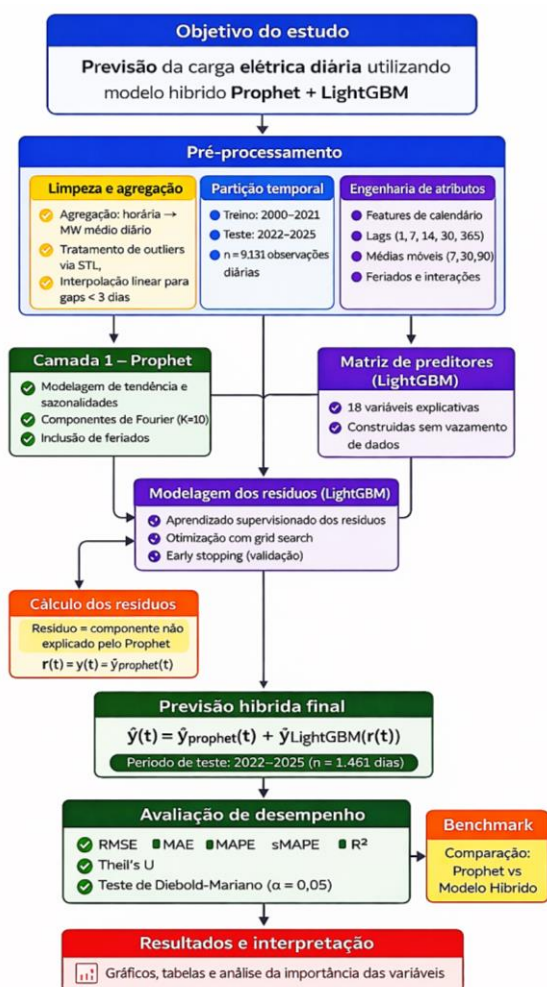


Figura 1. Fluxo metodológico: da coleta dos dados ONS à avaliação comparativa entre Prophet e o modelo híbrido Prophet+LightGBM.

Dados e preparação da série

Os dados correspondem à carga verificada horária do subsistema Sudeste/Centro-Oeste, obtidos via API pública do ONS (2025) para o período de 1 de janeiro de 2000 a 31 de dezembro de 2025. Após agregação para a média diária, a série totaliza 9.131 observações em MW médio.

O pré-processamento envolveu: (i) interpolação linear em lacunas de até três dias, método apropriado para falhas curtas em séries com alta autocorrelação temporal, uma vez que sua precisão decresce à medida que os blocos de dados ausentes aumentam, tornando-o inadequado para lacunas longas (DORJI *et al.*, 2024); (ii) detecção de outliers via decomposição STL (*Seasonal and Trend decomposition using Loess*), substituindo observações com resíduo superior a três desvios padrão pela mediana dos sete dias adjacentes, de modo a preservar a estrutura sazonal da série sem distorcer os parâmetros do modelo (CLEVELAND *et al.*, 1990).

A partição seguiu protocolo estrito: treinamento em 2001-2018, validação em 2019-2021 e teste em 2022-2025, sem contaminação retroativa de qualquer preditor. O período de teste foi escolhido por representar um regime de demanda pós-pandemia estruturalmente distinto do período de treinamento, com padrões de consumo alterados e episódios climáticos extremos que impõem maior exigência aos modelos de previsão (HEMBRAM *et al.*, 2023). Para o ajuste dos modelos finais, os conjuntos de treinamento e validação foram reunidos, cobrindo o período de 2001 a 2021.

Modelo Prophet: Camada 1

O Prophet decompõe a série temporal de forma aditiva, descrita abaixo:

$$\widehat{Carga}(t) = Tendência(t) + Sazonalidade(t) + Feriados(t) + \varepsilon(t) \quad (1)$$

Onde $\widehat{Carga}(t)$ é a demanda elétrica observada no instante t ; $Tendência(t)$ representa a tendência linear por partes, com pontos de mudança detectados automaticamente ao longo da série; $Sazonalidade(t)$ incorpora as sazonalidades periódicas via séries de Fourier truncadas, utilizando 3 pares de harmônicos para a sazonalidade semanal e 10 pares para a anual, onde cada par de harmônicos corresponde a uma função seno e cosseno de frequência múltipla da frequência base da sazonalidade, permitindo representar formas sazonais complexas de maneira parcimoniosa; o maior número de pares na sazonalidade anual reflete a maior variabilidade de forma ao longo dos 365 dias em comparação ao ciclo semanal de 7 dias; $Feriados(t)$ captura os efeitos de



feriados nacionais como regressores explícitos; e $\varepsilon(t)$ é o termo de erro, assumido como normalmente distribuído com média zero e variância constante, isto é, $\varepsilon(t) \sim N(0, \sigma^2)$ (TAYLOR; LETHAM, 2018).

A estrutura aditiva implica que os componentes contribuem de forma independente para o nível da série, o que pode ser uma restrição relevante quando a amplitude sazonal varia proporcionalmente à tendência.

O desempenho do Prophet depende da calibração de seus hiperparâmetros, sendo a Escala de Flexibilidade da Tendência aquele de maior impacto sobre a qualidade das previsões.

Esse hiperparâmetro controla a flexibilidade da tendência, regulando a intensidade com que os pontos de mudança alteram a inclinação local, valores maiores permitem ajustes mais bruscos, enquanto valores menores produzem tendências mais suaves. Tecnicamente, ele define a escala de uma distribuição de Laplace colocada como prior sobre as magnitudes das mudanças de tendência, atuando como um parâmetro de regularização.

Seu valor ótimo de 0,05, selecionado por minimização do RMSE no subconjunto de validação (2019-2021), coincide com o padrão recomendado pelos autores do modelo (TAYLOR; LETHAM, 2018), indicando que as mudanças estruturais da série do subsistema Sudeste/Centro-Oeste são suficientemente graduais para serem capturadas com regularização moderada, sem exigir flexibilidade excessiva que pudesse comprometer a generalização no período de teste. A previsão para cada dia do período de teste é calculada diretamente, sem realimentação recursiva.

LightGBM: Camada 2

O LightGBM é treinado sobre os resíduos da primeira camada, definidos como:

$$\text{Resíduo}(t) = \text{Carga}(t) - \widehat{\text{CargaProphet}}(t) \quad (2)$$

Calculados exclusivamente no conjunto de treino. O conjunto de preditores é composto por 18 variáveis, organizadas em quatro grupos.

O primeiro grupo reúne indicadores de calendário: (ano, mês, dia, dia da semana e dia do ano), que capturam padrões de diferentes granularidades; tendência linear, que aproxima o crescimento secular da demanda; e indicadores binários de feriado, véspera de feriado e pós-feriado, que capturam a queda e a recuperação do consumo em torno de datas especiais.

O segundo grupo é composto por defasagens da série observada: Defasagem de 1 dia, que captura a autocorrelação de curto prazo; Defasagem de 7 dias, que reflete o padrão do mesmo dia na semana anterior; Defasagens de 14 e 30 dias, que incorporam dependências de médio prazo; e Defasagem de 365 dias, que ancora o modelo no nível de demanda do mesmo período no ano anterior.

O terceiro grupo inclui médias móveis de 7, 30 e 90 dias, que suavizam flutuações de curto prazo e capturam a tendência recente em diferentes horizontes temporais.

O quarto grupo é composto pelo termo de interação entre a Defasagem de 7 dias e o Dia da Semana, que permite ao modelo capturar que o efeito do consumo semanal anterior varia conforme o dia da semana.

No LightGBM, as folhas correspondem aos nós terminais das árvores de decisão, cada folha representa uma região do espaço de preditores com uma previsão associada. Diferentemente de algoritmos tradicionais que crescem as árvores nível a nível, o LightGBM adota uma estratégia de crescimento folha a folha, selecionando a cada iteração a folha que oferece a maior redução da função de perda, o que permite capturar padrões mais complexos com maior eficiência computacional (KE *et al.*, 2017).

A Taxa de Aprendizado, por sua vez, controla a contribuição de cada árvore adicionada sequencialmente ao ensemble, funcionando como um fator de encolhimento que reduz o impacto de cada atualização individual, tornando o modelo mais conservador e robusto.

A seleção dos hiperparâmetros foi conduzida por busca em grade, uma vez que a acurácia do LightGBM em previsão de carga depende significativamente dos valores adotados, podendo o uso de valores padrão resultar em erros elevados (NGUYEN *et al.*, 2024).

Foram explorados o Número de Folhas em {31, 63, 127}, partindo do valor padrão do algoritmo (31) e dobrando progressivamente a capacidade de particionamento, com valores maiores aumentando a acurácia no treino ao custo de maior risco de *overfitting*, fenômeno em que o modelo se ajusta excessivamente aos dados de treinamento, perdendo capacidade de generalização para novos dados.

Adicionalmente, explorou-se a Taxa de Aprendizado em {0,05; 0,01}, valores abaixo do padrão de 0,1, pois taxas menores tendem a produzir melhor generalização, embora exijam mais iterações para convergência (KE *et al.*, 2017), e o Mínimo de



Amostras por Folha em {20, 50}, onde o valor padrão de 20 foi testado junto a um valor mais restritivo de 50, avaliando o efeito de uma regularização mais conservadora sobre o risco de overfitting (KE *et al.*, 2017).

A busca foi monitorada com Parada Antecipada de 50 iterações no subconjunto de validação, mecanismo que interrompe o treinamento quando a métrica de avaliação não apresenta melhora por um número determinado de rodadas consecutivas, prevenindo *overfitting* e reduzindo o custo computacional do treinamento.

A previsão final do modelo híbrido é obtida pela soma das duas camadas:

$$\widehat{Carga}(t) = Carga\widehat{Prophet}(t) + CargaLightGBM(\widehat{Resíduo}(t)) \quad (3)$$

Onde $\widehat{Carga}(t)$ é a previsão final do modelo híbrido no instante t ; $\widehat{CargaProphet}(t)$ é a componente determinística estimada pelo Prophet na primeira camada; e $CargaLightGBM(\widehat{Resíduo}(t))$ é a correção estimada pelo LightGBM sobre os resíduos da primeira camada, capturando a estrutura não linear remanescente não explicada pelo Prophet.

Métricas de desempenho e teste de Diebold-Mariano

O desempenho é medido no período de teste (2022-2025; $n = 1.461$ dias) pelas métricas RMSE, MAE, MAPE, sMAPE, R^2 e Theil's U, este último normaliza o RMSE pelo erro de uma previsão ingênua (random walk), com $U < 1$ indicando superioridade sobre a referência. As métricas são definidas pelas equações a seguir:

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (Y_i - \hat{Y}_i)^2} \quad (4)$$

$$MAE = \frac{1}{n} \sum_{i=1}^n |Y_i - \hat{Y}_i| \quad (5)$$

$$MAPE = \frac{100}{n} \sum_{i=1}^n \left| \frac{Y_i - \hat{Y}_i}{Y_i} \right| \quad (6)$$

$$sMAPE = \frac{200}{n} \sum_{i=1}^n \frac{|Y_i - \hat{Y}_i|}{|Y_i| + |\hat{Y}_i|} \quad (7)$$

$$R^2 = 1 - \frac{\sum_{i=1}^n (Y_i - \hat{Y}_i)^2}{\sum_{i=1}^n (Y_i - \bar{Y})^2} \quad (8)$$

$$MBE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i) \quad (9)$$

$$U = \frac{\sqrt{\frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n}}}{\sqrt{\frac{\sum_{i=1}^n (y_i - y_{i-1})^2}{n}}} \quad (10)$$

Onde Y_i denota os valores observados, $\hat{Y}_i$ os valores previstos e n o número de observações (HYNDMAN; ATHANASOPOULOS, 2021), $\bar{Y}$ é a média dos valores observados. O MBE mede o viés sistemático, com valores negativos indicando subestimação. O sMAPE é limitado ao intervalo [0%, 200%] e trata sobre e subestimações de forma simétrica (MAKRIDAKIS, 1993). O Theil's U normaliza o RMSE pelo erro da previsão ingênua (passeio aleatório), com $U < 1$ indicando superioridade sobre essa referência.

A significância estatística das diferenças foi testada pelo teste de Diebold-Mariano com função de perda quadrática e correção de Newey-West para autocorrelação nos diferenciais de erro (DIEBOLD; MARIANO, 1995).

RESULTADOS E DISCUSSÃO

Série histórica e contexto

A Figura 2 apresenta os 25 anos de demanda diária no subsistema Sudeste/Centro-Oeste. Três regimes são claramente distintos: crescimento contínuo entre 2000 e 2014; estabilidade e ligeira queda entre 2015 e 2020, período que inclui a crise econômica de 2015-2016 e a pandemia de 2020; e retomada a partir de 2021, com patamares que superam os máximos históricos nos verões de 2023 e 2024.

A sazonalidade intra-semanal é pronunciada em todo o período, com reduções de 15-20% nos fins de semana em relação aos dias úteis (Soares e Medeiros, 2008).

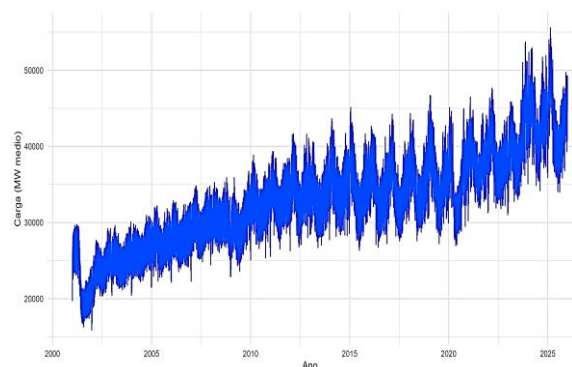


Figura 2. Demanda diária média (MW médio) – Subsistema Sudeste/Centro-Oeste (2000–2025).

Fonte: Adaptado de ONS (2025).

Decomposição do Prophet

As Figuras 3, 4 e 5 mostram os componentes estimados pelo Prophet. A tendência (Figura 3) revela crescimento até 2014, seguida de estabilidade em 2015-2020 e retomada pós-pandemia.

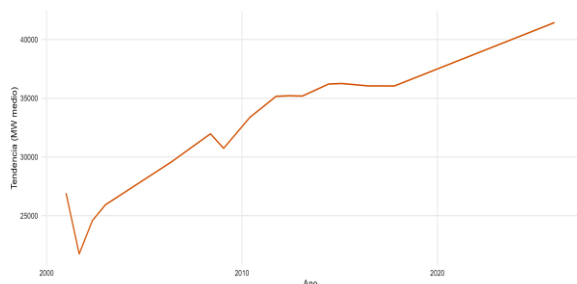


Figura 3. Tendência estimada pelo Prophet – subsistema Sudeste/Centro-Oeste (2000-2025).

A sazonalidade semanal (Figura 4) exibe efeito de domingo de até -4.500 MW em relação à média dos dias úteis.

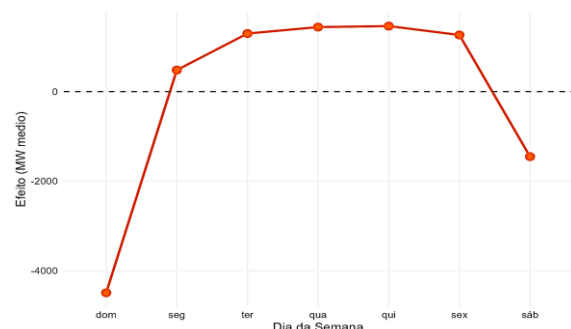


Figura 4. Sazonalidade semanal estimada pelo Prophet (efeito em MW médio em relação à segunda-feira).

A sazonalidade anual (Figura 5) oscila em torno de ± 2.200 MW, com picos de verão austral (climatização) e vales de inverno. No período de teste, contudo, o R^2 do Prophet isolado é de apenas 0,66 e o viés médio (MBE) de -2.659 MW, revelando subestimação sistemática que se concentra nos meses de verão.

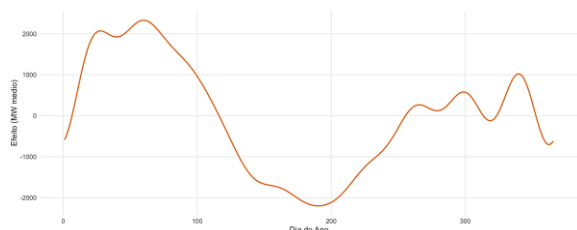


Figura 5. Sazonalidade anual estimada pelo Prophet (efeito em MW médio ao longo do ano).

Mecanismo em duas camadas

As Figuras 6, 7 e 8 ilustram o funcionamento do modelo híbrido em duas camadas no período de teste.

A Figura 6 mostra que o Prophet acompanha a tendência geral da série, mas opera consistentemente abaixo dos valores reais, sobretudo nos verões de 2023-2024, acumulando um viés de subestimação que o modelo determinístico não consegue corrigir por si só.

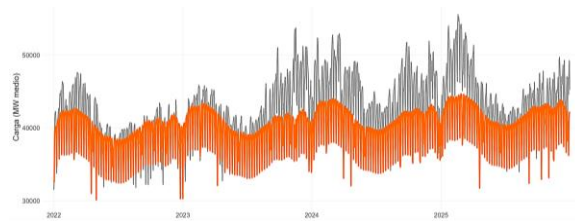


Figura 6. Previsão do Prophet vs. demanda observada no período de teste (2022-2025).

A Figura 7 evidencia que os resíduos resultantes dessa primeira camada não se comportam como ruído branco: apresentam autocorrelação pronunciada e sazonalidade residual com amplitude de até 13.000 MW, estrutura que o LightGBM aprende a rastrear sistematicamente.

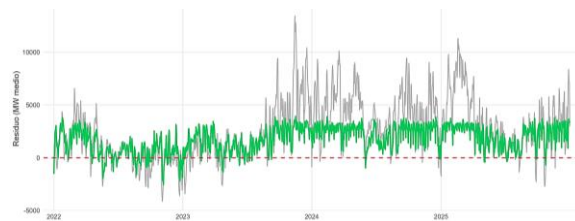


Figura 7. Resíduos do Prophet (cinza) e correção estimada pelo LightGBM (verde) no período de teste (2022-2025).

A Figura 8 mostra o resultado da soma das duas camadas, a previsão híbrida adere à demanda real com nitidamente maior fidelidade, corrigindo tanto o viés sistemático quanto os desvios episódicos de maior magnitude.

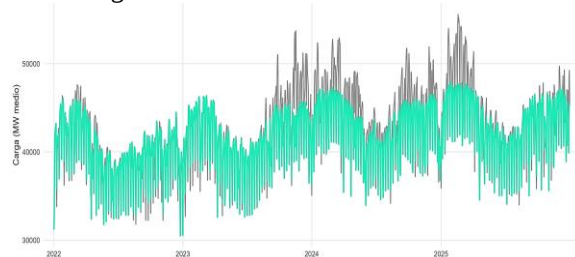


Figura 8. Previsão do modelo híbrido Prophet+LightGBM vs. demanda observada no período de teste (2022-2025).

Desempenho comparativo

A Tabela 1 sumariza as métricas fora da amostra. O modelo híbrido supera o Prophet isolado em todas as dimensões: o RMSE cai de 3.816 para 2.211 MW (-42,1%), o MAE de 2.944 para 1.558 MW (-47,1%) e o MAPE de 6,55% para 3,49% (-46,7%). O sMAPE

confirma a mesma tendência, caindo de 6,89% para 3,57% (–48,2%), resultado consistente com o MAPE e que reforça que a melhoria do híbrido é robusta a diferentes formas de mensurar o erro relativo.

O Theil's U passa de 1,24 para 0,72, cruzando o limiar unitário e confirmando superioridade sobre a previsão ingênua. O MBE cai de –2.659 para –886 MW, indicando redução expressiva do viés sistemático de subestimação identificado no Prophet isolado.

Tabela 1. Métricas de desempenho fora da amostra – período de teste (2022–2025).

Modelo	Métrica	Valor
Prophet	RMSE	3.816,11
	MAE	2.943,63
	MAPE (%)	6,55
	sMAPE (%)	6,89
	R ²	0,662
	MBE	-2.659,30
	Theil's U	1,236
Prophet+LGB	RMSE	2.210,53
	MAE	1.558,05
	MAPE (%)	3,49
	sMAPE (%)	3,57
	R ²	0,810
	MBE	-885,99
	Theil's U	0,716

RMSE: raiz do erro quadrático médio; MAE: erro absoluto médio; MAPE (%): erro percentual absoluto médio; sMAPE (%): erro percentual absoluto médio simétrico; R²: coeficiente de determinação; MBE: erro médio de viés (valores negativos indicam subestimação); Theil's U: razão entre o erro do modelo e o de uma previsão ingênua ($U < 1$ indica melhor desempenho).

A Figura 9 apresenta a comparação direta entre os valores observados e previstos pelos dois modelos no período de teste. O híbrido exibe maior aderência à linha de 45°, evidenciando menor dispersão e ausência do viés de subestimação sistemática presente no Prophet isolado.

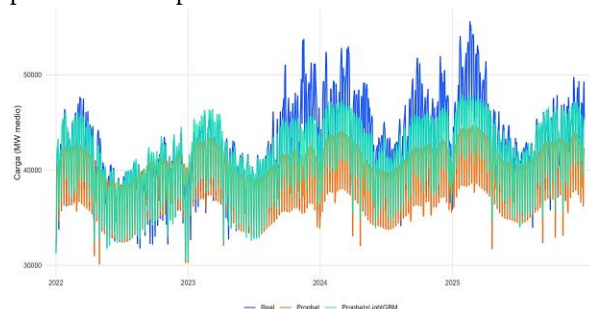


Figura 9. Valores observados vs. previstos - Prophet (esquerda) e Prophet+LightGBM (direita). Linha tracejada = 45°. Período de teste 2022–2025.

A Tabela 2 apresenta os resultados do teste de Diebold-Mariano. A estatística $DM = 27,47$ para a comparação Prophet vs. Prophet+LightGBM implica rejeição irrefutável da hipótese de igual acurácia ($p < 0,001$), resultado robusto a qualquer forma razoável de autocorrelação dos diferenciais de erro.

A magnitude da estatística, muito acima dos limiares convencionais, confirma que a melhoria de 42% no RMSE não é artefato amostral, mas ganho estrutural obtido pela correção supervisionada dos resíduos.

Tabela 2. Teste de Diebold-Mariano - perda quadrática, $\alpha = 0,05$.

Comparação	DM	p-valor	Conclusão
Prophet vs. Prophet+LGB	27,47	< 0,001	Híbrido sig. melhor

A Figura 10 apresenta a comparação visual das três principais métricas de erro entre os dois modelos. A redução percentual exibida confirma de forma imediata os valores da Tabela 1: o modelo híbrido supera o Prophet isolado em todas as dimensões avaliadas, com destaque para a queda de 46,7% no MAPE e de 42,1% no RMSE, reforçando que o ganho obtido pela correção supervisionada dos resíduos é consistente tanto em termos absolutos quanto relativos.

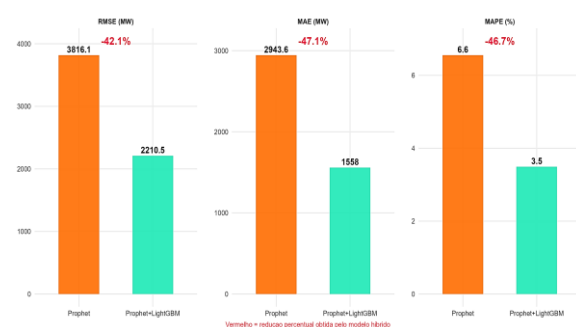


Figura 10. Comparação das métricas de desempenho entre Prophet e Prophet+LightGBM no período de teste (2022–2025).

Padrão temporal dos erros

A Figura 11 mostra o MAE médio por mês. Ambos os modelos concentram os maiores erros no verão (novembro-março), mas a diferença relativa é maior no inverno: em julho, o Prophet erra em média 1.388 MW e o híbrido apenas 580 MW, redução de 58%. Esse padrão sugere que o LightGBM aprende com maior facilidade os resíduos do inverno, quando o consumo segue ciclos semanais mais regulares. A variabilidade residual nos meses de verão ainda

impõe um teto ao ganho do híbrido, motivando a inclusão de temperatura máxima diária como extensão.

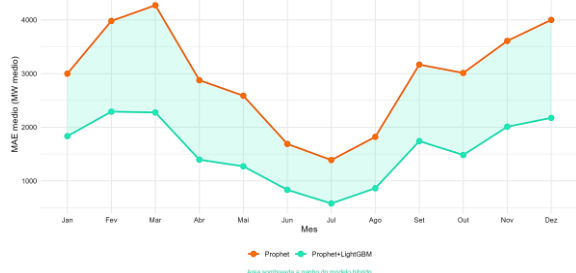


Figura 11. MAE médio por mês – Prophet (laranja) e Prophet+LightGBM (turquesa). Área sombreada = ganho absoluto do modelo híbrido.

As Figuras 12, 13 e 14 analisam a estrutura dos resíduos por meio da função de autocorrelação e do teste de Ljung-Box. As Figuras 15 e 16 comparam os valores observados e previstos pelos dois modelos, e as Figuras 17, 18 e 19 complementam a análise com o gráfico violin e os diagramas Q-Q dos resíduos.

A Figura 12 apresenta a função de autocorrelação dos resíduos do Prophet no período de teste. A presença de autocorrelação estatisticamente significativa em múltiplos defasamentos confirma que os resíduos da primeira camada não constituem ruído branco: existe estrutura temporal remanescente que o modelo determinístico não consegue capturar por meio de seus componentes de tendência, sazonalidade e feriados. É precisamente essa estrutura que o LightGBM aprende a modelar na segunda camada do modelo híbrido.

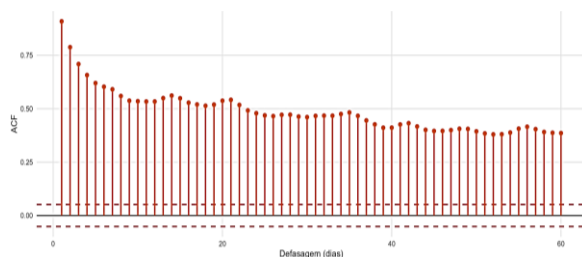


Figura 12. ACF dos resíduos do Prophet. Período de teste (2022-2025).

A Figura 13 apresenta a ACF dos resíduos do modelo híbrido. Em contraste com a Figura 12, a autocorrelação é expressivamente reduzida na maioria dos defasamentos, indicando que o LightGBM absorveu grande parte da estrutura temporal que o Prophet deixou nos resíduos. O resultado aproxima os resíduos do híbrido de um processo não autocorrelacionado, o que representa uma melhoria qualitativa relevante além das métricas de erro reportadas na Tabela 1.

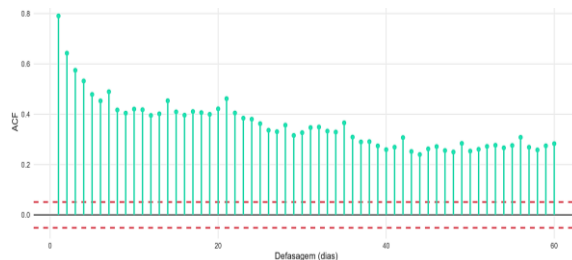


Figura 13. ACF dos resíduos do modelo híbrido Prophet+LightGBM. Período de teste (2022-2025).

A Figura 14 apresenta os p-valores do teste de Ljung-Box para os resíduos dos dois modelos em diferentes defasamentos. O Prophet rejeita sistematicamente a hipótese de ruído branco, com p-valores próximos de zero na quase totalidade dos defasamentos testados, corroborando formalmente a estrutura autocorrelacionada evidenciada na Figura 12.

O modelo híbrido, por sua vez, apresenta p-valores substancialmente maiores, indicando que os resíduos da segunda camada se aproximam de um processo não autocorrelacionado. Esse resultado confere respaldo estatístico formal à melhoria qualitativa observada nas ACFs e reforça a validade da arquitetura híbrida proposta.

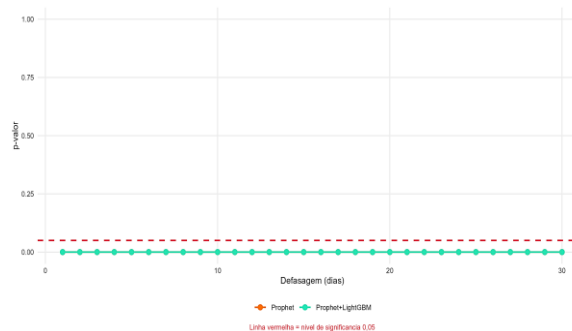


Figura 14. Teste de Ljung-Box para os resíduos do Prophet e do modelo híbrido Prophet+LightGBM. Período de teste (2022-2025).

A Figura 15 apresenta o diagrama de dispersão entre os valores observados e previstos pelo Prophet. A nuvem de pontos exibe desvio sistemático abaixo da linha de referência perfeita (45°), especialmente nos patamares mais elevados de demanda, confirmando o viés de subestimação identificado nas métricas da Tabela 1. O ajuste linear (linha vermelha) reforça essa tendência, indicando que o Prophet subestima progressivamente à medida que a demanda cresce.

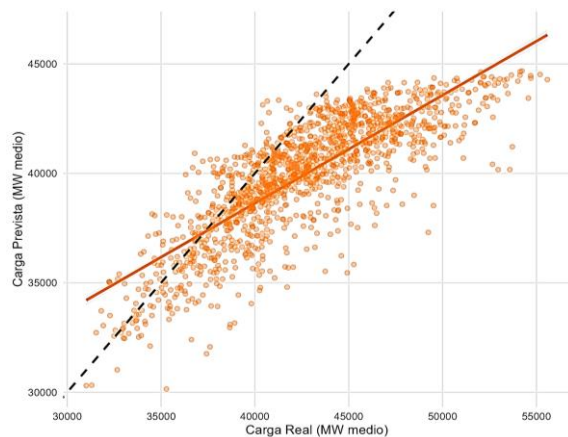


Figura 15. Dispersão entre valores observados e previstos pelo Prophet. Linha tracejada = referência perfeita (45°). Período de teste (2022-2025).

A Figura 16 apresenta o diagrama de dispersão do modelo híbrido. Em contraste com a Figura 15, a nuvem de pontos se concentra consideravelmente mais próxima da linha de referência perfeita, e o ajuste linear se aproxima da diagonal de 45°, indicando ausência do viés sistemático de subestimação presente no Prophet isolado.

A melhoria é mais pronunciada nos patamares elevados de demanda, exatamente onde o Prophet mais errava, confirmando que o LightGBM corrige de forma eficaz a estrutura não linear remanescente nos resíduos da primeira camada.

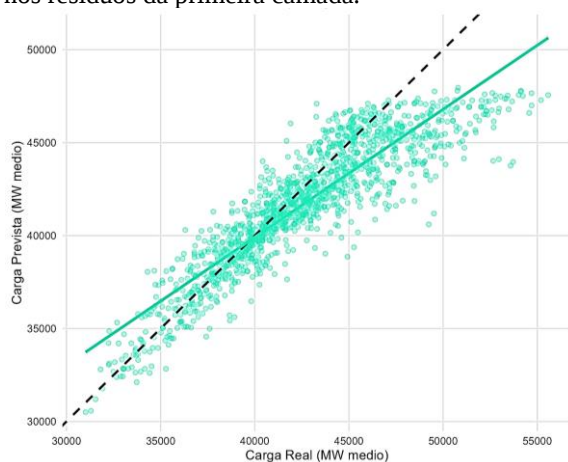


Figura 16. Dispersão entre valores observados e previstos pelo modelo híbrido Prophet+LightGBM. Linha tracejada = referência perfeita (45°). Período de teste (2022-2025).

A Figura 17 apresenta o gráfico violim dos resíduos dos dois modelos. A distribuição do Prophet é mais larga e assimétrica, com cauda positiva pronunciada associada aos episódios de subestimação nos verões de 2023-2024. O híbrido apresenta distribuição notavelmente mais concentrada em torno de zero, com menor dispersão e assimetria reduzida, indicando que o LightGBM não apenas reduz a

magnitude dos erros mas também melhora a simetria da distribuição residual, aproximando-a das hipóteses desejáveis para um bom modelo preditivo.

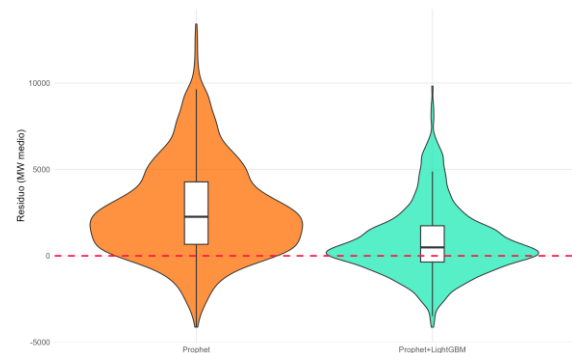


Figura 17. Gráfico violim dos resíduos do Prophet e do modelo híbrido Prophet+LightGBM. Período de teste (2022-2025).

A Figura 18 apresenta o gráfico Q-Q dos resíduos do Prophet. Os desvios sistemáticos nas caudas, tanto inferior quanto superior, indicam que a distribuição dos resíduos se afasta da normalidade, com presença de valores extremos mais frequentes do que o esperado. Esse padrão é consistente com a estrutura não linear remanescente identificada nas Figuras 12 e 14, e reflete a incapacidade do modelo determinístico de capturar os picos de demanda associados aos episódios de calor intenso nos verões de 2023-2024.

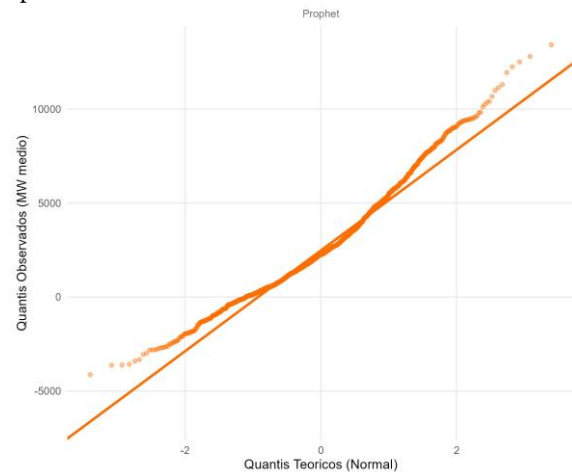


Figura 18. Gráfico Q-Q dos resíduos do Prophet. Período de teste (2022-2025).

A Figura 19 apresenta o gráfico Q-Q dos resíduos do modelo híbrido. Em comparação com a Figura 18, a aderência à linha teórica é consideravelmente maior, especialmente na região central da distribuição, confirmando a redução da estrutura não linear nos resíduos após a correção pelo LightGBM. Persiste leve assimetria positiva nos extremos superiores, associada aos episódios de demanda excepcionalmente elevada nos verões de 2023-2024, o que sugere que a incorporação de variáveis

climáticas como temperatura máxima diária poderia reduzir ainda mais esses desvios residuais.

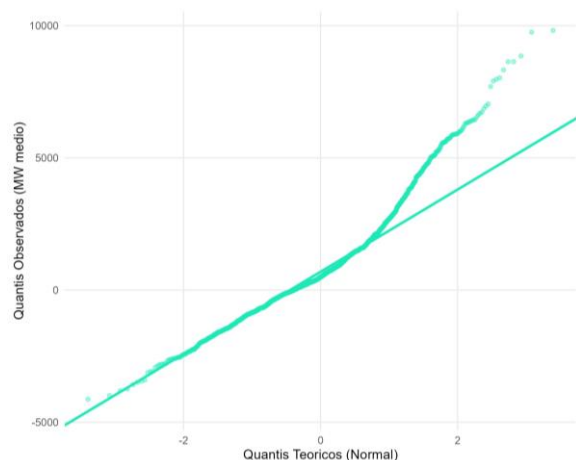


Figura 19. Gráfico Q-Q dos resíduos do modelo híbrido Prophet+LightGBM. Período de teste (2022-2025).

Importância das variáveis

A Tabela 3 lista as dez variáveis de maior ganho no LightGBM de correção. A liderança da Média Móvel de 7 dias (21,38%) e da Defasagem de 1 dia (18,84%) revela que grande parte dos resíduos do Prophet é explicada pelo nível recente da série, ponto cego de um modelo que ajusta a tendência globalmente. A Tendência Linear em terceiro lugar (13,16%) confirma que o Prophet acumula erro nos períodos de aceleração de demanda, que o LightGBM corrige localmente. A própria Previsão do Prophet em quinto lugar (8,07%) indica que o LightGBM captura sazonalidades residuais finas que a série de Fourier não resolve.

Tabela 3. Top 10 variáveis por Ganho – LightGBM de correção de resíduos.

Rank	Variável	Ganho (%)
1	Média Móvel 7 dias	21,38%
2	Defasagem 1 dia	18,84%
3	Tendência Linear	13,16%
4	Dia do Ano	10,23%
5	Previsão Prophet	8,07%
6	Média Móvel 30 dias	4,20%
7	Defasagem 365 dias	3,35%
8	Ano	3,17%
9	Defasagem 7 dias × Dia da Semana	3,06%
10	Média Móvel 90 dias	2,84%

CONCLUSÃO

O modelo híbrido Prophet+LightGBM apresentou ganhos substanciais na previsão da demanda elétrica do subsistema Sudeste/Centro-Oeste, com reduções de 42,1% no RMSE, 47,1% no MAE e 46,7% no MAPE em relação ao Prophet isolado, com superioridade estatisticamente confirmada pelo teste de Diebold-Mariano.

Os resultados demonstram que os resíduos do Prophet contêm estrutura não linear explorável, cuja modelagem via *boosting* permite corrigir sistematicamente erros associados à dinâmica de curto prazo e à tendência.

Apesar dos avanços, a persistência de erros nos picos de verão evidencia a necessidade de incorporação de variáveis climáticas, indicando um caminho direto para aprimoramento do modelo e ampliação de sua aplicabilidade em sistemas reais.

AGRADECIMENTOS

À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) e ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), pela concessão de bolsas de pesquisa.

À Universidade Estadual da Paraíba (UEPB), pelo suporte institucional.

REFERÊNCIAS

- ARIAS, N. B. et al. An overview of Brazil's electricity market: planning, dispatch models, pricing, and modernization. *Utilities Policy*, v. 83, p. 101621, 2023.
- DIEBOLD, F. X.; MARIANO, R. S. Comparing predictive accuracy. *Journal of Business & Economic Statistics*, v. 13, n. 3, p. 253–263, 1995.
- DUDEK, G. A comprehensive study of random forest for short-term load forecasting. *Energies*, v. 15, n. 20, p. 7547, 2022.
- LÚCIO, L. C. et al. Electric power load in Brazil: view on the long-term forecasting models. *Production*, v. 28, p. e20170122, 2018.
- ONS – OPERADOR NACIONAL DO SISTEMA ELÉTRICO. Curva de carga verificada por subsistema. Disponível em: <https://dados.ons.org.br>. Acesso em: mar. 2025.
- SOARES, L. J.; MEDEIROS, M. C. Modeling and forecasting short-term electricity load: a comparison of methods with an application to Brazilian data. *International Journal of Forecasting*, v. 24, n. 4, p. 630–644, 2008.



TAYLOR, S. J.; LETHAM, B. Forecasting at scale. *The American Statistician*, v. 72, n. 1, p. 37–45, 2018.

VELASQUEZ, C. E.; ZOCATELLI, M.; ESTANISLAU, F. B. G. L. Analysis of time series models for Brazilian electricity demand forecasting. *Energy*, v. 247, p. 123483, 2022.

VIVAS, E.; ALLENDE-CID, H.; SALAS, R. A systematic review of statistical and machine learning methods for electrical power forecasting with reported MAPE score. *Entropy*, v. 22, n. 12, p. 1412, 2020.

TAYLOR, J. W.; DE MENEZES, L. M.; McSHARRY, P. E. A comparison of univariate methods for forecasting electricity demand up to a day ahead. *International Journal of Forecasting*, v. 22, n. 1, p. 1–16, 2006.

CHADALAVADA, R. J.; RAGHAVENDRA, S.; REKHA, V. Electricity requirement prediction using time series and Facebook's PROPHET. *Indian Journal of Science and Technology*, v. 13, n. 47, p. 4631–4645, 2020.

DORJI, K. et al. Electricity load forecasting using hybrid datasets with linear interpolation and synthetic data. *Engineering, Technology & Applied Science Research*, v. 14, n. 6, p. 17931–17938, 2024.

CLEVELAND, R. B. et al. STL: A seasonal-trend decomposition procedure based on loess. *Journal of Official Statistics*, v. 6, n. 1, p. 3–73, 1990.

HEMBRAM, P. et al. Probabilistic forecasting of electricity demand incorporating mobility data. *Applied Sciences*, v. 13, n. 11, p. 6520, 2023.

CHEN, K. et al. Short-term load forecasting in power systems based on the Prophet–BO–XGBoost model. *Energies*, v. 18, n. 2, p. 227, 2025.

NGUYEN, K. T.; TRAN, T. N.; NGUYEN, H. T. Research on the influence of hyperparameters on the LightGBM model in load forecasting. *Engineering, Technology & Applied Science Research*, v. 14, n. 5, p. 17005–17010, 2024.

HYNDMAN, R. J.; ATHANASOPOULOS, G. *Forecasting: Principles and Practice*. 3. ed. Melbourne: OTexts, 2021. Disponível em: <https://otexts.com/fpp3>. Acesso em: mar. 2025.

KE, G. et al. LightGBM: a highly efficient gradient boosting decision tree. *Advances in Neural Information Processing Systems*, v. 30, p. 3146–3154, 2017.

MAKRIDAKIS, S. Accuracy measures: theoretical and practical concerns. *International Journal of Forecasting*, v. 9, n. 4, p. 527–529, 1993.