EFICÁCIA DE MODELOS DE SÉRIES TEMPORAIS NA PREVISÃO DE CONSUMO PARA ENERGIA FOTOVOLTAICA: SARIMA e PROPHET

Henrique Teixeira Silva[1] (IC), Rodrigo Aparecido da Silva Braga (PQ)[1]
[1]Universidade Federal de Itajubá

Palavras-chave: Energia fotovoltaica. Previsão de consumo. SARIMA. Prophet. Séries temporais

Introdução

Este estudo compara a eficácia dos modelos SARIMA e Prophet na previsão do consumo de energia elétrica para sistemas fotovoltaicos, com foco em cenários de dados históricos limitados. O objetivo principal é otimizar a operação desses sistemas, melhorar a eficiência da rede e reduzir custos operacionais. A pesquisa se justifica pelo crescimento exponencial do setor fotovoltaico no Brasil, que ultrapassou 30 GW de potência instalada em 2023, representando mais de 10% da matriz elétrica do país (ABSOLAR, 2024). Os desafios de previsão enfrentados incluem a variabilidade na demanda, a intermitência da geração solar e a escassez de dados históricos em muitos casos (WEI et al., 2019). O método envolve a análise de dados de várias empresas, incluindo tratamento de dados, análise de características como sazonalidade e estacionaridade, e otimização de parâmetros dos modelos. Este estudo visa preencher uma lacuna importante na literatura, focando na comparação e otimização de modelos de previsão em contextos de dados escassos, oferecendo insights valiosos para empresas de energia fotovoltaica aprimorarem suas estratégias de previsão de consumo.

Metodologia

A metodologia segue um fluxo de procedimentos detalhado: 1) Aquisição e tratamento de dados, incluindo a exclusão de instalações com menos de um a dois anos de dados; 2) Análise das características dos dados, utilizando a função seasonal_decompose da biblioteca StatsModels para analisar sazonalidade e o teste Augmented Dickey-Fuller (ADF) para estacionaridade (RAMENAH et al., 2018); 3) Extração de parâmetros usando a função auto_arima da biblioteca pmdarima; 4) Teste inicial de performance usando divisão treino/teste (90% treino, 10% teste); 5) Medição de performance usando métricas como RMSE, MAE e MAPE, com considerações sobre suas aplicabilidades em diferentes cenários (CHAI; DRAXLER, 2014; DE MYTTENAERE et al., 2016); 6) Otimização iterativa dos parâmetros para encontrar um modelo satisfatório; 7) Criação de modelo final com todos os dados para testar performance com dados futuros; 8) Salvamento do modelo usando bibliotecas como Pickle ou Prophet para integração em outras bases de código. A metodologia também inclui técnicas de data augmentation para lidar com a escassez de dados, embora com cautela devido ao risco de sobreajuste (IWANA; UCHIDA, 2021).

Resultados e discussão

Os resultados da pesquisa são apresentados em tabelas comparativas, mostrando o desempenho dos modelos SARIMA e Prophet para diferentes instalações, com e sem aumento de dados. As métricas de erro utilizadas incluem RMSE (Root Mean Squared Error), MAPE (Mean Absolute Percentage Error) e MAE (Mean Absolute Error).

Para o modelo SARIMA sem aumento de dados, os resultados foram:

SARIMA			
Instalação	RMSE	MAPE (%)	MAE
Empresa 1	838.51	42.28	692.01
Empresa 2	705.83	9.13	589.66
Empresa 3	140.61	Erro	112.09
Empresa 4	111.94	10.44	88.95
Empresa 5	433.01	15.54	351.91

Tabela 1. Métricas de erro para os modelos criados com SARIMA.

Após o aumento de dados, os resultados para o SARIMA foram:

SARIMA - Augmented			
Instalação	RMSE	MAPE (%)	MAE
Empresa 1	1073.57	54.82	822.93
Empresa 2	849.05	10.81	661.72
Empresa 3	188.91	Erro	112.09
Empresa 4	111.94	11.62	139.63
Empresa 5	138.72	22.77	115.63

Tabela 2. Métricas de erro para os modelos criados com SARIMA para a base de dados aumentada.

Para o modelo Prophet sem aumento de dados, os resultados foram:

Prophet			
Instalação	RMSE	MAPE (%)	MAE
Empresa 1	170.77	5.58	170.77
Empresa 2	661.61	8.75	661.61
Empresa 3	144.57	Erro	144.57
Empresa 4	257.24	33.84	257.24
Empresa 5	310.13	11.57	310.13

Tabela 3. Métricas de erro para os modelos criados com Prophet.

E após o aumento de dados para o Prophet:

Prophet - Augmented			
Instalação	RMSE	MAPE (%)	MAE
Empresa 1	486.45	12.66	486.45
Empresa 2	320.02	4.23	320.02
Empresa 3	115.83	Erro	115.83
Empresa 4	79.71	10.48	79.71
Empresa 5	159.97	7.54	159.97

Tabela 4. Métricas de erro para os modelos criados com Prophet para a base de dados aumentada.

Nota-se que o MAPE da instalação Empresa 3 consta como "Erro" em todas as tabelas devido à presença de zeros nos dados, o que leva o MAPE ao infinito ou a valores inconsistentes com as outras métricas.
A análise dos resultados revela que o Prophet geralmente se sai melhor em bases de dados com forte sazonalidade, como observado na instalação Empresa 1. Por outro lado, o SARIMA mostrou-se mais eficaz em padrões de consumo menos definidos, como na instalação Empresa 4.
O aumento de dados resultou em performances variadas. Em alguns casos, houve melhoria nas previsões futuras, mas piora nas métricas de erro imediatas, indicando potencial sobreajuste. Este fenômeno foi particularmente notável na instalação Empresa 1 para o modelo SARIMA, onde o RMSE aumentou de 838.51 para 1073.57 após o aumento de dados.
A análise da instalação Empresa 2 revelou que tanto o SARIMA quanto o Prophet podem ser eficazes em dados estacionários sem fortes tendências. O Prophet, no entanto, demonstrou maior confiabilidade na previsão de tendências futuras para esta instalação.
Uma comparação dos melhores resultados de cada método mostra:

Melhores Resultados			
Instalação	RMSE	MAPE (%)	MAE
SARIMA - Empresa 2	705.83	9.13	589.66
SARIMA Aug. - Empresa 2/4*	849.05	10.81	661.72
	111.94	11.62	139.63
Prophet - Empresa 1	170.77	5.58	170.77
Prophet Aug. - Empresa 2	320.02	4.23	320.02

*As duas instalações possuem um erro percentual bem próximo
Tabela 5. Melhores resultados de cada método

Estes resultados destacam a complexidade da previsão de consumo de energia em instalações fotovoltaicas e a importância de considerar as características específicas de cada conjunto de dados ao selecionar e otimizar modelos de previsão.
Além disso, uma análise mais detalhada foi realizada para a instalação Empresa 2, que apresentou resultados interessantes tanto para o SARIMA quanto para o Prophet. A Figura 14 mostra uma comparação das previsões SARIMA com e sem aumento de dados para esta instalação:

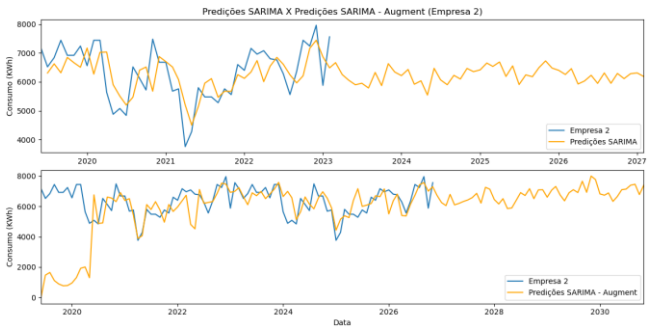

Figura 14. Comparação das previsões SARIMA com e sem aumento de dados

Esta figura demonstra que, para a instalação Empresa 2, o modelo SARIMA apresentou um desempenho relativamente bom, mas com sinais de possível sobreajuste, especialmente após o aumento de dados. As previsões futuras não parecem refletir completamente o comportamento dos dados reais, indicando limitações do modelo neste caso específico.
Por outro lado, a Figura 15 apresenta uma comparação similar para o modelo Prophet:

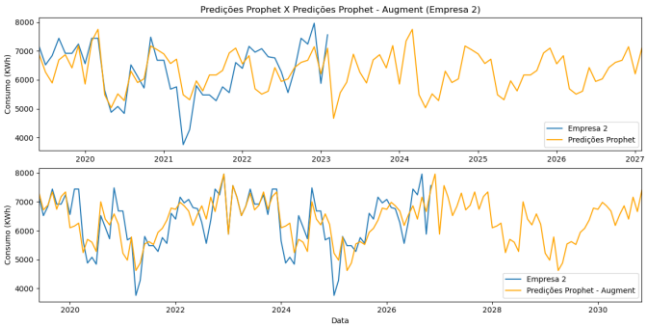

Figura 15. Comparação das previsões Prophet com e sem aumento de dados

A Figura 15 revela que o Prophet demonstrou uma melhor capacidade de capturar as tendências gerais dos dados da instalação Empresa 2, tanto com quanto sem o aumento de dados. O modelo parece ser mais robusto às variações nos dados e apresenta previsões futuras mais alinhadas com o comportamento histórico observado.
Estas análises visuais complementam as métricas quantitativas apresentadas anteriormente, oferecendo insights adicionais sobre o desempenho relativo dos modelos SARIMA e Prophet em diferentes cenários de dados. Elas reforçam a conclusão de que a escolha do modelo mais apropriado depende fortemente das características específicas dos dados de cada instalação.

Conclusões

O estudo conclui que o desempenho dos modelos SARIMA e Prophet varia significativamente conforme as características dos dados analisados. O Prophet demonstrou eficiência tanto em dados com forte sazonalidade quanto em dados estacionários, enquanto o SARIMA mostrou-se mais adequado para padrões de consumo menos definidos. A ampliação artificial da base de dados resultou em performances variadas, com risco de sobreajuste em alguns casos, destacando o desafio de realizar previsões precisas com dados históricos limitados. O estudo ressalta a importância da análise individual de cada instalação devido à grande variabilidade nos padrões de consumo entre diferentes clientes. As implicações práticas incluem a otimização da alocação de recursos, melhor planejamento da produção de energia e gestão mais eficiente de picos de demanda no setor fotovoltaico. Para pesquisas futuras, sugere-se a incorporação de variáveis exógenas como condições climáticas, o desenvolvimento de modelos híbridos combinando SARIMA e Prophet com técnicas de aprendizado de máquina, e a extensão da análise para previsões de longo prazo considerando mudanças tecnológicas e regulatórias no setor. Adicionalmente, propõe-se a exploração de técnicas de ensemble, combinando os resultados do SARIMA e do Prophet para potencialmente melhorar a precisão geral das previsões (TAYLOR; LETHAM, 2017; CHATURVEDI et al., 2022).

Agradecimentos

Gostaria de expressar minha profunda gratidão à Universidade Federal de Itajubá por fornecer o ambiente e os recursos necessários para a realização desta pesquisa. Agradeço também ao CNPq pelo apoio financeiro que permitiu a realização deste trabalho. Um agradecimento especial ao meu orientador, Rodrigo Aparecido da Silva Braga, por sua orientação inestimável, paciência e apoio contínuo ao longo deste projeto. Sua experiência e conhecimento foram fundamentais para o sucesso deste trabalho

Referências

ABSOLAR. Retrospectiva esg 2023: o ano da virada para a energia solar no brasil - absolar. 2024. Disponível em: https://www.absolar.org.br/noticia/retrospectiva-esg-2023-o-ano-da-virada-para-a-energia-solar-no-brasil/. Acesso em: 8 set. 2024.

WEI, N. et al. Conventional models and artificial intelligence-based models for energy consumption forecasting: A review. Journal of Petroleum Science and Engineering, v. 181, p. 106187, 2019.

RAMENAH, H. et al. Accurate determination of parameters relationship for photovoltaic power output by augmented dickey fuller test and engle granger method. AIMS Energy, v. 6, n. 1, p. 19-48, 2018.

CHAI, T.; DRAXLER, R. R. Root mean square error (rmse) or mean absolute error (mae)? – arguments against avoiding rmse in the literature. Geoscientific Model Development, v. 7, n. 3, p. 1247-1250, 2014.

IWANA, B. K.; UCHIDA, S. An empirical survey of data augmentation for time series classification with neural networks. Plos one, v. 16, n. 7, p. e0254841, 2021.

TAYLOR, S. J.; LETHAM, B. Forecasting at scale. The American Statistician, v. 72, n. 1, p. 37-45, 2017.

CHATURVEDI, S. et al. A comparative assessment of sarima, lstm rnn and fb prophet models to forecast total and peak monthly energy demand for india. Energy Policy, v. 160, p. 112648, 2022.