

Modelagem preditiva de dados eólicos offshore para o litoral da Paraíba: Uma abordagem multifonte com Holt-Winters e Machine Learning

Isaac Gomes Veras (IFPB, Campus João Pessoa), Thiago Gouveia (IFPB, Campus João Pessoa)

E-mails: isaac-veras.iv@academico.ifpb.edu.br, thiago.gouveia@ifpb.edu.br

Área de conhecimento (Tabela CNPq): 1.03.03.04-9 Sistemas de Informação.

Palavras-chave: energia renovável; segurança energética; análise de séries temporais; random forest; predição.

1. Introdução

A transição para uma matriz energética sustentável é um dos principais desafios globais, impulsionada pela crescente preocupação com as mudanças climáticas e pela necessidade de diversificar as fontes de energia (IEA, 2023; IRENA, 2024). Nesse cenário, a energia eólica *offshore* surge como uma alternativa promissora, utilizando ventos marítimos que são, em geral, mais intensos e constantes que os terrestres (GWEC, 2023). Embora o Brasil possua um vasto potencial eólico em seu litoral, a exploração de parques *offshore* ainda é incipiente, representando uma fronteira estratégica para a expansão de energias renováveis no país (EPE, 2020). O litoral da Paraíba, com suas características de vento favoráveis, apresenta o potencial para contribuir com essa diversificação.

O principal desafio da energia eólica, contudo, reside em sua natureza intermitente. Nesse contexto, a modelagem preditiva baseada em dados de satélite e *Machine Learning* surge como uma abordagem promissora para a avaliação eficiente e operação otimizada de projetos *offshore*, impactando diretamente sua viabilidade e a segurança energética (Liu et al., 2021; Majidi Nezhad et al., 2021; Majidi Nezhad et al., 2022). Para superar essa barreira e a lacuna no acesso a dados eólicos nacionais de qualidade, esta pesquisa fundamenta-se em uma abordagem multifonte, utilizando séries temporais de alta granularidade do projeto *NASA POWER DAV* e empregando técnicas de Ciência de Dados, como a suavização exponencial de *Holt-Winters* e algoritmos de *Machine Learning* (*Random Forest*), para desenvolver um modelo preditivo robusto.

Portanto, o presente estudo tem como objetivo geral desenvolver e validar um modelo preditivo de alta precisão para a velocidade do vento *offshore* na costa da Paraíba. A pesquisa visa não apenas preencher uma lacuna de conhecimento sobre a dinâmica eólica local, mas também fornecer uma ferramenta de suporte à decisão para futuros investimentos no setor de energia renovável. Com isso, espera-se contribuir para a transição e segurança energética do Estado, alinhando-se aos esforços de desenvolvimento sustentável e inovação tecnológica no Brasil.

Além desta seção introdutória, o presente trabalho detalha, na Seção 2, os Materiais e Métodos empregados na pesquisa. Na Seção 3, são apresentados e discutidos os principais Resultados e, por fim, na Seção 4 são tecidas as Considerações Finais do estudo.

2. Materiais e métodos

A pesquisa empregou uma abordagem quantitativa, estruturada com base no *framework* de Análise Exploratória de Dados (EDA). Este *framework* foi selecionado por sua eficácia em extrair *insights*, identificar padrões, testar hipóteses preliminares e diagnosticar a qualidade dos dados por meio de técnicas estatísticas e de visualização. As etapas de aquisição, verificação de qualidade, análise estatística e visualização gráfica guiaram o desenvolvimento da pesquisa.

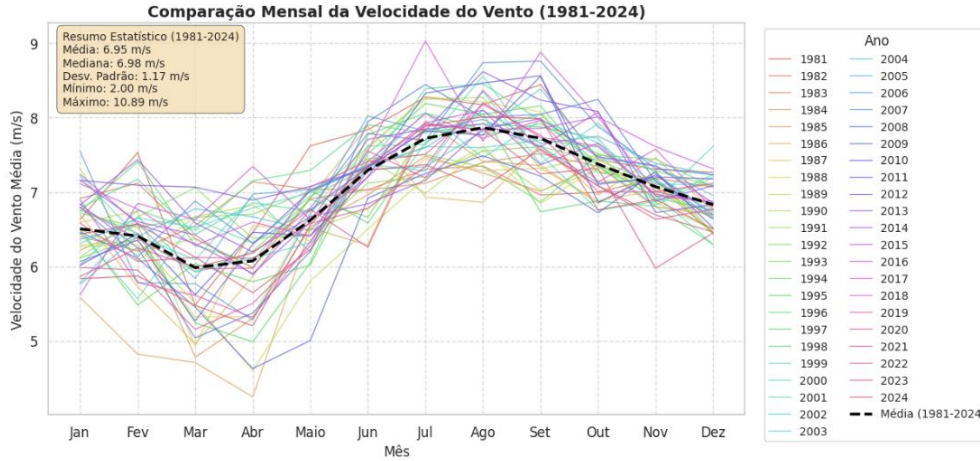
Para a análise, foram utilizados dados de séries temporais diárias da velocidade (WS50M) e direção (WD50M) do vento a 50 metros de altitude, obtidos através da API do projeto *NASA Langley Research Center (LaRC) POWER*. A escolha da altitude de 50 metros foi definida pela disponibilidade deste parâmetro padrão no dataset da NASA, o que garantiu a integridade da longa série temporal. Reconhece-se que turbinas *offshore* operam em alturas superiores, e para aplicações de engenharia os resultados podem ser extrapolados para a altura de cubo desejada através da Lei de Potência do Perfil de Vento, prática padrão na indústria eólica (SERGIENKO et al., 2022). O ponto de coleta foi definido pelas coordenadas geográficas (-7, 34.8) do litoral de João Pessoa, Paraíba, abrangendo um período histórico de janeiro de 1981 a dezembro de 2024 para a modelagem.

A etapa de preparação dos dados envolveu a estruturação da série temporal em uma frequência diária e a subsequente divisão em conjuntos de treino e teste para validação do modelo. Na fase de modelagem, adotou-se uma abordagem multifonte, iniciando com a aplicação do modelo de suavização exponencial de *Holt-Winters* e, em seguida, algoritmos de *Machine Learning Random Forest*, que se mostrou mais eficaz em identificar e prever os padrões sazonais anuais observados na análise exploratória. A avaliação dos modelos preditivos foi realizada de forma quantitativa, comparando as previsões com os dados reais do conjunto de teste. Para medir a acurácia, foram utilizadas métricas estatísticas padrão para séries temporais, como o Erro Quadrático Médio (RMSE) e o Erro Absoluto Médio (MAE). Esse processo de validação, que incluiu planos para validação cruzada temporal, foi fundamental para aferir a capacidade de generalização e a confiabilidade do modelo final.

3. Resultados e discussão

A análise da série histórica de 1981 a 2024 para o litoral da Paraíba revelou um padrão sazonal anual bem definido, com uma velocidade média geral de 6,95 m/s (desvio padrão de 1,17 m/s) e picos de velocidade nos meses de inverno (junho a setembro), conforme ilustrado no Gráfico 1.

Gráfico 1 – Comparação mensal da velocidade do vento (1981 e 2024).

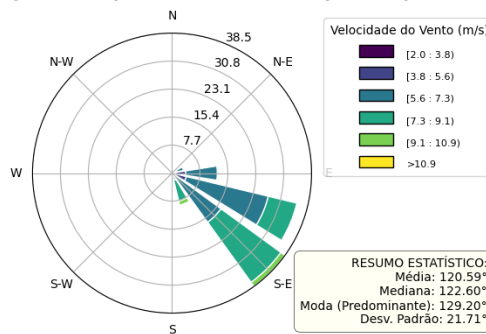


Fonte: Dados da pesquisa (2025)

Adicionalmente, a análise direcional, apresentada no Gráfico 2, demonstrou uma forte e consistente predominância dos ventos provenientes do quadrante Sudeste (moda de 129,20°). Essa baixa variabilidade na direção é um indicador altamente positivo para a viabilidade de projetos, pois simplifica o design do *layout* do parque eólico para maximizar a captura de energia e minimizar o efeito de esteira entre as turbinas.

Gráfico 2 – Distribuição de frequência e intensidade por direção (1981-2024).

Distribuição de frequência e intensidade por direção (1981-2024)

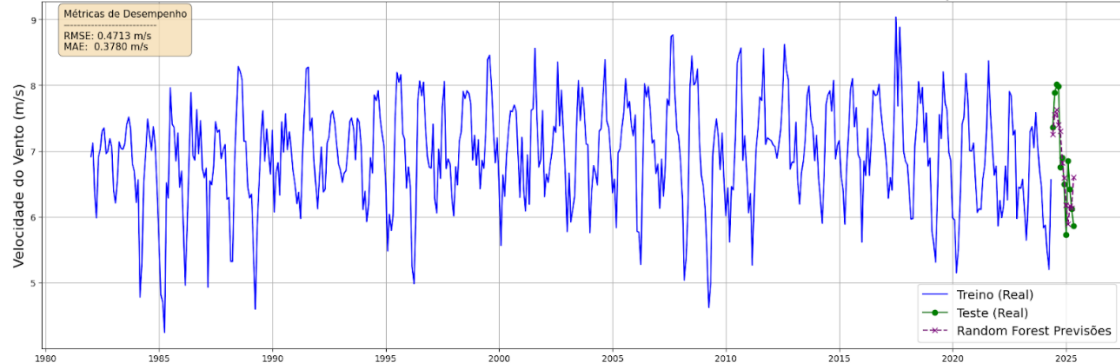


Fonte: Dados da pesquisa (2025)

No que tange à modelagem preditiva, a seleção do modelo final foi baseada em uma análise comparativa. O modelo *Holt-Winters*, aplicado como um *baseline* estatístico, obteve um Erro Quadrático Médio (RMSE) de 0,5557 m/s e o Erro Absoluto Médio (MAE) de 0,5021 m/s.

Gráfico 3 - Previsão da velocidade média do vento com *Random Forest* e métricas de desempenho.

Previsão da velocidade média do vento com *Random Forest* e métricas de desempenho



Fonte: Dados da pesquisa (2025)

Em contraste, o modelo de *Machine Learning* baseado em *Random Forest* demonstrou uma capacidade preditiva superior, alcançando um RMSE de 0,4713 m/s e um Erro Absoluto Médio (MAE) de 0,3780 m/s. A redução de mais de 58% no RMSE valida a abordagem de Ciência de Dados obtida pelo modelo *Random Forest* para o estudo. Conforme ilustra o Gráfico 3, as previsões do modelo final acompanharam de perto os dados reais do período de teste, confirmando sua confiabilidade para o planejamento operacional.

4. Considerações finais

Este estudo atingiu seu objetivo principal ao desenvolver e validar um modelo preditivo para a velocidade e direção dos ventos *offshore* na costa da Paraíba. A metodologia de Análise Exploratória de Dados permitiu identificar padrões cruciais do recurso eólico, incluindo uma forte sazonalidade anual na intensidade do vento, com média geral de 6,95 m/s, e uma consistente predominância da direção Sudeste (SE), com moda de 129,20°. A capacidade de modelos de *Machine Learning* baseados em ensemble, com destaque para o *Random Forest*, em capturar essas dinâmicas foi comprovada por uma alta precisão, validada por um Erro Quadrático Médio (RMSE) de 0.4713 m/s e um Erro Absoluto Médio (MAE) de 0.3780 m/s.

A principal contribuição da pesquisa reside em fornecer subsídios quantitativos para a tomada de decisão no setor de energias renováveis. A caracterização do regime de ventos e a validação de uma ferramenta preditiva de alta acurácia geram informações estratégicas para a análise de viabilidade econômica e otimização do *layout* de futuros parques eólicos, além de fortalecer a segurança energética regional. Dessa forma, o trabalho preenche uma lacuna no conhecimento sobre o potencial eólico específico do litoral paraibano, possuindo tanto relevância acadêmica quanto aplicação prática.

Como desdobramento futuro, sugere-se a expansão da análise para outros pontos da costa nordestina e a validação de outros algoritmos de ensemble, como o LightGBM, que também demonstrou performance de ponta em análises preliminares, bem como a exploração de arquiteturas de *Deep Learning* (LSTM) para aprimorar ainda mais as previsões. Adicionalmente, planeja-se o desenvolvimento de uma ferramenta de consulta pública, possivelmente uma aplicação web, para a visualização dos resultados gerados por esta pesquisa. Tais ações poderão consolidar os achados aqui apresentados e ampliar a contribuição da pesquisa para o desenvolvimento sustentável do setor energético no Brasil.

Referências

- EPE (Empresa de Pesquisa Energética). Roadmap Eólica Offshore Brasil. Brasília: EPE, 2020. Disponível em: https://www.epe.gov.br/sites-pt/publicacoes-dados-abertos/publicacoes/PublicacoesArquivos/publicacao-456/Roadmap_Eolica_Offshore_EPE_versao_R2.pdf. Acesso em: 12 abr. 2025.
- GWEC (Global Wind Energy Council). Global Wind Report 2023. Brussels: GWEC, 2023. Disponível em: https://sawea.org.za/sites/default/files/content-files/Market%20Reports/GWEC-2023_interactive.pdf. Acesso em: 12 abr. 2025.
- IEA (International Energy Agency). Renewables 2023. Paris: IEA, 2023. Disponível em: https://iea.blob.core.windows.net/assets/96d66a8b-d502-476b-ba94-54ffda84cf72/Renewables_2023.pdf. Acesso em: 12 abr. 2025.
- IRENA (International Renewable Energy Agency). World Energy Transitions Outlook 2024. Abu Dhabi: IRENA, 2024. Disponível em: https://www.irena.org/-/media/Files/IRENA/Agency/Publication/2024/Nov/IRENA_World_energy_transitions_outlook_2024_Summary.pdf. Acesso em: 12 abr. 2025.
- LIU, Xiaolei; LIN, Zi; FENG, Ziming. Short-term offshore wind speed forecast by seasonal ARIMA - A comparison against GRU and LSTM. *Energy*, v. 227, p. 120492, 2021. DOI: <https://doi.org/10.1016/j.energy.2021.120492>. Acesso em: 12 abr. 2025.
- MAJIDI NEZHAD, Meysam; HEYDARI, Azim; NESHAT, Mehdi; KEYNIA, Farshid; PIRAS, Giuseppe; ASTIASO GARCIA, Davide. A Mediterranean Sea Offshore Wind classification using MERRA-2 and machine learning models. *Renewable Energy*, v. 190, p. 156-166, 2022. DOI: <https://doi.org/10.1016/j.renene.2022.03.110>. Acesso em: 12 abr. 2025.
- MAJIDI NEZHAD, M.; HEYDARI, A.; PIRSHAYAN, E.; GROPPY, D.; ASTIASO GARCIA, D. A novel forecasting model for wind speed assessment using sentinel family satellites images and machine learning method. *Renewable Energy*, v. 179, p. 2198-2211, 2021. DOI: <https://doi.org/10.1016/j.renene.2021.08.013>. Acesso em: 12 abr. 2025.
- SERGIENKO, N. Y.; DA SILVA, L. S. P.; BACHYNSKI-POLIĆ, E. E.; CAZZOLATO, B. S.; ARJOMANDI, M.; DING, B. Review of scaling laws applied to floating offshore wind turbines. *Renewable and Sustainable Energy Reviews*, v. 162, p. 112477, 2022. DOI: <https://doi.org/10.1016/j.rser.2022.112477>. Acesso em: 11 jul. 2025.