



UNIVERSIDADE FEDERAL DO CEARÁ
CENTRO DE TECNOLOGIA
DEPARTAMENTO DE ENGENHARIA MECÂNICA
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA MECÂNICA

FRANCISCO DIEGO VIDAL BEZERRA

**AVALIAÇÃO DE MÉTODOS DE *ENSEMBLE* DINÂMICO EM APRENDIZADO DE
MÁQUINA PARA PREVISÕES DE IRRADIÂNCIA SOLAR E VELOCIDADE DO
VENTO**

FORTALEZA

2024

FRANCISCO DIEGO VIDAL BEZERRA

AVALIAÇÃO DE MÉTODOS DE *ENSEMBLE* DINÂMICO EM APRENDIZADO DE MÁQUINA PARA PREVISÕES DE IRRADIÂNCIA SOLAR E VELOCIDADE DO VENTO

Dissertação submetida à Coordenação do Curso de Pós-Graduação em Engenharia Mecânica da Universidade Federal do Ceará, como requisito parcial à obtenção do Título de Mestre em Engenharia Mecânica. Área de Concentração: Processos, Equipamentos e Sistemas para Energias Renováveis.

Orientador(a): Prof. Dr. Paulo Alexandre Costa Rocha

FORTALEZA

2024

Dados Internacionais de Catalogação na Publicação
Universidade Federal do Ceará
Sistema de Bibliotecas
Gerada automaticamente pelo módulo Catalog, mediante os dados fornecidos pelo(a) autor(a)

B469a Bezerra, Francisco Diego Vidal.

Avaliação de métodos de ensemble dinâmico em aprendizado de máquina para previsões de irradiância solar e velocidade do vento / Francisco Diego Vidal Bezerra. – 2024.
61 f. : il. color.

Dissertação (mestrado) – Universidade Federal do Ceará, Centro de Tecnologia, Programa de Pós-Graduação em Engenharia Mecânica, Fortaleza, 2024.

Orientação: Prof. Dr. Paulo Alexandre Costa Rocha.

1. Energia eólica. 2. Energia solar. 3. Energia renovável. 4. Aprendizado de máquina. 5. Métodos ensemble. I. Título.

CDD 620.1

FRANCISCO DIEGO VIDAL BEZERRA

AVALIAÇÃO DE MÉTODOS DE *ENSEMBLE* DINÂMICO EM APRENDIZADO DE MÁQUINA PARA PREVISÕES DE IRRADIÂNCIA SOLAR E VELOCIDADE DO VENTO

Dissertação submetida à Coordenação do Curso de Pós-Graduação em Engenharia Mecânica da Universidade Federal do Ceará, como requisito parcial à obtenção do Título de Mestre em Engenharia Mecânica. Área de Concentração: Processos, Equipamentos e Sistemas para Energias Renováveis.

Aprovada em: 07 / 08 / 2024

BANCA EXAMINADORA

Prof. Dr. Paulo Alexandre Costa Rocha
Universidade Federal do Ceará (UFC)

Prof. Dr. André Valente Bueno
Universidade Federal do Ceará (UFC)

Profª. Dra. Tatiane Carolyne Carneiro
Universidade Federal do Maranhão (UFMA)
Membro Externo

Aos meus pais, Bezerra e Lúcia. à
minha irmã Lydice e à minha
namorada Amanda.

AGRADECIMENTOS

Agradeço a Deus pelos caminhos aos quais me permitiu ser agraciado com a oportunidade da realização de um mestrado, que em meio a condições atípicas de uma pandemia mundial, parecia um projeto distante, mas em virtude do meu constante esforço pelo crescimento nos desígnios divinos, tornou-se capaz.

Agradeço aos meus pais Bezerra e Lúcia, ainda à minha irmã Lydice pelo amor infinito de bondade e inspiração, que me fez corajoso e forte diante das muitas adversidades.

Agradeço a minha namorada Amanda por seu cuidado e determinação em me auxiliar e motivar na busca do meu progresso.

Agradeço ao Prof. Dr. Paulo Alexandre por suas orientações e exemplo moral.

Agradeço aos meus colegas Felipe Pinto e Victor Oliveira, que me suportaram em todas as necessidades acadêmicas e me instruíram no sentido de sempre continuar a desenvolver meus estudos.

“Quando o discípulo está pronto, o mestre aparece.”

RESUMO

Nos últimos anos, com o aumento do interesse pela descarbonização dos processos geradores de energias por vias alternativas limpas e a conseqüente busca por novas fontes de recursos energéticos, há uma crescente exploração dos recursos da irradiância solar e da velocidade do vento, sendo fundamental o bom desempenho em sua previsibilidade face a natural intermitência destes recursos. O objetivo desta dissertação é a implementação de uma arquitetura de *ensemble* dinâmico para horizonte de tempo variando de 10 a 60 minutos para dados em intervalos de tempo de 10 minutos. A irradiância horizontal global (GHI) e a velocidade do vento foram calculadas usando quatro modelos de previsão independentes (*Random forest*, *k-nearest neighbors*, *support vector regression* e *elastic net*) para comparar seu desempenho com dois métodos de conjunto dinâmico, *windowing* e *arbitrating*, os quais combinam os resultados de modelos independentes. Os modelos autônomos e os métodos de *ensemble* dinâmico foram avaliados usando as métricas de erro RMSE, MAE, R^2 e MAPE. Os resultados deste trabalho mostraram que o método de *ensemble* dinâmico *windowing* foi o método com melhor desempenho quando comparado aos outros modelos avaliados. Para ambos os casos de previsão da velocidade do vento e da irradiância solar, o modelo de *windowing* atingiu os melhores valores de erro em termos de RMSE para todos os horizontes de previsão avaliados. Usando esta abordagem, o ganho na previsão da velocidade do vento foi de 0,56% quando comparado com o segundo melhor modelo de previsão, enquanto o ganho na previsão do GHI foi de 1,96%, considerando a métrica RMSE. O desenvolvimento de um modelo *ensemble* capaz de fornecer estimativas precisas pode ser implementado em aplicações de previsão em tempo real, auxiliando na avaliação da operação de parques eólicos e solares.

Palavras chaves: energia eólica; energia solar; energia renovável; aprendizado de máquina; métodos *ensemble*.

ABSTRACT

In recent years, with increased interest in decarbonizing energy-generating processes through clean alternatives and the consequent search for new energy sources there has been a growing exploitation of solar irradiance and wind speed resources, and good predictability performance is essential given the natural intermittency of these resources. The aim of this thesis is to implement a dynamic ensemble architecture for time horizons ranging from 10 to 60 minutes for data at 10-minute time intervals. Global horizontal irradiance (GHI) and wind speed were calculated using four independent prediction models (Random forest, k-nearest neighbors, support vector regression and elastic net) to compare their performance with two dynamic ensemble methods, windowing and arbitrating, which combine the results of independent models. The autonomous models and dynamic ensemble methods were evaluated using the RMSE, MAE, R^2 and MAPE error metrics. The results of this work showed that the dynamic ensemble windowing method was the best performing method when compared to the other models evaluated. For both wind speed and solar irradiance forecasts, the windowing model achieved the best error values in terms of RMSE for all the forecast horizons evaluated. Using this approach, the gain in wind speed forecasting was 0.56% when compared to the second best forecasting model, while the gain in GHI forecasting was 1.96%, considering the RMSE metric. The development of an ensemble model capable of providing accurate estimates can be implemented in real-time forecasting applications, helping to evaluate the operation of wind and solar farms.

Keywords: wind energy; solar energy; renewable energy; machine learning; forecasting ensembles.

LISTA DE FIGURAS

Figura 1 - Fluxograma de dados para a metodologia aplicada	28
Figura 2 - Mapa do Nordeste do Brasil. O local de medição de Petrolina está em destaque (Google, 2023).....	29
Figura 3 - Influência da variação do parâmetro λ de windowing no RMSE para diferentes horizontes de tempo na análise de dados de velocidade do vento para todos os horizontes de tempo estudados.	33
Figura 4 - A influência do parâmetro λ de windowing no valor RMSE para o horizonte de tempo $t+10$	34
Figura 5 - Influência do parâmetro λ de windowing no valor MAE para o horizonte de tempo $t+10$	35
Figura 6 - Influência do parâmetro λ de windowing no valor R^2 para o horizonte de tempo $t+10$	36
Figura 7 - Influência do parâmetro λ de windowing no valor MAPE para o horizonte de tempo $t+10$	37
Figura 8 - Efeito da variação do parâmetro λ no desempenho do método windowing. O resultado SVR é mostrado para referência.....	37
Figura 9 - A influência da variação do parâmetro λ de windowing Parâmetro no RMSE para todos os horizontes de tempo estudados na análise de dados de irradiação solar. ...	39
Figura 10 - Influência do parâmetro λ de windowing no valor de RMSE no horizonte de tempo $t+10$	40
Figura 11 - Influência da variação do parâmetro λ de windowing no valor de R^2 no horizonte de tempo $t+10$	41
Figura 12 - Influência da variação do parâmetro λ de windowing no valor de MAE no horizonte de tempo $t+10$	42
Figura 13 - Influência da variação do parâmetro λ de windowing no valor de MAPE no horizonte de tempo $t+10$	43

LISTA DE TABELAS

Tabela 1 - Coordenadas geográficas, altitude em relação ao nível do mar, intervalos de medição e períodos de medição dos dados foram coletados na estação Petrolina. IM e PM significam, respectivamente, “intervalo de medição” e “período de medição”.	29
Tabela 2 - Parâmetros de pesquisa e valores de grid aplicados aos métodos testados.	31
Tabela 3 - Melhores parâmetros para cada método de aprendizado de máquina.	32
Tabela 4 - Comparação dos valores de RMSE (m/s), utilizando diferentes métodos para diferentes horizontes de tempo e windowing da variação dos parâmetros λ . Os melhores resultados para cada horizonte temporal estão em negrito.	33
Tabela 5 - Comparação dos valores do MAE (m/s), utilizando diferentes métodos em diferentes horizontes de tempo e windowing da variação do parâmetro λ . Os melhores resultados para cada horizonte temporal estão em negrito.	34
Tabela 6 - Comparação dos valores de R^2 (m/s), utilizando diferentes métodos em diferentes horizontes de tempo e da variação do parâmetro λ de windowing. Os melhores resultados para cada horizonte de tempo estão em negrito.	35
Tabela 7 - Comparação dos valores de MAPE (m/s), utilizando diferentes métodos em diferentes horizontes de tempo e da variação do parâmetro λ de windowing. Os melhores resultados para cada horizonte de tempo estão em negrito.	36
Tabela 8 - Melhores parâmetros para cada método de aprendizado de máquina.	38
Tabela 9 - Comparação dos valores de RMSE (W/m^2), utilizando diferentes Métodos em diferentes horizontes de tempo variando o parâmetro λ de windowing. Os melhores resultados para cada Horizonte de Tempo estão em negrito.	38
Tabela 10 - Comparação dos valores de R^2 (W/m^2), utilizando diferentes métodos em diferentes horizontes de tempo variando o parâmetro λ de windowing. Os melhores resultados para cada horizonte de tempo estão em negrito.	40
Tabela 11 - Comparação dos valores de MAE (W/m^2), utilizando diferentes métodos em diferentes horizontes de tempo variando o parâmetro λ de windowing. Os melhores resultados para cada horizonte de tempo estão em negrito.	41
Tabela 12 - Comparação dos valores de MAPE (W/m^2), utilizando diferentes métodos em diferentes horizontes de tempo variando o parâmetro λ de windowing. Os melhores resultados para cada horizonte de tempo estão em negrito.	42
Tabela 13 - Compilação dos resultados obtidos pelo método windowing em diferentes horizontes de tempo.	44

Tabela 14 - Compilação de resultados da literatura para previsão da velocidade do vento. ... 44

Tabela 15 - Compilação de resultados para previsão de irradiação global horizontal (GHI).. 46

LISTA DE ABREVIATURAS E SIGLAS

CNN-1D	Convolutional Neural Network (Rede Neural Convolucional)
DELNET	Dynamic Elastic Net (Rede Elástica Dinâmica)
DL	<i>Deep Learning</i> (Aprendizagem Profunda)
DNI	<i>Direct Normal Irradiance</i> (Irradiância Normal Direta)
DNN	<i>Deep Neural Network</i> (Rede Neural Profunda)
EN	<i>Elastic Net</i> (Rede Elástica)
GAT	<i>Graph Attention Network</i> (Rede de Atenção Gráfica)
GB	<i>Gradient Boosting</i> (Aumento de Gradiente)
GHI	<i>Global Horizontal Irradiance</i> (Irradiância Horizontal Global)
GSR	<i>Global Solar Radiation</i> (Radiação Solar Global)
GRU	<i>Gated Recurrent Unit</i> (Unidade Recorrente Fechada)
IM	Intervalo de Medição
INPE	Instituto Nacional de Pesquisas Espaciais
KNN	<i>K-Nearest Neighbors</i> (k vizinhos mais próximos)
LSTM	<i>Long Short-Term Memory</i> (Memória de Longo Prazo)
MAE	<i>Mean Absolute Error</i> (Erro Médio Absoluto)
MAPE	<i>Mean Absolute Percentage Error</i> (Erro Médio Absoluto Percentual)
MBE	<i>Mean Bias Error</i> (Erro Médio de Polarização)
MCG	Modelo Climático Global
MEMD	<i>Multivariate Empirical Mode Decomposition</i> (Decomposição do Modo Empírico Multivariado)
ML	<i>Machine Learning</i> (Aprendizado de Máquina)
PCA	Principal Component Analysis (Análise de Componente Principal)
PM	Período Medido
RF	<i>Random Forest</i> (Florestas Aleatórias)
RMSE	<i>Root Mean Squared Error</i> (Raíz do Erro Médio Quadrático)
SONDA	Sistema de Organização de Dados Ambientais
SVM	<i>Support Vector Machine</i> (Máquina de Vetor Suporte)
SVR	<i>Support Vector Regression</i> (Regressão de Vetor Suporte)
WS	<i>Wind Speed</i> (Velocidade do Vento)

LISTA DE SÍMBOLOS

I_{cs}	Fator de céu claro
kt	Razão entre o valor de irradiância horizontal global (GHI) (I) e irradiância de céu claro (I_{cs}),
I	Irradiância horizontal global (GHI)
λ	Parâmetro dimensional da janela de dados (<i>windowing</i>)

SUMÁRIO

1 INTRODUÇÃO	16
1.1 Objetivos	19
2 FUNDAMENTAÇÃO TEÓRICA	20
2.1 Recursos naturais intermitentes	20
<i>2.1.1 Recurso eólico</i>	21
<i>2.1.2 Recurso solar</i>	21
2.2 Modelos de aprendizado de máquina	22
<i>2.2.1 Random Forest</i>	22
<i>2.2.2 K-nearest neighbors</i>	23
<i>2.2.3 Support vector regression</i>	23
<i>2.2.4 Elastic net</i>	24
<i>2.2.5 Windowing</i>	24
<i>2.2.6 Arbitrating</i>	25
2.3 Critério de comparação das métricas de performance	25
<i>2.3.1 Coeficiente de determinação (R^2)</i>	26
<i>2.3.2 Raiz do erro quadrático médio (RMSE)</i>	26
<i>2.3.3 Erro absoluto médio (MAE)</i>	26
<i>2.3.4 Erro percentual absoluto médio (MAPE)</i>	27
3 METODOLOGIA	27
3.1 Aquisição de Dados	28
<i>3.1.1 Dados de velocidade do vento</i>	30
<i>3.1.2 Dados de irradiância solar</i>	30
3.2 Modelos de previsão de aprendizado de máquina e parâmetros de métodos de ensemble dinâmico	30
3.3 Critério comparativo por métricas de performance	31
4 RESULTADOS E DISCUSSÃO	32
4.1 Previsões de velocidade de vento	32

4.2 Previsões de irradiância.....	38
4.3 Comparação de resultados com a literatura.....	43
5 CONCLUSÃO	48
REFERÊNCIAS.....	51

1 INTRODUÇÃO

A eletricidade gerada por fontes de combustíveis fósseis tem sido o principal motor das alterações climáticas, provavelmente responsável por mais de 70% das emissões de gases geradores do efeito de estufa e por mais de 90% de todas as emissões de gases de carbono. A alternativa de descarbonizar o sistema de geração de eletricidade mundial é uma tendência focada em fontes de energia renovável, cujos custos de geração estão cada vez mais acessíveis (Osman *et al.*, 2022). Apesar disso, o impacto da geração intermitente na rede elétrica é um efeito indesejável quando se trata de geração elétrica a partir de recursos energéticos alternativos, como a velocidade do vento e a radiação solar global (Calif *et al.*, 2016). Como esta geração é dependente das condições climáticas, um dos meios para eliminar ou reduzir suas incertezas é a disponibilidade de bons métodos de previsão para estes recursos (Carneiro *et al.*, 2022).

A busca por parâmetros que possam descrever o comportamento atmosférico e sua previsibilidade tem levado ao desenvolvimento de pesquisas em aprendizado de máquina, baseadas nos mais diversos tipos de preditores, e à criação de modelos para utilização em diversas áreas. Em Shikhovtsev *et al.* (2023), o aprendizado de máquina multicamadas é usado para melhorar a resolução de telescópios astronômicos baseados em terra. Em Yuval *et al.* (2023) parâmetros são usados para construir um modelo de circulação atmosférica.

As influências dos fatores atmosféricos na geração de energia elétrica a partir de fontes solar e eólica são normalmente o principal problema na geração de redes inteligentes, onde as centrais de produção de grande escala precisam de ser integradas na rede elétrica, o que afeta diretamente os objetivos de planejamento, investimento e tomada de decisão. Os modelos de previsão podem minimizar esse problema por meio de modelos de aprendizado de máquina (Meenal *et al.*, 2022).

Os benefícios da otimização da previsão de geração a partir de fontes eólica e solar por meio de modelos matemáticos também é um fator de influência econômica, pois confere maior segurança ao setor elétrico por meio da melhoria dos contratos de compra de energia renovável em virtude da maior acurácia das previsões (Mesa-Jiménez *et al.*, 2023).

Um conjunto de dados de 14 anos foi explorado em Rocha *et al.* (2019), contendo valores diários de variáveis meteorológicas. Este conjunto de dados foi usado para treinar três arquiteturas de *Deep Neural Networks* (DNNs) em vários horizontes de tempo para prever a radiação solar global para Fortaleza, na região Nordeste do Brasil. A precisão das previsões foi considerada considerável de acordo com seus valores de erro quadrático médio normalizado

(nRMSE) e boa em relação aos valores de *Mean Absolute Percentage Error* (MAPE). A variabilidade dos modelos matemáticos de predição tem importância individual inerente a cada um dos métodos empregados e, neste cenário, modelos de conjuntos dinâmicos emergem, que apresentam potencial de melhorar o desempenho quando comparados aos modelos individuais, pois buscam a otimização máxima ao considerar o melhor dos modelos individuais. Esta abordagem é atualmente utilizada com muito sucesso nas áreas de pesquisa e industrial.

Diversos métodos de conjunto dinâmico foram desenvolvidos para prever a geração de energia a partir de fontes renováveis, nos quais utilizam a presença de modelos de previsão bem conhecidos, como *Random Forest* (RF), *Support vector regression* (SVR) e *k-nearest neighbors* (kNN), que são aplicados para integrar otimizações para uso em métodos de conjuntos dinâmicos (Du *et al.*, 2022).

O SVR é um algoritmo de regressão que usa coordenadas para observações individuais e hiperplanos para segregar conjuntos de dados. Este é um método amplamente utilizado para categorizar e classificar *clusters*. Este modelo foi desenvolvido pela primeira vez para fins de classificação e é amplamente testado em Vapnik *et al.* (1995), Smola *et al.* (1996) e ainda em abordagens recentes, como as de Mahesh *et al.* (2023), onde trata do desenvolvimento de um novo método para o rastreamento do ponto de potência máxima de um painel fotovoltaico e em Demir *et al.* (2023), onde é discutida a estimativa da radiação solar por meio de cinco abordagens diferentes de aprendizado de máquina.

O método kNN é um algoritmo de aprendizado supervisionado amplamente utilizado como classificador que, utiliza como parâmetro uma distância a dados vizinhos mais próximos, realiza a categorização por similaridade e prevê uma nova amostra usando as k amostras mais próximas. Recentemente, esta abordagem foi usada em Schwegmann *et al.* (2023), onde estações meteorológicas virtuais usam dados numéricos calibrados para fornecer estimativas precisas do vento durante todas as fases de um projeto de energia eólica para reproduzir condições ambientais ideais específicas do local.

Diversos estudos se concentram na previsão mais apurada da geração da energia elétrica através do recurso eólico, onde são consideradas as flutuações aleatórias e as incertezas envolvidas. Um estudo em Che *et al.* (2023) propõe uma nova previsão probabilística de energia eólica de ultracurto prazo usando uma modelagem de correção de erros pela abordagem de floresta aleatória.

O *Elastic Net* (EN) é um método de regressão regularizado que combina linearmente as penalidades dos métodos LASSO e Ridge. Em Nikodinoska *et al.* (2022) o estudo utiliza combinações de previsões obtidas pela aplicação de dados regionais da Alemanha para energia

solar fotovoltaica e eólica através do modelo *Elastic Net*, com validação cruzada e estimativa de janela rolante, no contexto de previsões de energias renováveis.

O estado da arte utiliza métodos de conjuntos dinâmicos a partir da combinação de resultados obtidos de modelos tradicionais agrupados, como a abordagem de meta-aprendizagem *arbitrating*, a qual utiliza a combinação de resultados de acordo com as previsões da perda resultantes e ainda a abordagem de *windowing*, que possui parâmetro de ajuste da quantidade de dados a serem considerados (Cerqueira *et al.*, 2019). Em Lakku *et al.* (2022), um modelo climático global (MCG) é estudado para melhorar a simulação da velocidade do vento próximo à superfície (WS) através da comparação de 28 modelos acoplados usando componentes dinâmicos.

Em Ji *et al.* (2022), um modelo híbrido de aprendizagem por transferência baseado em uma rede neural convolucional e uma rede neural recorrente fechada é proposta para prever a velocidade do vento em um cânião de curto prazo com menos dados de observação. O método usa uma janela deslizante de tempo para extrair séries temporais de dados históricos de velocidade do vento e dados de temperatura de cidades adjacentes como entrada da rede neural.

Em Su *et al.* (2020), os autores estudaram o Multi-GRU-RCN, um modelo de conjunto para obter informações significativas, como precipitação e irradiância solar, por meio de previsões de movimento de nuvens de curto prazo a partir de uma imagem de nuvem. A modelagem *ensemble* utilizada em Carneiro *et al.* (2022) integra metodologias de previsão eólica e solar aplicadas a dois locais com diferentes latitudes e perfis climáticos. Os resultados obtidos reduzem os erros de previsão e podem ser úteis na otimização do planejamento para utilização de recursos solares e eólicos intermitentes nas matrizes elétricas.

Um novo modelo de conjunto proposto em Santos *et al.* (2023) foi baseado em *Graph Attention Network* (GAT) e *GraphSAGE* para prever a velocidade do vento em uma abordagem bidimensional, usando um conjunto de dados holandês, considerando vários horizontes de tempo, atrasos de tempo e influências climáticas. Os resultados mostraram que o modelo *ensemble* proposto foi equivalente ou superou todos os modelos analisados e apresentou valores de erro menores que os encontrados na literatura de referência. Sob um intervalo de tempo de 5 minutos, em Marinho *et al.* (2023) foram estudados horizontes de tempo variando de 5 a 30 minutos na avaliação das previsões de curto prazo da irradiância solar para *Global Horizontal Irradiance* (GHI) e *Direct Normal Irradiance* (DNI) usando *Convolutional Neural Network* (CNN-1D), *Long Short-Term Memory* (LSTM) e CNN-LSTM. As métricas utilizadas foram o *Mean Absolute Error* (MAE), o *Mean Bias Error* (MBE), *Root Mean Squared Error* (RMSE), *Relative Root Mean Squared Error* (rRMSE) e o coeficiente de determinação (R^2). A melhor

acurácia foi obtida para um horizonte de 10 minutos, melhorando 11,15% nesta métrica de erro em comparação ao modelo de persistência.

Reconhecendo que existem estudos empregando diferentes arquiteturas DNN, como GNN, CNN e LSTM, alcançando resultados satisfatórios em diferentes campos da ciência (Wu *et al.*, 2022; Oliveira *et al.*, 2023; Tabrizi *et al.*, 2021; ZHANG *et al.*, 2022), o presente trabalho centra-se no *Machine Learning* (ML) clássico, uma vez que o objetivo principal é identificar a melhor abordagem de conjunto de apoio aos procedimentos de ML, avaliando a influência das abordagens *arbitrating* e *windowing*, que são aplicadas em conjuntos dinâmicos em comparação aos algoritmos de aprendizado de máquina tradicionais, com foco na previsão dos parâmetros de recursos naturais para a geração de energia elétrica. Os modelos de conjunto dinâmico apresentam, em diversos estudos, avaliação de eficiência na previsão, utilizando dados de interesse para produção de energia com variáveis de entrada de velocidade do vento e irradiância solar. Seguiu-se nesta dissertação uma abordagem que explora métodos de conjuntos dinâmicos, uma vez que estes buscam a melhor eficiência de modelos pré-existentes para gerar um resultado de previsibilidade único e mais eficaz.

1.1 Objetivos

O objetivo dessa dissertação é avaliar o desempenho de seis abordagens (*Random Forest*, *k-Nearest Neighbours*, *Support Vector Regression*, *Elastic Net*, *Windowing* e *Arbitrating*) de aprendizado de máquina para previsão de irradiância solar e velocidade de vento, utilizando bases de dados coletadas de estações anemométrica e solarimétrica da cidade de Petrolina, PE, sendo que dois destes modelos utilizam a abordagem de *ensemble* dinâmico (*arbitrating* e *windowing*). Os objetivos específicos deste trabalho são:

- Implementar em linguagem *Python* os modelos de aprendizado de máquina;
- Obter e tratar os dados de velocidade de vento e irradiância, fornecidos pelo Sistema de Organização Nacional de Dados Ambientais (SONDA) de estação anemométrica e solarimétrica localizada na cidade de Petrolina - PE, localizada na região do Nordeste brasileiro;
- Realizar o comparativo de desempenho entre os modelos de aprendizado de máquina, através da avaliação de métricas de erros determinísticos *Root Mean Squared Error (RMSE)*, *Mean Absolute Error (MAE)*, R^2 e *Mean absolute percentage error (MAPE)*;

- Apresentação dos resultados obtidos pelas técnicas individuais e pelos dois modelos que utilizam *ensemble* dinâmico;
- Avaliar a variação de desempenho do modelo de maior eficiência a partir da variação de parâmetros próprios;
- Comparar os resultados com os já apresentados na literatura.

Na seção fundamentação teórica são abordados os principais conceitos que formam a base da presente dissertação, uma introdução ao local de aquisição de dados, importância da previsibilidade da velocidade do vento e irradiância solar, os métodos de aprendizado de máquina que foram aplicados e as métricas de comparação de performance.

Na seção metodologia são apresentados como os dados de velocidade de vento e irradiância solar são utilizados pelos modelos de aprendizado de máquina e seus parâmetros, o comparativo entre os métodos, inclusive aos métodos de *ensemble* dinâmico, seus parâmetros e utilização das métricas comparativas de performance a partir da obtenção dos resultados.

Na seção resultados e discussão, os resultados concentram-se na análise de métricas de eficiência para os métodos de aprendizado de máquina empregados. Esta análise determina quais métodos/parâmetros obtêm o melhor desempenho na aplicação dos dados de velocidade do vento e irradiância solar.

2 FUNDAMENTAÇÃO TEÓRICA

A seção fundamentação teórica está dividida em quatro subseções. Na primeira subseção são apresentados os recursos eólico e solar, suas características e a forma de modelagem. Na segunda subseção são apresentadas as distribuições de probabilidade utilizadas e o estado da arte do uso dessas distribuições em análises estatísticas de velocidade de vento e irradiância solar. Na terceira subseção são apresentados os métodos determinísticos utilizados na estimativa dos parâmetros. Na última subseção, os testes estatísticos usados para medir a qualidade dos resultados são apresentados.

2.1 Recursos naturais intermitentes

A produtividade ao longo dos anos está firmemente ligada aos fatores que impulsionam a economia mundial e um dos fatores de relevância é a energia, que tem na atualidade o

incentivo a exploração de novos recursos alternativos e menos agressivos ao meio ambiente, tais como energia eólica e solar (Kwasney *et al.*, 2017).

A partir da busca por novas fontes de energias renováveis, deveu-se também conciliar a demanda de crescimento econômico à produção de energia por meio de fontes alternativas baseadas em recursos naturais, quais seja, água, ar, sol e derivados vegetais, aumentando os investimentos na otimização dos meios destas fontes de geração (Bermann, 2008; Vichi *et al.*, 2009).

2.1.1 Recurso eólico

Diversas tecnologias foram desenvolvidas a partir da necessidade de conter a ação antropogênica na contribuição pela utilização dos combustíveis fósseis. Diversas tecnologias alternativas surgiram, como a geração através da energia eólica, e estão agora competitivas frente às tradicionais (IPCC, 2022)

O recurso eólico, apesar de ter se destacado nos últimos anos e atingido considerável participação na composição da matriz elétrica brasileira, ainda não atingiu seu potencial máximo de expansão, devido às características dos investimentos econômicos atualmente serem de retração. O setor eólico contabiliza na atualidade cerca de 32 GW de potência instalada e distribuída em mais de mil parques eólicos operacionais no Brasil. (ABEEólica, 2024).

A característica dos regimes de vento no Brasil é sazonal (regimes sazonais diurnos e anuais, apresentando maior intensidade no segundo semestre) e possui peculiaridades de cada região. Soma-se a isso as incertezas nas estimativas de produção de energia, que é usual e apresentada nos leilões de energia (EPE, 2023).

Com o intuito de reduzir as incertezas inerentes aos perfis de velocidade do vento na estimativa de geração eólica, foram desenvolvidos diversos trabalhos ao longo dos anos que tratam da previsibilidade deste recurso através de algoritmos de aprendizado de máquina, como a incidência das cargas de vento a partir da velocidade de vento em Ding *et al.* (2023) e ainda a melhora na eficiência operacional dos sistemas de geração de energia eólica pela previsão de sua geração em Sun *et al.* (2024).

2.1.2 Recurso solar

A energia solar tem se mostrado mais eficiente em sua aplicação devido às vantagens inerentes de sua captação, possui vantagens propícias a uma geração distribuída e localizada, o

que não exige alto custo, possui baixo impacto ambiental e possui a característica de capacidade de geração em locais remotos (Oliveira *et al.*, 2023).

A capilaridade deste tipo de recurso energético tem gerado trabalhos que se baseiam em acessos onde não há acesso e fornecimento de energia convencional, impulsionando políticas públicas que fomentam e incentivam a produção de energia elétrica por fonte solar de modo remoto como se analisa em Ottavianelli *et al.* (2021).

2.2 Modelos de aprendizado de máquina

Os métodos de aprendizado de máquina ao longo do tempo têm se desenvolvido através da evolução dos algoritmos matemáticos com auxílio do processamento computacional. As aplicações ao mundo real são diversas, Em Saito *et al.* (2019) tem-se a importância do desenvolvimento de modelo de predição no diagnóstico de carcinoma hepático, já em Monkgonyane *et al.* (2019) é tratado o reconhecimento automático de fala baseado em algoritmo de aprendizado de máquina, a classificação de textos pelo método de vetor suporte (Kim *et al.*, 2005; Burdisso *et al.*, 2019) e ainda a detecção de fraude em transações financeiras (Dighe *et al.*, 2018).

2.2.1 *Random Forest*

As chamadas árvores de decisão tratam de um dos algoritmos mais comuns e originários do *Random Forest*, um algoritmo desenvolvido por Breiman *et al.* (2001), e que é mais um dos algoritmos de aprendizado de máquina em que se compõe por várias árvores de decisão agrupadas e independentes e, portanto, não relacionadas e possui a vantagem de poder ser um método utilizado tanto em naturezas de classificação, assim como em regressões.

A ideia por trás do *Random Forest* é sua vantagem em relação às árvores de decisão é a de que várias árvores de decisão não correlacionadas performam melhor do que uma só. A parametrização deste método se baseia em 3 fatores em que são considerados a abrangência do nó, a quantificação de árvores, assim como a quantidade de recursos considerados para exposição.

O grau de aleatoriedade utilizado neste método possui maior resistência ao *overfitting*, frente ao de árvores de decisão, maior poder de decisão e variedade de análise, uma vez que pode ser utilizado em questões de classificação e regressão. A aleatoriedade devida a este método se baseia por este ser um modelo com *bagging*, uma estrutura que usa método *ensemble*,

por associar a individualidade de vários modelos, no caso as árvores de decisão, a uma abordagem de maior relevância, complexidade e precisão, consolidado nas florestas aleatórias. A importância do método *ensemble* é retratado em James *et al.* (2013).

Em Lahouar *et al.* (2017) um estudo de previsão de geração de energia eólica utiliza uma abordagem com RF de quantis, aplicada a uma previsão com antecedência horária, no qual o conceito se baseia em ao invés de registrar um resultado médio em cada ramo da árvore, é retornada uma distribuição completa dos valores de resposta para cada entrada, onde os resultados demonstram a vantagem do RF frente a outros métodos clássicos, uma vez que se utilizam entradas exógenas (velocidade do vento e direção) e sua estabilidade frente a possíveis perturbações que possam ser originadas quando se utilizam entradas irrelevantes, uma das principais vantagens do modelo, que não necessita de ajustes ou otimizações.

2.2.2 K-nearest neighbors

O algoritmo de KNN é um dos métodos mais difundidos devido a facilidade de sua implementação e proposto por Fukunaga *et al.* (1975). Como classificador, a variável k se trata da quantidade de rótulos classificados a serem considerados nas proximidades de uma nova amostra. As métricas aplicadas a este algoritmo se resumem a distância das amostras e a variável k . A métrica de distância pode ser calculada por vários métodos, porém a distância euclidiana é a mais comumente utilizada. Já para o valor de k , este sensivelmente influencia no desempenho da análise e deve objetivar-se sua escolha com base na melhor precisão, podendo-se utilizar de otimização.

Já para regressão, o método também é passível de aplicação, porém fazendo o uso da média das classes mais próximas e adquiridas através da parametrização da variável k .

Em Pedro *et al.* (2018) utiliza-se o modelo KNN e *gradient boosting* (GB) para previsão de GHI e DNI por meio de aquisição de registro de dados a partir de piranômetro e imagens do céu. O modelo KNN, mesmo sendo um modelo mais simples, apresentou neste estudo apresentou melhor desempenho que GB quando se avaliou previsões probabilísticas.

2.2.3 Support vector regression

O SVR é um algoritmo de maior simplicidade e de execução relativamente adequada por utilizar um menor custo computacional quando comparado a outros métodos tradicionais. Sua origem remonta ao desenvolvimento do aprendizado estatístico (Vapnik,

1999) O método de baseia na busca de um hiperplano que pode ser encontrado em um espaço de N dimensões, sendo este N o número de recursos.

O método pode ser aplicando tanto como classificador, quanto em modelos de regressão e utiliza-se da função kernel, que basicamente converte dados em dimensões superiores. Na classificação, que é a abordagem mais usual para este algoritmo, é se aplica um hiperplano e margens de tolerância. A margem é interpretada como sendo o menor distanciamento do hiperplano aos grupos classificadores. O hiperplano que melhor classificar os agrupamentos é tratado como o de melhor ajuste, portanto, o que apresenta maior margem possível.

Em Dahmani *et al.* (2024) há o objetivo de aprimorar as capacidades preditivas de dois modelos robustos de máquinas de vetores de suporte, nomeadamente SSVM e BASVM (empilhamento de 30 ISVM). Esses modelos são projetados para a previsão precisa da radiação solar global horária com o uso de *Support Vector Machine* (SVM) para prever radiação média horária. Dados experimentais de *Global Solar Radiation* (GSR) foram comparados com o GSR calculado, e excelentes coeficientes de correlação (R^2) foram encontrados (0,9913) durante a fase de testes.

2.2.4 Elastic net

O Elastic Net (Zou *et al.*, 2005) é um algoritmo que faz o uso relacional entre a combinação dos modelos de penalização Ridge (Hoerl *et al.*, 1970) e Lasso (Tibshirani *et al.*, 1996), para implementar uma regularização em grupos de preditores correlacionados. O comparativo entre as penalidades é tratado por meio de um parâmetro α , que vai de 0 a 1, realizando a mistura das duas penalidades, na função de perda se o parâmetro de retorno for 0, este se aproxima da aplicação penalidade *ridge* e se for de maior proximidade de 1, a função de penalidade mais se adequa ao lasso.

Em Nikodinoska *et al.* (2022) é utilizado um método de *Dynamic Elastic Net* (DELNET), com validação cruzada e estimativa de janelas rolantes, no contexto de previsões de energias renováveis. Para irradiância solar as previsões obtidas atingiram melhoria de 13,4% em relação a média simples, enquanto para a previsão de velocidade do vento alcançou-se uma previsão 6,1% melhor.

2.2.5 Windowing

O *windowing* é uma abordagem em que se agrupam os resultados de outros modelos de previsão. A diversidade dos modelos torna a análise de previsão rica e complexa, pois cada modelo possui pontos fortes e outros pontos fracos, no sentido de que a partir desta combinação os melhores resultados podem ser tratados e considerados para obter previsões mais precisas.

Para realizar essa combinação é necessário estimar em quais pontos determinados modelos específicos apresentam melhor desempenho. O *windowing* (Cerqueira *et al.*, 2019) é uma abordagem de conjunto dinâmico, onde os pesos são calculados com base no desempenho de cada modelo individual, avaliado em uma janela de temporal referente a dados imediatamente anteriores. A dimensão desta janela é parametrizada pelo valor λ . Isso significa que os pesos de cada modelo são reavaliados a cada passo de tempo, e depois são classificados para catalogar apenas os melhores resultados de desempenho, gerando um modelo híbrido.

2.2.6 *Arbitrating*

O método *arbitrating* (Cerqueira *et al.*, 2017) usa o método *metalearning* para aprender e prever os classificadores. Neste estudo, consideram-se os pesos baseados no desempenho de cada modelo para um determinado intervalo de tempo. A cada instante de simulação, o modelo mais confiável é selecionado e incluído no processo de predição.

2.3 Métricas de avaliação das metodologias de previsão

Em modelos de regressão é fundamental a avaliação de suas performances, tais como os utilizados neste estudo, que usa como objetivo a comparação direta entre a eficácia dos métodos para seleção de melhor desempenho e inserção em um modelo de *ensemble* dinâmico.

As métricas de performance em que se avaliam regressões levam em conta o diferencial entre um valor que seja tido como real e seu equivalente previsto, onde y se trata do valor real e \hat{y} é associado a resultados preditos, tal qual como descrito na Equação 1.

$$e = y - \hat{y} \quad (1)$$

Cada modelo possui vantagens e desvantagens que podem ser visualizadas através de métricas específicas de performance sendo, portanto, não ideal a interpretação de dados ligada a apenas uma métrica de performance para a avaliação de eficiência de um modelo (Kuhn, 2018).

2.3.1 Coeficiente de determinação (R^2)

O coeficiente de determinação é uma métrica de avaliação de performance comumente medida em valor percentual que pode ser interpretado como uma porção do resultado dos dados adquiridos que podem ser explicados pelo modelo. Um valor de R^2 de 0,5, por exemplo, significa que o modelo pode explicar metade da variação do resultado. Apesar de existirem variadas equações para calcular esta métrica (Kvålseth, 1985), a sua versão mais simples resulta no coeficiente de determinação utilizando a correlação entre os valores observados e previstos.

2.3.2 Raiz do erro quadrático médio (RMSE)

Quando o resultado de um modelo é um número, o método mais comum para caracterizar as capacidades preditivas de um modelo é usar a raiz do erro quadrático médio (RMSE). Esta métrica é determinada a partir de função dos resíduos gerados pelo modelo, uma interpretação que pode ser dada pela subtração entre os resultados observados as previsões do modelo.

O RMSE é calculado a partir da raiz do erro quadrático médio MSE. O valor é interpretado como sendo a distância dos resíduos a partir do zero ou como a distância média entre os valores observados e o resultado das previsões do modelo. O RMSE é uma métrica que tem se mostrado como de maior preferência quando comparada ao MSE, por mais que as duas sejam populares por sua relevância em publicações de resultados estatísticos (Hyndman *et al.*, 1985).

2.3.3 Erro absoluto médio (MAE)

O erro absoluto médio é uma das métricas que possui seu foco na interpretação do desempenho a partir da média da diferença absoluta entre valores previstos e os de observação. Enquanto ao se utilizar a métrica RMSE, seus os erros são elevados ao quadrado antes mesmo do cálculo de sua média, gerando maior sensibilidade, por exemplo, a presença de *outliers*, já na métrica MAE a sensibilidade a este tipo de erros de maior relevância não é tão presente. Isto significa que havendo presença de outliers presente nos dados, este terá maior influência no aumento de RMSE, uma vez que este tem a característica da penalização de erros de maior magnitude.

Uma boa estratégia na avaliação do erro é a utilização em conjunto da análise de RMSE e MAE para que se verifiquem os casos em que haja uma grande discrepância ou crescimento do RMSE frente ao MAE, demonstrando mais uma vez que as métricas devem ser complementares e não analisadas individualmente (Willmott; Matsuura, 2005)

2.3.4 Erro percentual absoluto médio (MAPE)

O MAPE se trata de um erro percentual absoluto médio, sendo uma medição relativa a partir de valores absolutos, não sendo, portanto, levados em conta o contrabalanceamento de valores positivos e negativos, de modo a comparação da precisão entre métodos que usam séries temporais.

Essa métrica é difundida por ser intuitiva quando se fazem avaliações em termos de erros relativos. Sua utilização é muito relevante, por exemplo, em finanças onde em valores relativos se medem ganhos e perdas (De Myttenaere *et al.*, 2016).

3 METODOLOGIA

Inicialmente, foram adquiridos dados de velocidade do vento e irradiância solar e determinados os intervalos para os conjuntos de teste e treinamento. Para os dados de velocidade do vento e irradiância solar, num período de medição de 2007 a 2010, os primeiros três anos foram utilizados como conjunto de dados de treinamento e o último ano como conjunto de testes.

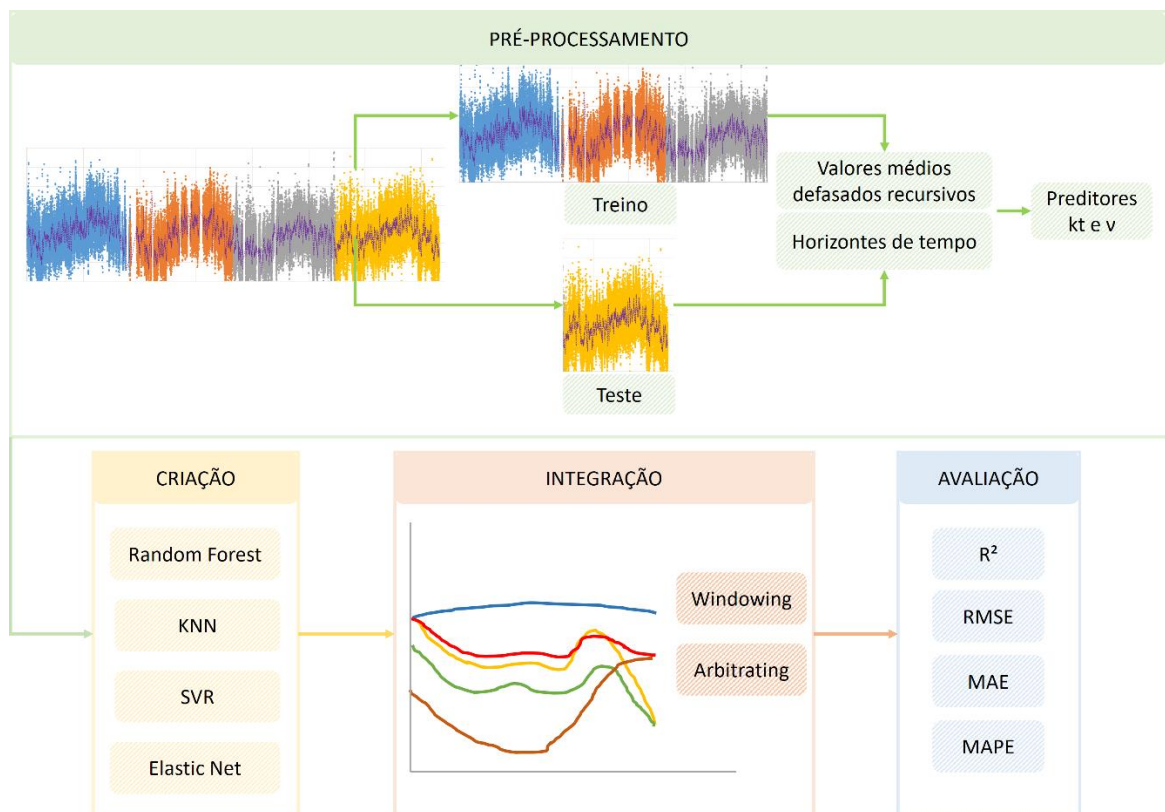
De modo a permitir a avaliação do desempenho dos modelos de previsão testados, assim como dos métodos de *ensemble* dinâmico, este estudo desenvolveu um código computacional em Python para avaliar os valores de saída obtidos pelos conhecidos métodos de previsão de aprendizado de máquina: *Random forest (RF)*, *k-nearest neighbours (kNN)*, *support vector regression (SVR)* e *elastic net (EN)*. Para cada um dos métodos foram avaliados os melhores parâmetros de desempenho (menor raiz do erro quadrático médio (RMSE)).

Logo após a etapa de aquisição e determinação dos parâmetros ótimos para cada um dos modelos, foram executados os métodos de *windowing* e *arbitrating* de conjuntos dinâmicos, dos quais também foram obtidos valores de métricas de desempenho: coeficiente de determinação (R^2), raiz do erro quadrático médio (RMSE), erro médio absoluto (MAE) e erro percentual médio absoluto (MAPE). Esses valores foram comparados para avaliar a eficiência dos métodos de conjunto dinâmico em comparação com outros modelos independentes.

Também foi avaliada a variação do parâmetro λ para *windowing*, que é o comprimento utilizado para a extensão dos valores considerados na previsão dos dados.

A metodologia utilizada pode ser visualizada na Figura 1. No pré-processamento dos dados, foi aplicada uma abordagem recursiva de valores médios defasados para séries temporais kt e v : esta característica é dada pelo vetor $L(t)$ com componentes calculados através da Equação 2.

Figura 1 - Fluxograma de dados para a metodologia aplicada



Fonte: Elaborada pelo autor.

3.1 Aquisição de Dados

Neste trabalho foram utilizados dois tipos de dados para a realização da análise, os quais foram adquiridos em estações solarimétricas e anemométricas localizadas em Petrolina – PE. Os dados foram recolhidos a partir da rede SONDA (Sistema de Organização de Dados Ambientais) (INPE, 2012) que foi uma colaboração conjunta entre diversas instituições e foi criada para a implementação de infraestruturas físicas e de recursos humanos, visando levantar e melhorar a base de dados de energia solar e recursos de energia eólica no Brasil.

O tempo de amostragem utilizado neste estudo foi de 10 minutos, e a duração da coleta de dados foi de janeiro de 2007 a dezembro de 2010. As informações detalhadas sobre os dados

da estação solarimétrica e anemométrica são mostradas na Tabela 1, onde IM (min) é o “intervalo de medição” e a duração da coleta de dados é apresentada como PM, “período medido”. Sua localização no mapa é mostrada na Figura 2.

Figura 2 - Mapa do Nordeste do Brasil. O local de medição de Petrolina está em destaque (Google, 2023)



Fonte: Elaborada pelo autor.

Tabela 1 - Coordenadas geográficas, altitude em relação ao nível do mar, intervalos de medição e períodos de medição dos dados foram coletados na estação Petrolina. IM e PM significam, respectivamente, “intervalo de medição” e “período de medição”.

Tipo	Lat. (°)	Long. (°)	Alt. (m)	IM (min)	PM
Anemométrica	09° 04' 08" S	40° 19' 11" O	387	10	01/01/2007 a 12/12/2010
Solarimétrica					01/01/2010 a 12/12/2010

Fonte: Elaborada pelo autor.

A região de Petrolina é classificada como zona climática BSh Koppen (Peel *et al.*, 2007). Existem diferenças consideráveis no ciclo anual entre a radiação solar e o vento. A velocidade média do vento e a irradiância solar em Petrolina sofrem variações sazonais significativas ao longo do seu ciclo anual. O intervalo mais ventoso do ano ocorre de maio a novembro, com velocidades médias de vento acima de 5,4 m/s. O mês com ventos mais fortes é agosto, com velocidade média horária do vento de 6,7 m/s. O período com menor volume de ventos do ano

é de novembro a maio. O mês com ventos mais calmos é março, com velocidade média horária de 4,1 m/s.

O período de maior radiação solar do ano é de setembro a novembro, com média diária superior a 7,2 kWh/m², sendo outubro o pico com média de 7,5 kWh/m². O período com menor radiação solar do ano é de maio a julho, com média diária de 6,1 kWh/m², sendo junho o mês com menor radiação solar, com média de 5,7 kWh/m².

3.1.1 Dados de velocidade do vento

A velocidade do vento foi obtida em m/s a partir de uma estação meteorológica, que possui sensores anemométricos nas altitudes de 25 m e 50 m do solo. A altitude mais alta foi escolhida para este estudo, tanto para reduzir os efeitos do terreno quanto para estar mais próxima das altitudes atualmente praticadas para turbinas eólicas (Landberg *et al.*, 2003)

3.1.2 Dados de irradiância solar

Os dados de irradiância horizontal global (GHI) adquiridos da estação solarimétrica foram utilizados neste estudo, e o coeficiente de céu claro foi considerado, a fim de remover a dependência da massa de ar nos valores de irradiância que atingem os sensores (KASTEN *et al.*, 1980), através da utilização do fator céu claro (I_{cs}) (Ineichen *et al.*, 2022) utilizando o modelo de ajuste polinomial (Marquez *et al.*, 2013). O trabalho (Rocha *et al.*, 2022) obteve resultados promissores no mesmo banco de dados usando dois modelos de estimativa de aprendizado de máquina para (GHI).

Para obter dados de irradiância independentes das variações da massa de ar, utilizou-se k_t , que é definido pela razão entre o valor de irradiância horizontal global (GHI) (I) e a irradiância de céu claro (I_{cs}), conforme mostrado na Equação (1).

$$k_t = \frac{I}{I_{cs}} \quad (2)$$

3.2 Modelos de previsão de aprendizado de máquina e parâmetros de métodos de ensemble dinâmico

Na etapa de treinamento dos dados, foi utilizado o *GridSearch* com validação cruzada de 5 vezes. Os parâmetros de busca são mostrados na Tabela 2.

Tabela 2 - Parâmetros de pesquisa e valores de *grid* aplicados aos métodos testados.

Método	Parâmetro	Grid Values
RF	maxdepth	[2, 5, 7, 9, 11, 13, 15, 21, 35]
KNN	nearest neighbours k	$1 \leq k \leq 50$, k integer
SVR	penalty term C	[0.1, 1, 10, 100, 1000]
	coefficient λ	[1, 0.1, 0.01, 0.001, 0.0001]
EN	regularization term λ	[1, 0.1, 0.01, 0.001, 0.0001]
Windowing	λ	[1, 3, 6, 12, 25, 50, 100]
Arbitrating		*

*: Devido à utilização de uma metodologia meta-heurística, o parâmetro inicial não foi necessário.
Fonte: Elaborada pelo autor.

3.3 Critério comparativo por métricas de performance

Como o objetivo deste trabalho é avaliar o desempenho de métodos de conjuntos dinâmicos em relação a outros métodos, métricas de desempenho tiveram que ser determinadas para permitir isso. As métricas utilizadas foram as das Equações (3)–(6).

- Coeficiente de determinação (R^2)

$$R^2 = 1 - \frac{\sum_{i=1}^N (y_i - \hat{y}_i)^2}{\sum_{i=1}^N (y_i - \bar{y})^2} \quad (2)$$

- Raiz do erro quadrático médio (RMSE)

$$RMSE = \sqrt{\frac{\sum_{i=1}^N (y_i - \hat{y}_i)^2}{n}} \quad (3)$$

- Erro médio absoluto (MAE)

$$MAE = \frac{1}{N} \sum_{i=1}^N |\hat{y}_i - y_i| \quad (4)$$

- Erro percentual médio absoluto (MAPE)

$$MAPE = \frac{1}{n} \sum_{i=1}^n \left| \frac{y_i - \hat{y}_i}{y_i} \right| \quad (5)$$

4 RESULTADOS E DISCUSSÃO

Esta seção discute os resultados gerados neste trabalho. Ele se concentra na análise de métricas de eficiência para os métodos de aprendizado de máquina empregados. Esta análise determina quais métodos/parâmetros obtêm o melhor desempenho na aplicação dos dados de velocidade do vento e irradiância solar.

4.1 Previsões de velocidade de vento

Durante a busca pelos métodos de melhor desempenho, os parâmetros otimizados para cada um dos métodos testados precisaram ser identificados. Isso permitiu a elaboração do conjunto dinâmico, que foi construído a partir da fusão dos resultados de melhor desempenho em cada etapa de tempo e para todos os métodos em questão. Os parâmetros ótimos encontrados para cada um dos horizontes temporais são apresentados na Tabela 3.

Tabela 3 - Melhores parâmetros para cada método de aprendizado de máquina.

Método	Parâmetro	t+10	t+20	t+30	t+60
RF	best max depth	7	7	7	7
	best n estimators	20	20	20	20
KNN	best n neighbors	49	49	49	49
SVR	best C	1	1	1	1
	best epsilon	0,1	0,1	1	0,1
EN	best ll_ratio	1	1	1	1

Fonte: Elaborada pelo autor.

As avaliações de eficiência para cada um dos métodos de previsão foram baseadas em avaliações de métricas de desempenho para cada horizonte temporal em estudo (t+10, t+20, t+30 e t+60). Inicialmente, para todos os horizontes temporais, o *windowing* provou ser o método mais eficiente. Em seguida, foi realizada uma avaliação de ajuste fino com base na variação do parâmetro de *windowing* para avaliar sua influência no desempenho. A predominância de melhor desempenho para *windowing* em todos os horizontes de tempo e suas comparações pode ser observada na Tabela 4 e na Figura 3.

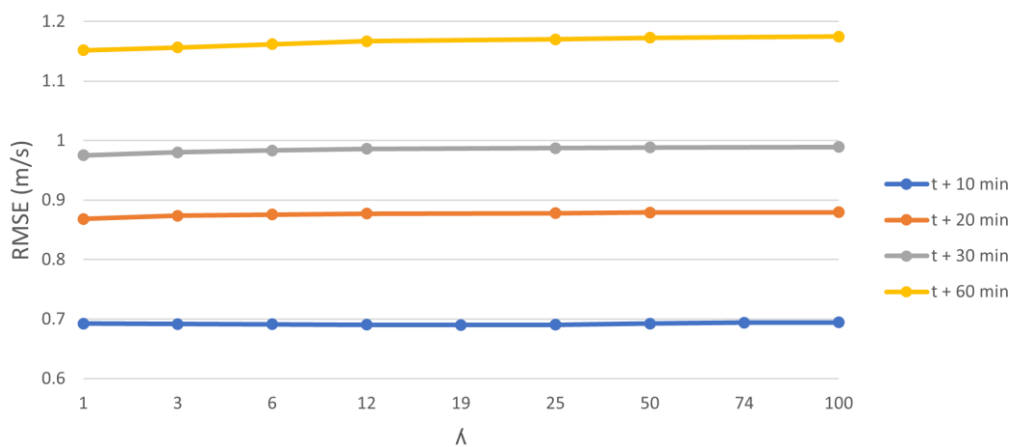
Elastic net é um modelo de regressão linear penalizado que é uma combinação de regressão Ridge e LASSO em um único algoritmo e usa *best_ll_ratio* como parâmetro de penalidade durante a etapa de treinamento, sendo 0 para Ridge e 1 valor para regressão LASSO. Na Tabela 3, o parâmetro obteve o valor 1, o que significa que a regressão LASSO foi utilizada em sua totalidade.

Tabela 4 - Comparação dos valores de RMSE (m/s), utilizando diferentes métodos para diferentes horizontes de tempo e *windowing* da variação dos parâmetros λ . Os melhores resultados para cada horizonte temporal estão em negrito.

Horizonte de Tempo	λ	RF	KNN	SVR	EN	Windowing	Arbitrating
t+10 min	1					0,692	
	3					0,691	
	6					0,691	
	12					0,690	
	19	0,694	0,710	0,693	0,698	0,690	0,694
	25					0,690	
	50					0,692	
	74					0,694	
	100					0,694	
t+20 min	1					0,868	
	3					0,873	
	6					0,875	
	12	0,883	0,893	0,883	0,885	0,876	0,883
	25					0,878	
	100					0,879	
t+30 min	1					0,974	
	3					0,980	
	6					0,983	
	12	0,994	0,998	0,991	0,996	0,985	0,990
	25					0,987	
	100					0,989	
t+60 min	1					1,151	
	3					1,156	
	6					1,161	
	12	1,180	1,195	1,177	1,182	1,166	1,181
	25					1,169	
	100					1,174	

Fonte: Elaborada pelo autor.

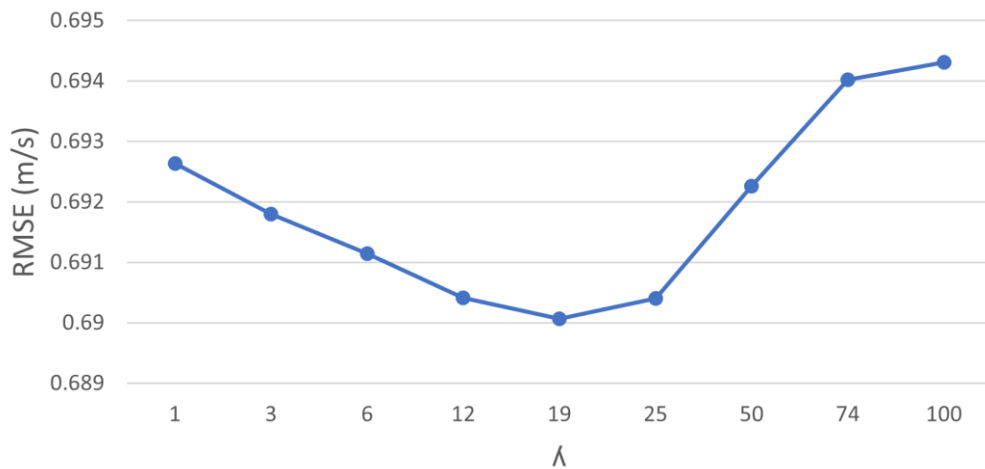
Figura 3 - Influência da variação do parâmetro λ de *windowing* no RMSE para diferentes horizontes de tempo na análise de dados de velocidade do vento para todos os horizontes de tempo estudados.



Fonte: Elaborada pelo autor.

Assim como na avaliação por meio do RMSE, também foram avaliados os valores de R^2 , MAE e MAPE. Uma vez encontrado o melhor desempenho para o método *windowing*, foi realizada uma análise aprofundada com base na variação do seu parâmetro λ para avaliar a influência no seu desempenho interno. Como o horizonte temporal que apresentou melhor desempenho foi $t+10$, este foi o foco da análise, conforme mostram as Figuras 4–7. Os dados detalhados para todos os horizontes são apresentados nas Tabelas 5–7.

Figura 4 - A influência do parâmetro λ de *windowing* no valor RMSE para o horizonte de tempo $t+10$



Fonte: Elaborada pelo autor.

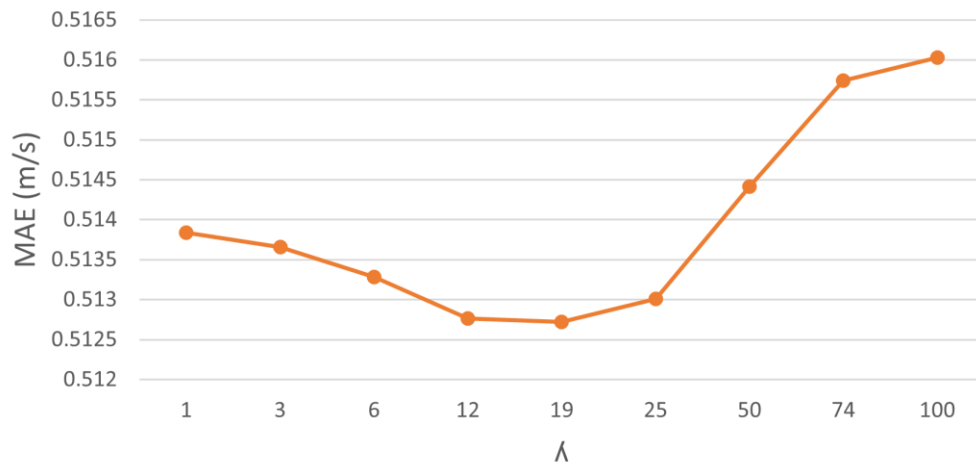
Tabela 5 - Comparação dos valores do MAE (m/s), utilizando diferentes métodos em diferentes horizontes de tempo e *windowing* da variação do parâmetro λ . Os melhores resultados para cada horizonte temporal estão em negrito.

Horizonte de Tempo	λ	RF	KNN	SVR	EN	Windowing	Arbitrating
t+10 min	1					0,513	
	3					0,513	
	6					0,513	
	12					0,512	
	19	0,515	0,532	0,514	0,518	0,512	0,517
	25					0,513	
	50					0,514	
	74					0,515	
	100					0,516	
t+20 min	1					0,646	
	3					0,651	
	6					0,653	
	12	0,658	0,668	0,660	0,659	0,654	0,659
	25					0,655	
	50					0,656	
	100					0,656	
t+30 min	1					0,725	
	3					0,731	
	6	0,742	0,747	0,741	0,743	0,734	0,740
	12					0,736	
	25					0,737	

	50					0,738	
	100					0,739	
	1					0,867	
	3					0,872	
	6					0,878	
t+60 min	12	0,894	0,907	0,891	0,895	0,883	0,895
	25					0,885	
	50					0,888	
	100					0,889	

Fonte: Elaborada pelo autor.

Figura 5 - Influência do parâmetro λ de *windowing* no valor MAE para o horizonte de tempo t+10.



Fonte: Elaborada pelo autor.

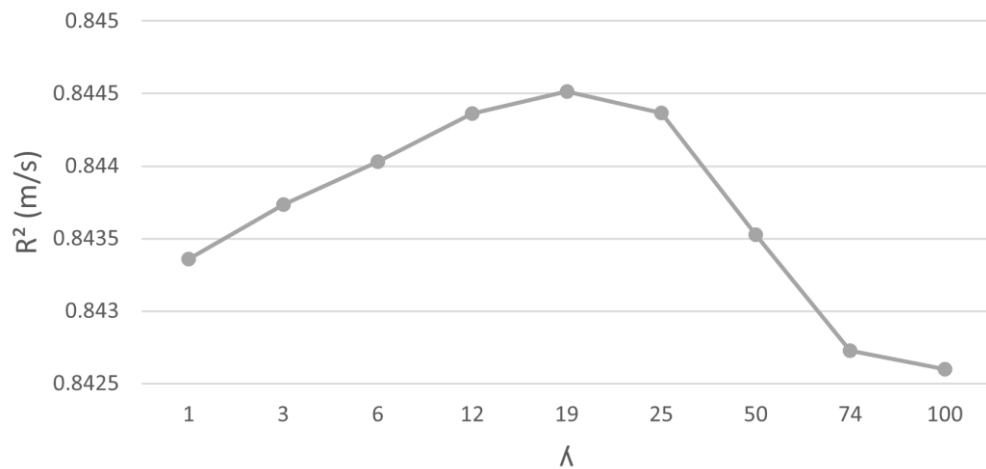
Tabela 6: Comparação dos valores de R^2 (m/s), utilizando diferentes métodos em diferentes horizontes de tempo e da variação do parâmetro λ de *windowing*. Os melhores resultados para cada horizonte de tempo estão em negrito

Horizonte de Tempo	λ	RF	KNN	SVR	EN	Windowing	Arbitrating
t+10 min	1					0,843	
	3					0,843	
	6					0,844	
	12					0,844	
	19	0,842	0,835	0,842	0,840	0,844	0,842
	25					0,844	
	50					0,843	
	74					0,842	
	100					0,842	
t+20 min	1					0,753	
	3					0,750	
	6					0,749	
	12	0,745	0,739	0,744	0,743	0,748	0,745
	25					0,748	
	100					0,747	
t+30 min	1					0,689	
	3					0,686	
	6	0,676	0,674	0,679	0,675	0,684	0,679
	12					0,682	
	25					0,681	

		<u>50</u>				<u>0,681</u>	
		100				0,680	
		1				0,566	
		3				0,563	
		6				0,559	
t+60 min	0,544	0,533	0,546	0,542		0,555	0,543
		25				0,552	
		50				0,550	
		100				0,549	

Fonte: Elaborada pelo autor.

Figura 6 - Influência do parâmetro λ de *windowing* no valor R^2 para o horizonte de tempo t+10.



Fonte: Elaborada pelo autor.

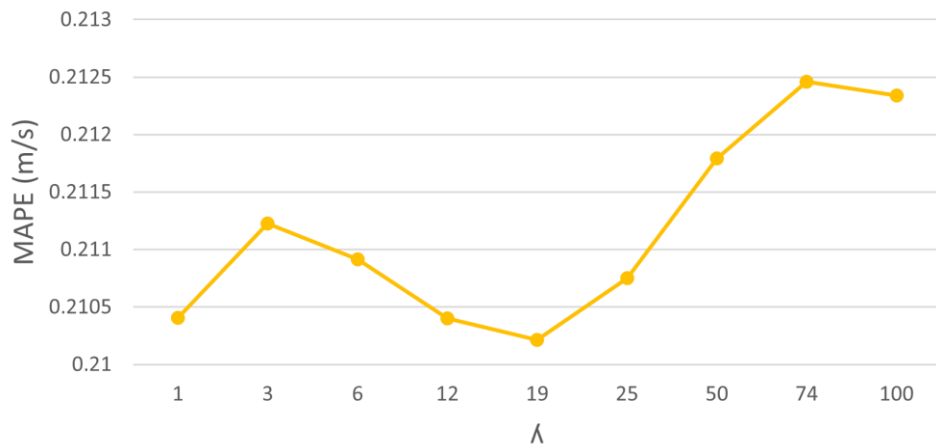
Tabela 7 - Comparação dos valores de MAPE (m/s), utilizando diferentes métodos em diferentes horizontes de tempo e da variação do parâmetro λ de *windowing*. Os melhores resultados para cada horizonte de tempo estão em negrito

Horizonte de Tempo	λ	RF	KNN	SVR	EN	Windowing	Arbitrating
t+10 min	1					0,210	
	3					0,211	
	6					0,210	
	12					0,210	
	19	0,212	0,253	0,202	0,218	0,210	0,216
	25					0,210	
	50					0,211	
	74					0,212	
	100					0,212	
	t+20 min	1					0,312
3						0,315	
6						0,316	
12		0,315	0,338	0,341	0,312	0,317	0,325
25						0,319	
50						0,319	
100						0,321	
t+30 min	1					0,367	
	3					0,369	
	6	0,380	0,397	0,375	0,370	0,372	0,384
	12					0,372	
	25					0,373	

		<u>50</u>				<u>0,373</u>	
		100				0,375	
		1				0,505	
		3				0,507	
		6				0,511	
t+60 min		12	0,523	0,535	0,517	0,512	0,524
		25				0,514	
		50				0,515	
		100				0,518	

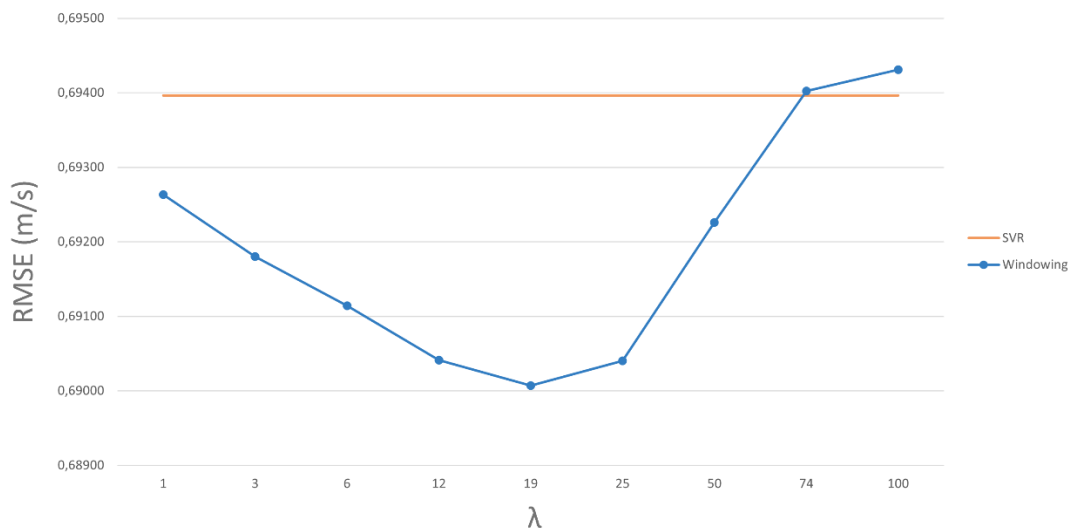
Fonte: Elaborada pelo autor.

Figura 7 - Influência do parâmetro λ de *windowing* no valor MAPE para o horizonte de tempo t+10.



Fonte: Elaborada pelo autor.

Figura 8 - Efeito da variação do parâmetro λ no desempenho do método *windowing*. O resultado SVR é mostrado para referência.



Fonte: Elaborada pelo autor.

Quando se verificou a influência do parâmetro λ no desempenho do método de *windowing*, constatou-se a partir de $\lambda = 74$ que ele não é mais o método mais eficiente, e o SVR

passa a ser o melhor, devido ao seu menor valor de RMSE. É importante destacar que o melhor valor de desempenho para o método de *windowing*, que é o melhor desempenho geral, foi encontrado para $\lambda = 19$. A comparação de desempenho entre os dois métodos pode ser vista na Figura 8.

4.2 Previsões de irradiância

Durante a busca pelos métodos de melhor desempenho, os parâmetros otimizados de cada um desses métodos precisaram ser conhecidos para permitir a elaboração do conjunto dinâmico, que é construído a partir da fusão dos resultados de melhor desempenho em cada instante e para cada uma das abordagens em questão. Os parâmetros ótimos para cada horizonte temporal são apresentados na Tabela 8.

Tabela 8 - Melhores parâmetros para cada método de aprendizado de máquina.

Método	Parâmetro	t+10	t+20	t+30	t+60
RF	best_max_depth	5	5	5	5
	best_n_estimators	20	20	20	20
KNN	best_n_neighbors	37	37	49	48
SVR	best_C	0,1	0,1	0,1	0,1
	best_epsilon	0,1	0,1	0,1	0,1
EN	best_l1_ratio	1	1	1	1

Fonte: Elaborada pelo autor.

A avaliação da eficiência de cada um dos métodos de previsão da irradiância solar foi baseada em métricas de desempenho para cada horizonte temporal em estudo (t+10, t+20, t+30 e t+60). Novamente, o *windowing* mostrou-se o método mais eficiente para todos os horizontes de tempo, sendo o melhor método encontrado para o horizonte de tempo t+10, tendo o menor valor de RMSE, utilizando suas parametrizações com $\lambda = 50$ inicialmente. Em seguida, foi realizado um ajuste fino com base na variação do parâmetro de *windowing* para avaliar sua influência no desempenho. A predominância de melhor desempenho para *windowing* em todos os horizontes de tempo e suas comparações podem ser observadas na Tabela 9 e na Figura 9.

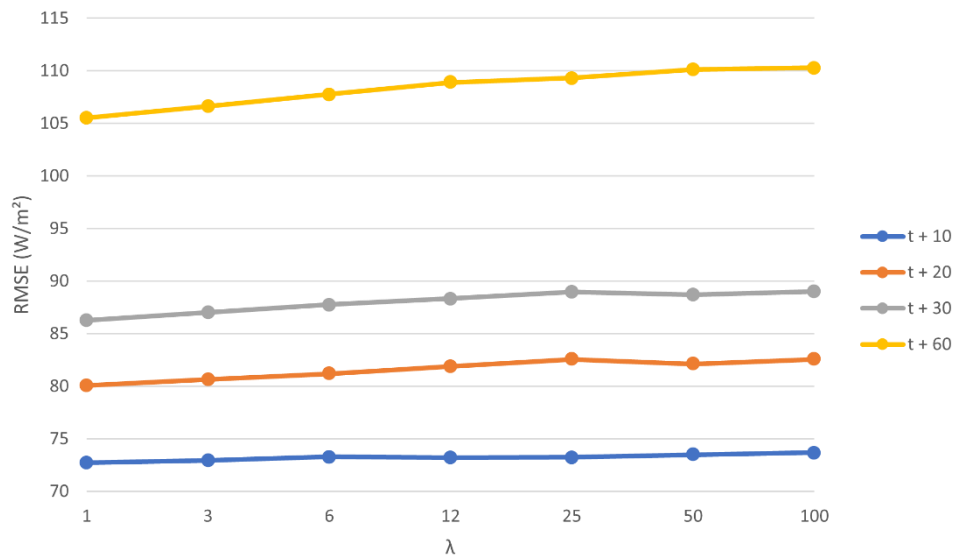
Tabela 9 - Comparação dos valores de RMSE (W/m²), utilizando diferentes Métodos em diferentes horizontes de tempo variando o parâmetro λ de *windowing*. Os melhores resultados para cada Horizonte de Tempo estão em negrito.

Horizonte de Tempo	λ	RF	KNN	SVR	Elastic Net	Windowing	Arbitrating
t+10 min	1					72,731	
	3					72,932	
	6					73,293	
	12	75,020	75,260	74,190	74,980	73,210	74,010
	25					73,246	
	50					73,480	
	100					73,693	
t+20 min	1	90,940	83,500	84,450	84,530	80,070	83,190

	3					80,630	
	6					81,190	
	12					81,870	
	25					82,560	
	50					82,110	
	100					82,570	
	1					86,250	
	3					87,000	
	6					87,750	
t+30 min	12	90,150	90,500	91,490	93,490	88,330	89,700
	25					88,950	
	50					88,700	
	100					89,010	
	1					105,510	
	3					106,620	
	6					107,760	
t+60 min	12	112,050	112,130	112,760	118,080	108,890	111,130
	25					109,320	
	50					110,120	
	100					110,300	

Fonte: Elaborada pelo autor.

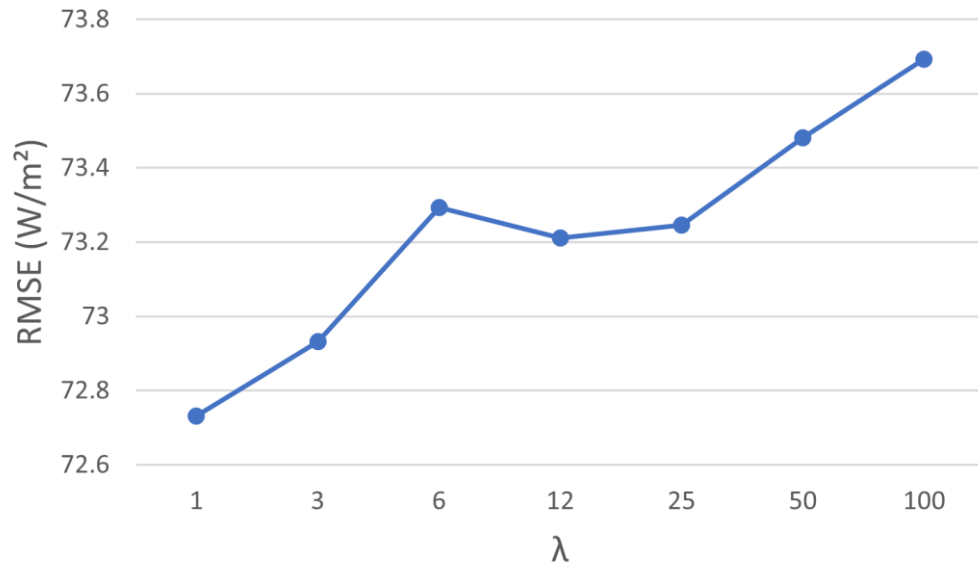
Figura 9 - A influência da variação do parâmetro λ de *windowing* Parâmetro no RMSE para todos os horizontes de tempo estudados na análise de dados de irradiância solar.



Fonte: Elaborada pelo autor.

Assim como na avaliação pelo RMSE, também foram analisados os valores de R^2 , MAE e MAPE. Depois de encontrado o melhor desempenho para o método de *windowing*, foi realizada uma análise aprofundada com base na variação do seu parâmetro λ para avaliar a influência no desempenho interno. Como o horizonte temporal que apresentou melhor desempenho foi t+10, este foi o foco da análise, conforme mostram as Figuras 9–12. Os dados detalhados para todos os horizontes de tempo testados são mostrados nas Tabelas 10–12.

Figura 10 - Influência do parâmetro λ de *windowing* no valor de RMSE no horizonte de tempo t+10.



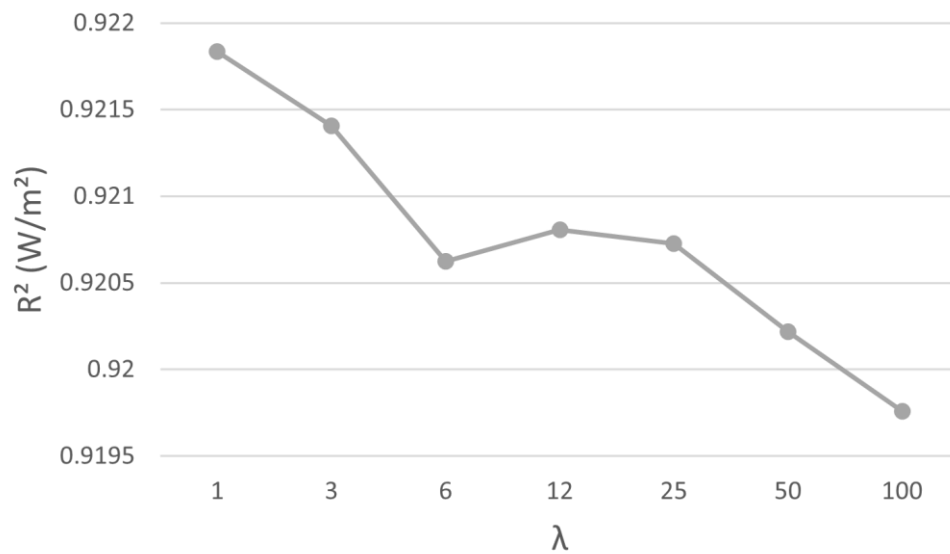
Fonte: Elaborada pelo autor.

Tabela 10 - Comparação dos valores de R^2 (W/m²), utilizando diferentes métodos em diferentes horizontes de tempo variando o parâmetro λ de *windowing*. Os melhores resultados para cada horizonte de tempo estão em negrito.

Horizonte de Tempo	λ	RF	KNN	SVR	Elastic Net	Windowing	Arbitrating
t+10 min	1					0,921	
	3					0,921	
	6					0,920	
	12	0,920	0,920	0,920	0,920	0,920	0,920
	25					0,920	
	50					0,920	
	100					0,919	
t+20 min	1					0,910	
	3					0,910	
	6					0,900	
	12	0,880	0,900	0,900	0,900	0,900	0,900
	25					0,900	
	50					0,900	
	100					0,900	
t+30 min	1					0,890	
	3					0,890	
	6					0,890	
	12	0,880	0,880	0,880	0,870	0,890	0,880
	25					0,890	
	50					0,880	
	100					0,890	
t+60 min	1					0,850	
	3					0,840	
	6					0,840	
	12	0,830	0,830	0,820	0,512	0,840	0,830
	25					0,830	
	50					0,830	
	100					0,830	

Fonte: Elaborada pelo autor.

Figura 11 - Influência da variação do parâmetro λ de *windowing* no valor de R^2 no horizonte de tempo $t+10$.



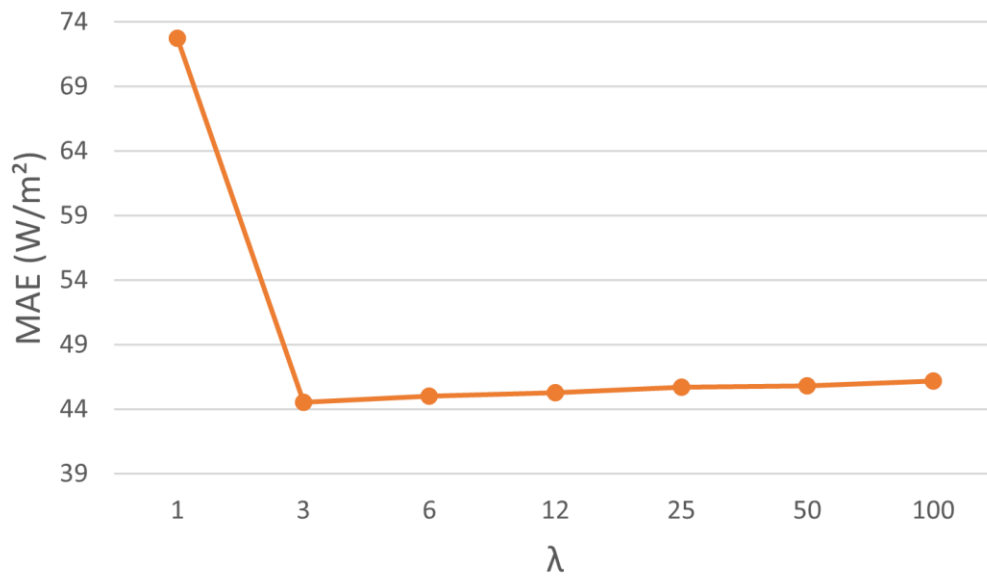
Fonte: Elaborada pelo autor.

Tabela 11 - Comparação dos valores de MAE (W/m²), utilizando diferentes métodos em diferentes horizontes de tempo variando o parâmetro λ de *windowing*. Os melhores resultados para cada horizonte de tempo estão em negrito.

Horizonte de Tempo	λ	RF	KNN	SVR	Elastic Net	Windowing	Arbitrating
t+10 min	1					72,731	
	3					44,523	
	6					45,007	
	12	48,290	48,470	44,160	49,310	45,277	46,240
	25					45,679	
	50					45,791	
	100					46,166	
t+20 min	1					52,530	
	3					53,310	
	6					54,120	
	12	65,190	55,630	59,670	58,860	55,270	56,200
	25					56,880	
	50					55,590	
	100					56,790	
t+30 min	1					58,140	
	3					59,020	
	6					59,910	
	12	62,090	61,580	64,770	67,130	60,850	60,910
	25					61,340	
	50					61,840	
	100					61,510	
t+60 min	1					74,590	
	3					75,920	
	6					77,110	
	12	81,280	79,840	81,440	89,070	78,470	79,800
	25					79,080	
	50					79,480	
	100					79,630	

Fonte: Elaborada pelo autor.

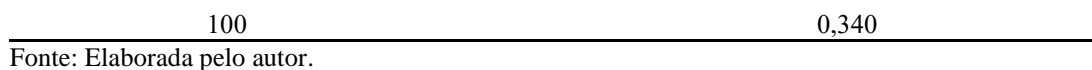
Figura 12 - Influência da variação do parâmetro λ de *windowing* no valor de MAE no horizonte de tempo $t+10$.



Fonte: Elaborada pelo autor.

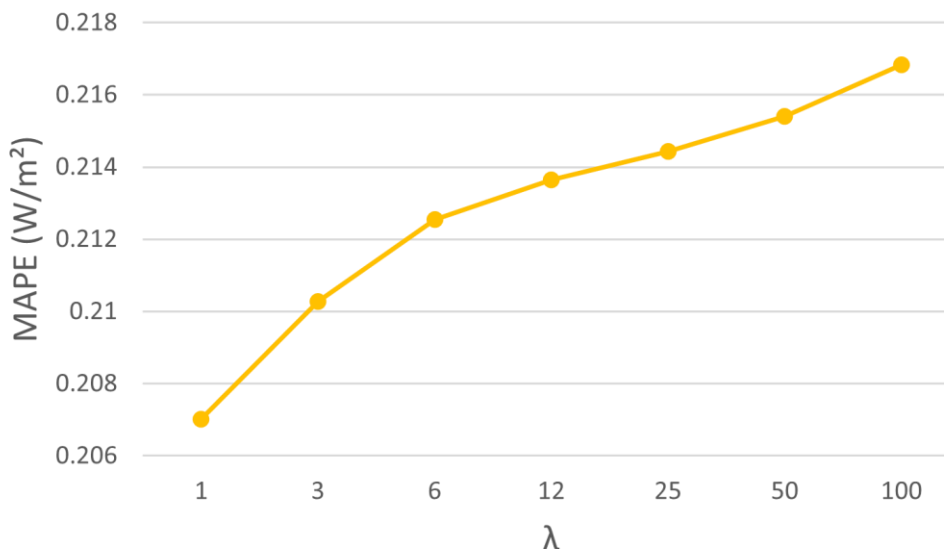
Tabela 12 - Comparação dos valores de MAPE (W/m²), utilizando diferentes métodos em diferentes horizontes de tempo variando o parâmetro λ de *windowing*. Os melhores resultados para cada horizonte de tempo estão em negrito.

Horizonte de Tempo	λ	RF	KNN	SVR	Elastic Net	Windowing	Arbitrating
t+10 min	1					0,207	
	3					0,210	
	6					0,212	
	12	0,220	0,240	0,210	0,230	0,213	0,220
	25					0,214	
	50					0,215	
	100					0,216	
	t+20 min	1					0,250
3						0,250	
6						0,260	
12		0,320	0,280	0,280	0,270	0,260	0,270
25						0,270	
50						0,260	
100						0,270	
t+30 min		1					0,270
	3					0,280	
	6					0,280	
	12	0,290	0,300	0,290	0,330	0,280	0,290
	25					0,290	
	50					0,290	
	100					0,290	
	t+60 min	1					0,320
3						0,320	
6						0,330	
12		0,340	0,350	0,340	0,547	0,330	0,340
25						0,340	
50						0,340	



Fonte: Elaborada pelo autor.

Figura 13 - Influência da variação do parâmetro λ de *windowing* no valor de MAPE no horizonte de tempo $t+10$.



Fonte: Elaborada pelo autor.

Alguns autores aplicaram *Elastic Net* em combinações variantes no tempo (Nikodinoska *et al.*, 2022), usando RMSE como métrica de desempenho. Descobriram que, para as previsões fotovoltaicas, obtiveram previsões 13,4% mais precisas do que a média simples e para a previsão do vento, obteve previsões 6,1% melhores. Em Carneiro (2022), foi estudado um método *ensemble* que utilizou MAPE como métrica comparativa de eficiência para dados de velocidade do vento com valor de 9,345%, e solar com 7,186%, que se mostrou o mais eficiente. Neste estudo foram obtidas melhorias de desempenho para o método mais eficiente (*windowing*) em comparação ao segundo mais eficiente para velocidade do vento de 0,56% e, para irradiância solar, 1,86%.

4.3 Comparação de resultados com a literatura

O desempenho da abordagem de *windowing* foi comparado com outros modelos de previsão de vento encontrados na literatura. É importante ressaltar que uma comparação direta entre diferentes modelos preditivos não é uma tarefa fácil, uma vez que cada abordagem aplicada possui seus próprios objetivos, hiperparâmetros e dados de entrada (Santos *et al.*, 2023). Para facilitar a comparação com os resultados encontrados na literatura, a Tabela 13 compila os resultados apresentados anteriormente para o modelo de *windowing* proposto. Os resultados

encontrados na literatura para previsão da velocidade do vento são compilados e apresentados na Tabela 14, onde RMSE e MAE estão em m/s.

Tabela 13 - Compilação dos resultados obtidos pelo método *windowing* em diferentes horizontes de tempo.

Métrica	Horizonte de Tempo	Velocidade do Vento	GHI
RMSE	t+10	0,690 m/s	72,731 W/m ²
	t+20	0,868 m/s	80,076 W/m ²
	t+30	0,974 m/s	86,254 W/m ²
	t+60	1,151 m/s	105,511 W/m ²
nRMSE	t+10	0,098	0,073
	t+20	0,124	0,080
	t+30	0,139	0,086
	t+60	0,164	0,106
R ²	t+10	0,844m/s	0,921 W/m ²
	t+20	0,753 m/s	0,911 W/m ²
	t+30	0,689 m/s	0,894 W/m ²
	t+60	0,566 m/s	0,853 W/m ²
MAE	t+10	0,512 m/s	44,523 W/m ²
	t+20	0,646 m/s	52,532 W/m ²
	t+30	0,725 m/s	58,141 W/m ²
	t+60	0,867 m/s	74,592 W/m ²
MAPE	t+10	0,210 m/s	0,207 W/m ²
	t+20	0,312 m/s	0,253 W/m ²
	t+30	0,367 m/s	0,271 W/m ²
	t+60	0,505 m/s	0,324 W/m ²

Fonte: Elaborada pelo autor

Tabela 14 - Compilação de resultados da literatura para previsão da velocidade do vento.

Modelo	Valor da métrica	Autor
GNN SAGE GAT	RMSE	Oliveira Santos et al. (2023)
	0.638 for t+60 forecasting horizon	
	MAE	
ED-HGNDO-BiLSTM	0.458 for t+60 forecasting horizon	Neshat et al. (2021)
	RMSE	
	0.696 average for t+10 forecasting horizon	
	1.445 average for t+60 forecasting horizon	
	MAE	
	0.717 average for t+10 forecasting horizon	
	0.953 average for t+60 forecasting horizon	
	MAPE	
0.590 average for t+10 forecasting horizon		
Statistical model for wind speed forecasting	9.769 average for t+60 forecasting horizon	Dowell et al. (2014)
	RMSE	
Hybrid wind speed forecasting model using area division (DAD) Method and a deep learning neural network	1.090 for t+60 forecasting horizon	Liu et al. (2021)
	RMSE	
	0.291 average for t+10 forecasting horizon	
	0.355 average for t+30 forecasting horizon	
	0.426 average for t+60 forecasting horizon	

	MAE	
	0.221 average for t+10 forecasting horizon	
	0.293 average for t+30 forecasting horizon	
	0.364 average for t+60 forecasting horizon	
	RMSE	
	0.547 for t+10 forecasting horizon	
	0.802 for t+20 forecasting horizon	
	0.895 for t+30 forecasting horizon	
Hybrid model CNN-LSTM	1.114 for t+60 forecasting horizon	Zhu et al. (2020)
	MAPE	
	4.385 for t+10 forecasting horizon	
	6.023 for t+20 forecasting horizon	
	7.510 for t+30 forecasting horizon	
	11.127 for t+60 forecasting horizon	

Fonte: Elaborada pelo autor.

Analisando os resultados para referência (Santos *et al.*, 2023), nos quais a velocidade do vento foi prevista na Holanda usando uma abordagem de conjunto mesclando teoria de grafos e aprendizado profundo baseado em atenção, podemos observar que o modelo *windowing* proposto não é capaz de superar os resultados nem para RMSE nem MAE para o horizonte de previsão t+60. A diferença acentuada entre esses dois modelos pode ser explicada porque o modelo GNN SAGE GAT, sendo desenvolvido para lidar com estruturas de dados semelhantes a grafos, se destaca na recuperação de relações espaço-temporais complexas subjacentes ao conjunto de dados, melhorando drasticamente sua capacidade de previsão quando comparado com outros modelos de *Machine Learning* (ML) e *Deep Learning* (DL).

Em Neshat *et al.* (2021), os autores propuseram uma previsão de vento para um local na Suécia, com um modelo baseado em uma rede neural recorrente bidirecional, uma técnica de decomposição hierárquica e um algoritmo de otimização. Quando comparado com os seus resultados, o modelo de *windowing* proposto neste artigo oferece uma melhoria em relação aos resultados de referência para o horizonte de previsão t+10 em 1% e em 20% para t+60. Quando analisados MAE e MAPE, o *windowing* indica melhoria em relação a essas métricas para t+10 e t+60, aumentando em 28% o valor do MAE para t+10 e 9% para t+60. Em relação ao MAPE, a melhoria é de 64% para t+10 e 95% para t+60.

No trabalho de Liu (2021), outro modelo preditivo baseado em aprendizagem profunda foi proposto. Eles usaram uma abordagem híbrida composta por divisão de área de dados para extrair informações históricas da velocidade do vento e uma camada LSTM otimizada por meio de um algoritmo genético para processar o aspecto temporal do conjunto de dados para prever a velocidade do vento no Japão. Comparado com esta referência, o modelo *windowing* não apresentou melhorias na previsão da velocidade do vento. Contudo, a abordagem de *windowing*

oferece previsões competitivas para as janelas de tempo avaliadas, sendo da mesma ordem de grandeza das da referência.

No trabalho de Zhu (2020) os autores propuseram o emprego de outra arquitetura híbrida de previsão composta por modelos de aprendizagem profunda CNN e LSTM para estimativa da velocidade do vento nos EUA. Seus resultados, quando comparados com a metodologia de *windowing*, são muito semelhantes para todos os horizontes de previsão, mostrando que tanto o *windowing* quanto o CNN-LSTM oferecem bons resultados para estimativa da velocidade do vento para esses intervalos de tempo.

Em Dowell (2014), foi proposto um modelo estatístico para estimativa dos valores futuros da velocidade do vento na Holanda. Para o horizonte temporal t+60 disponível, observa-se que, mais uma vez, as velocidades do vento previstas para os modelos de referência e de *windowing* propostos são muito semelhantes, sugerindo ambos os modelos como ferramentas valiosas para a previsão da velocidade do vento. Para a previsão do GHI, os resultados encontrados na literatura são apresentados na Tabela 15.

Tabela 15 - Compilação de resultados para previsão de irradiação global horizontal (GHI)

Modelo	Valor da métrica	Autor
CNN-1D	RMSE (R^2)	Marinho <i>et al.</i> (2023)
	36.24 (0.98) for t+10 forecasting horizon	
	39.00 (0.98) for t+20 forecasting horizon 38.46 (0.98) for t+30 forecasting horizon	
MEMD-PCA-GRU	RMSE (R^2) 31.92 (0.99) for t+60 forecasting horizon	Gupta e Singh (2023)
Physical-based forecasting model	RMSE	Yang <i>et al.</i> (2022)
	75.91 for t+30 forecasting horizon 89.81 for t+60 forecasting horizon	
	MAE 48.85 for t+30 forecasting horizon 57.01 for t+60 forecasting horizon	
Physical-based forecasting model	RMSE 114.06 for t+60 forecasting horizon	Kallio-Meyers <i>et al.</i> (2020)
Deep learning transformer-based forecasting model	MAE	Liu <i>et al.</i> (2023)
	34.21 for t+10 forecasting horizon 43.64 for t+20 forecasting horizon	
	49.53 for t+30 forecasting horizon	

Fonte: Elaborada pelo autor.

No trabalho de Marinho (2023), um modelo autônomo de aprendizagem profunda CNN foi aplicado para estimar valores futuros de GHI nos EUA. Comparando os resultados de previsão do GHI obtidos via *windowing* com esta referência, observa-se que o modelo proposto não foi capaz de fornecer desempenho de previsão superior. No entanto, os resultados do *windowing* ainda são competitivos, uma vez que ambas as abordagens conseguiram atingir

valores elevados de coeficiente de determinação para todos os horizontes de previsão avaliados, com uma ligeira vantagem para o modelo de aprendizagem profunda.

Em Gupta e Singh (2023), os autores combinaram a análise *Principal Component Analysis* (PCA) com a decomposição do *Multivariate Empirical Mode Decomposition* (MEMD) e *Gated Recurrent Unit* (GRU) para prever o GHI na Índia. Em sua metodologia, o PCA extraiu as características mais relevantes do conjunto de dados após ele ter sido filtrado por meio do algoritmo MEMD. Por último, a irradiância futura foi estimada através do modelo de aprendizagem profunda do GRU. Comparado com a sua abordagem, o modelo de *windowing* não conseguiu melhorar a previsão do GHI dentro de uma janela de tempo $t+60$. Além disso, o modelo de referência MEMD-PCA-GRU forneceu um valor R^2 elevado de 99%, mostrando desempenho claramente superior ao modelo de conjunto proposto. Quando o presente modelo é comparado com os modelos de previsão de base física propostos em Yang e Gao (2022) e Kallio-Myers (2020), pode-se concluir que o *windowing* pode alcançar resultados semelhantes para horizontes temporais de $t+30$ e $t+60$.

Em Yang e Gao (2022), os autores usaram o método FY-4A-Heliosat de imagens de satélite para estimar o GHI na China. Embora o modelo de *windowing* não tenha conseguido melhorar a previsão do GHI para as janelas de tempo $t+30$ e $t+60$, o modelo proposto foi capaz de retornar resultados relevantes para estimativa de irradiância em ambos os casos.

O segundo modelo de base física proposto em Kallio-Myers (2020) foi aplicado para estimar o GHI na Finlândia. Em sua metodologia, o método Heliosat é novamente empregado, juntamente com dados meteorológicos geoestacionários de imagens de satélite. Em comparação com a abordagem proposta, o modelo de *windowing* pode melhorar a previsão do GHI para $t+60$ em 8%, proporcionando um avanço significativo na estimativa de irradiância. Em Liu e Zang (2023), os autores usaram a arquitetura de aprendizagem profunda de transformadores de última geração juntamente com imagens do céu (Vaswani *et al.*, 2017) para estimativa de GHI nos EUA. Analisando seus resultados e os fornecidos pelo método de *windowing*, observa-se que o modelo baseado em transformador atinge os melhores valores de previsão de GHI para RMSE em todas as janelas de tempo avaliadas.

Após a comparação da abordagem de *windowing* com modelos de referência encontrados na literatura, vê-se que a previsão da velocidade do vento é frequentemente competitiva e geralmente melhora a previsão da velocidade do vento para os horizontes de previsão avaliados. Os resultados para previsão da velocidade do vento usando o modelo *ensemble* corroboram os resultados encontrados na literatura, onde a abordagem *ensemble* frequentemente atinge previsões de última geração em aplicações de previsão de séries

temporais (Carneiro *et al.*, 2022; Peng *et al.*, 2020; Abdellatif *et al.*, 2022; Wu *et al.*, 2021). Seu melhor desempenho vem da combinação de modelos preditivos mais fracos para melhorar sua capacidade geral de previsão, reduzindo também a variância do modelo conjunto (Ghojogh *et al.*, 2023; Chen *et al.*, 2016).

No entanto, a abordagem dinâmica proposta enfrentou maior dificuldade na determinação dos valores futuros do GHI. Isto pode ser uma indicação de que a previsão de irradiância é um fenômeno natural não linear mais complexo, exigindo uma melhor extração de informações espaço-temporais do conjunto de dados. Como o modelo de conjunto proposto não possui um modelo de aprendizagem profunda em sua arquitetura, ele não consegue identificar e extrair adequadamente informações espaço-temporais subjacentes ao conjunto de dados, falhando assim em fornecer uma melhor estimativa de irradiância. Os modelos de aprendizagem profunda muitas vezes podem se destacar neste tipo de tarefa, como comprovado nos resultados da Tabela 15. Uma extensa literatura pode ser encontrada sobre melhorias em problemas de previsão de séries temporais quando abordagens complexas e profundas são empregadas (Santos *et al.*, 2023; Marinho *et al.*, 2023; Oliveira *et al.*, 2023).

5 CONCLUSÃO

Este trabalho propôs avaliar o desempenho de quatro modelos de aprendizado de máquina aplicados em duas abordagens que utilizam conjunto dinâmico de dados utilizando velocidade do vento e irradiância solar separadamente como entradas. Inicialmente foram coletados dados de velocidade do vento e irradiância solar da mesma estação meteorológica, determinados horizontes temporais a serem estudados ($t+10$ min, $t+20$ min, $t+30$ min e $t+60$ min) e, em seguida, desenvolvido algoritmo em linguagem Python para os resultados de previsão em várias abordagens, de modo a permitir o comparativo de suas eficiências com base em seus preditores.

Métodos de aprendizado de máquina bem conhecidos em outros trabalhos de pesquisa de previsão de energia relativos a dados de vento e irradiância foram selecionados para comparar sua eficiência com duas outras abordagens que utilizam conjunto dinâmico de dados (*windowing* e *arbitrating*). O código em linguagem de programação *Python* foi desenvolvido para catalogar os parâmetros ótimos de eficiência de cada modelo previamente conhecido, com base em métricas de erro e coeficiente de determinação. Os métodos de conjunto dinâmico (*windowing* e *arbitrating*), baseados nos parâmetros ótimos de cada modelo previamente

ajustados (RF, kNN, SVR e EN), geraram um modelo único com maior eficiência tanto para previsão de dados de irradiância solar, assim como para velocidade do vento.

Para a previsão dos dados de velocidade do vento, constatou-se que o método mais eficiente é o *windowing* para todos os horizontes temporais, quando avaliado pelo critério do menor valor RMSE, e especificamente para o horizonte temporal $t+10$, conforme evidenciado na Figura 3. A maior eficiência foi encontrada no intervalo de 1 a 74 para o parâmetro λ , atingindo desempenho máximo para o valor $\lambda = 19$, conforme visto na Figura 8, o que sugere que a parametrização do *windowing* influencia diretamente no desempenho do método.

Estruturalmente, os dados de irradiância solar são diferentes dos dados de velocidade do vento, pois possuem ciclos na natureza e são fenômenos físicos de características distintas, apresentando diferentes correlações com seus valores históricos e frequência, o que impacta diferentes tendências para o parâmetro *lambda* em cada uma das variáveis.

Para a previsão da irradiância solar, o método mais eficiente também foi o *windowing* e o horizonte temporal $t+10$ min atingiu o menor valor de RMSE. Diferentemente do que foi encontrado para os dados de velocidade do vento, percebeu-se uma maior linearidade na tendência a partir do gráfico de variação dos parâmetros de *windowing* ao analisar seus valores de RMSE. Olhando para o intervalo λ em estudo, foi encontrado o melhor valor de desempenho (usando critérios RMSE) de $\lambda = 1$, como pode ser visto na Figura 10. Ao contrário de todos os outros gráficos, na Figura 12, há um salto repentino entre *lambda* de 1 a 3. Embora a métrica de referência seja o RMSE, para algumas outras métricas o uso de $\lambda = 1$ pode significar informação insuficiente para o modelo, uma vez que terá como variável de entrada apenas um intervalo de tempo anterior (dimensão da janela).

Utilizando dados de velocidade do vento, o ganho de eficiência do modelo mais eficiente (*windowing* para o horizonte de tempo $t+10$ min e $\lambda = 19$, ver Tabela 4), quando comparado com a segunda maior eficiência (SVR), foi de 0,56% quando utilizado o métrica RMSE de valor mais baixo. Uma tendência semelhante pode ser observada para o modelo utilizando dados de irradiância solar. O aumento de eficiência, comparando o modelo mais eficiente (*windowing* para o horizonte temporal $t+10$ min e $\lambda = 1$, ver Tabela 9) com o segundo de maior eficiência (*arbitrating*), foi de cerca de 1,72%, e quando comparado com o terceiro mais eficiente método (SVR), foi de cerca de 1,96%.

Além disso, extensas comparações com modelos espaço-temporais encontrados na literatura mostram que o modelo de conjunto dinâmico para a velocidade do vento muitas vezes fornece desempenho de previsão superior para os horizontes de tempo avaliados, considerando a abordagem proposta como uma ferramenta valiosa para a estimativa da velocidade do vento.

Em relação à previsão de irradiância, a arquitetura de conjunto dinâmico proposta neste estudo não conseguiu superar os modelos baseados em aprendizagem profunda, que apresentaram identificação espaço-temporal superior e, conseqüentemente, melhores valores estimados de GHI. No entanto, a abordagem de *windowing* proposta pode fornecer resultados competitivos e previsão de GHI superior quando comparada a modelos preditivos baseados em física.

Avalia-se que há um grande potencial de melhoria da arquitetura do conjunto dinâmico através da adição de modelos de aprendizado de máquina mais complexos, como o aprendizado profundo e abordagens baseadas em gráficos de conjuntos neurais, como o dos trabalhos de Santos (2023). Isto pode aumentar a capacidade de previsão de *windowing* para o GHI e a estimativa da velocidade do vento, uma vez que seja capaz de se beneficiar das informações de dados espaço-temporais subjacentes ao conjunto de dados. Os modelos foram desenvolvidos para tratar o banco de dados de forma generalizada. Estudos específicos com delimitação de estações e/ou horários do dia poderão ser explorados. O desenvolvimento de um modelo conjunto capaz de fornecer estimativas precisas pode ser empregado no desenvolvimento de aplicações de previsão em tempo real, auxiliando na avaliação da operação de parques eólicos e solares.

REFERÊNCIAS

- ABDELLATIF, Abdallah *et al.* Forecasting Photovoltaic Power Generation with a Stacking Ensemble Model. **Sustainability**, v. 14, n. 17, p. 11083, 5 set. 2022. Disponível em: <https://doi.org/10.3390/su141711083>. Acesso em: 13 dez. 2024.
- ABEEÓLICA - ASSOCIAÇÃO BRASILEIRA DE ENERGIA EÓLICA. [Dados de parques eólicos nacionais]. Disponível em: <http://abeeolica.org.br>. Acesso em: 10 mar. 2024.
- AKDAĞ, Seyit Ahmet; GÜLER, Önder. A novel energy pattern factor method for wind speed distribution parameter estimation. **Energy Conversion and Management**, v. 106, p. 1124-1133, dez. 2015. Disponível em: <https://doi.org/10.1016/j.enconman.2015.10.042>. Acesso em: 13 dez. 2024.
- ALAVI, Omid; MOHAMMADI, Kasra; MOSTAFAEIPOUR, Ali. Evaluating the suitability of wind speed probability distribution models: A case of study of east and southeast parts of Iran. **Energy Conversion and Management**, v. 119, p. 101-108, jul. 2016. Disponível em: <https://doi.org/10.1016/j.enconman.2016.04.039>. Acesso em: 13 dez. 2024.
- AL-NASSAR, W. K. *et al.* Feasibility of offshore wind energy as an alternative source for the state of Kuwait. **Energy**, v. 169, p. 783-796, fev. 2019. Disponível em: <https://doi.org/10.1016/j.energy.2018.11.140>. Acesso em: 13 dez. 2024.
- AMIRINIA, Gholamreza; MAFI, Somayeh; MAZAHERI, Said. Offshore wind resource assessment of Persian Gulf using uncertainty analysis and GIS. **Renewable Energy**, v. 113, p. 915-929, dez. 2017. Disponível em: <https://doi.org/10.1016/j.renene.2017.06.070>. Acesso em: 13 dez. 2024.
- ANDRADE, Carla Freitas de *et al.* An efficiency comparison of numerical methods for determining Weibull parameters for wind energy applications: A new approach applied to the northeast region of Brazil. **Energy Conversion and Management**, v. 86, p. 801-808, out. 2014. Disponível em: <https://doi.org/10.1016/j.enconman.2014.06.046>. Acesso em: 13 dez. 2024.
- APPIOTT, Joseph; DHANJU, Amardeep; CICIN-SAIN, Biliiana. Encouraging renewable energy in the offshore environment. **Ocean & Coastal Management**, v. 90, p. 58-64, mar. 2014. Disponível em: <https://doi.org/10.1016/j.ocecoaman.2013.11.001>. Acesso em: 13 dez. 2024.
- ARIES, Nawel; BOUDIA, Sidi Mohammed; OUNIS, Houdayfa. Deep assessment of wind speed distribution models: A case study of four sites in Algeria. **Energy Conversion and Management**, v. 155, p. 78-90, jan. 2018. Disponível em: <https://doi.org/10.1016/j.enconman.2017.10.082>. Acesso em: 13 dez. 2024.
- BERMANN, Célio. **Crise ambiental e as energias renováveis**. Ciência e Cultura, v. 60, n. 3, p. 20-29, 2008.

BHATI, Deepesh et al. A new 3-parameter extension of generalized Lindley distribution. **arXiv preprint arXiv:1601.01045**, 2016.

BILGILI, Mehmet; YASAR, Abdulkadir; SIMSEK, Erdogan. Offshore wind power development in Europe and its comparison with onshore counterpart. **Renewable and Sustainable Energy Reviews**, v. 15, n. 2, p. 905-915, fev. 2011. Disponível em: <https://doi.org/10.1016/j.rser.2010.11.006>. Acesso em: 13 dez. 2024.

BOURLÈS, Bernard *et al.* THE PIRATA PROGRAM. **Bulletin of the American Meteorological Society**, v. 89, n. 8, p. 1111-1126, ago. 2008. Disponível em: <https://doi.org/10.1175/2008bams2462.1>. Acesso em: 13 dez. 2024.

BREIMAN, Leo. **Machine Learning**, v. 45, n. 1, p. 5-32, 2001. Disponível em: <https://doi.org/10.1023/a:1010933404324>. Acesso em: 13 dez. 2024.

BROWER, Michael. **Wind Resource Assessment: A Practical Guide to Developing a Wind Project**. [S. l.]: Wiley & Sons, Incorporated, John, 2012. 300 p. ISBN 9781118249840.

CALIF, Rudy; SCHMITT, François G.; MEDINA, O. Durán. $-5/3$ Kolmogorov Turbulent Behaviour and Intermittent Sustainable Energies. *In*: CALIF, Rudy; SCHMITT, François G.; MEDINA, O. Durán. **Sustainable Energy - Technological Issues, Applications and Case Studies**. [S. l.]: InTech, 2016. ISBN 9789535128397. Disponível em: <https://doi.org/10.5772/65755>. Acesso em: 14 dez. 2024.

CARNEIRO, Tatiane Carolyne *et al.* Review on Photovoltaic Power and Solar Resource Forecasting: Current Status and Trends. **Journal of Solar Energy Engineering**, v. 144, n. 1, 16 jul. 2021. Disponível em: <https://doi.org/10.1115/1.4051652>. Acesso em: 13 dez. 2024.

CARNEIRO, Tatiane C. *et al.* Ridge regression ensemble of machine learning models applied to solar and wind forecasting in Brazil and Spain. **Applied Energy**, v. 314, p. 118936, maio 2022. Disponível em: <https://doi.org/10.1016/j.apenergy.2022.118936>. Acesso em: 13 dez. 2024.

CELIK, Ali N.; KOLHE, Mohan. Generalized feed-forward based method for wind energy prediction. **Applied Energy**, v. 101, p. 582-588, jan. 2013. Disponível em: <https://doi.org/10.1016/j.apenergy.2012.06.040>. Acesso em: 13 dez. 2024.

CERQUEIRA, Vitor *et al.* Arbitrage of forecasting experts. **Machine Learning**, v. 108, n. 6, p. 913-944, 4 dez. 2018. Disponível em: <https://doi.org/10.1007/s10994-018-05774-y>. Acesso em: 14 dez. 2024.

CERQUEIRA, Vítor; TORGO, Luís; SOARES, Carlos. Arbitrated Ensemble for Solar Radiation Forecasting. *In*: CERQUEIRA, Vítor; TORGO, Luís; SOARES, Carlos. **Advances in Computational Intelligence**. Cham: Springer International Publishing, 2017. p. 720-732. ISBN 9783319591520. Disponível em: https://doi.org/10.1007/978-3-319-59153-7_62. Acesso em: 14 dez. 2024.

CHE, Jinxing *et al.* Ultra-short-term probabilistic wind power forecasting with spatial-temporal multi-scale features and K-FSDW based weight. **Applied Energy**, v. 331,

p. 120479, fev. 2023. Disponível em: <https://doi.org/10.1016/j.apenergy.2022.120479>. Acesso em: 13 dez. 2024.

CHEN, Tianqi; GUESTRIN, Carlos. XGBoost. *In: KDD '16: THE 22ND ACM SIGKDD INTERNATIONAL CONFERENCE ON KNOWLEDGE DISCOVERY AND DATA MINING*, San Francisco California USA. **KDD '16: The 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining**. New York, NY, USA: ACM, 2016. ISBN 9781450342322. Disponível em: <https://doi.org/10.1145/2939672.2939785>. Acesso em: 13 dez. 2024.

DAHMANI, Abdennasser; AMMI, Yamina; HANINI, Salah. A Novel Non-Linear Model Based on Bootstrapped Aggregated Support Vector Machine for the Prediction of Hourly Global Solar Radiation. **Smart Grids and Sustainable Energy**, v. 9, n. 1, 30 nov. 2023b. Disponível em: <https://doi.org/10.1007/s40866-023-00179-w>. Acesso em: 13 dez. 2024.

DE ASSIS TAVARES, Luiz Filipe *et al.* Assessment of the offshore wind technical potential for the Brazilian Southeast and South regions. **Energy**, v. 196, p. 117097, abr. 2020. Disponível em: <https://doi.org/10.1016/j.energy.2020.117097>. Acesso em: 14 dez. 2024.

DE MYTTENAERE, Arnaud *et al.* Mean Absolute Percentage Error for regression models. **Neurocomputing**, v. 192, p. 38-48, jun. 2016b. Disponível em: <https://doi.org/10.1016/j.neucom.2015.12.114>. Acesso em: 13 dez. 2024.

DEMIR, Vahdettin; CITAKOGLU, Hatice. Forecasting of solar radiation using different machine learning approaches. **Neural Computing and Applications**, 23 set. 2022. Disponível em: <https://doi.org/10.1007/s00521-022-07841-x>. Acesso em: 13 dez. 2024.

DIGHE, Deepti; PATIL, Sneha; KOKATE, Shrikant. Detection of Credit Card Fraud Transactions Using Machine Learning Algorithms and Neural Networks: A Comparative Study. *In: 2018 FOURTH INTERNATIONAL CONFERENCE ON COMPUTING COMMUNICATION CONTROL AND AUTOMATION (ICCUBEA)*, 2018, Pune, India. **2018 Fourth International Conference on Computing Communication Control and Automation (ICCUBEA)**. [S. l.]: IEEE, 2018. ISBN 9781538652572. Disponível em: <https://doi.org/10.1109/iccubea.2018.8697799>. Acesso em: 13 dez. 2024.

DING, Yang *et al.* Probabilistic method for wind speed prediction and statistics distribution inference based on SHM data-driven. **Probabilistic Engineering Mechanics**, p. 103475, jun. 2023. Disponível em: <https://doi.org/10.1016/j.probenmech.2023.103475>. Acesso em: 13 dez. 2024.

DOWELL, Jethro; WEISS, Stephan; INFIELD, David. Spatio-temporal prediction of wind speed and direction by continuous directional regime. *In: 2014 INTERNATIONAL CONFERENCE ON PROBABILISTIC METHODS APPLIED TO POWER SYSTEMS (PMAPS)*, 2014, Durham, United Kingdom. **2014 International Conference on Probabilistic Methods Applied to Power Systems (PMAPS)**. [S. l.]: IEEE, 2014. ISBN 9781479935611. Disponível em: <https://doi.org/10.1109/pmaps.2014.6960596>. Acesso em: 13 dez. 2024.

DU, Liang *et al.* Bayesian optimization based dynamic ensemble for time series forecasting. **Information Sciences**, v. 591, p. 155-175, abr. 2022. Disponível em: <https://doi.org/10.1016/j.ins.2022.01.010>. Acesso em: 13 dez. 2024.

EPE - Empresa de Pesquisa Energética. **Caderno de Tecnologias de Geração**. Brasília, DF: Ministério de Minas e Energia, Secretaria de Planejamento e Desenvolvimento Energético, 2023. Disponível em: <https://www.epe.gov.br/sites-pt/publicacoes-dados-abertos/publicacoes/PublicacoesArquivos/publicacao745/Caderno%20de%20Tecnologias%20de%20Gera%C3%A7%C3%A3o%202023.pdf>. Acesso em 26 mar. 2024.

FREITAS, B. B.; PALACIO, G. B. A.; SILVA, F. J. R. da; SILVA, F. E. M. da; FREITAS, D. M. de; QUEIROZ, D. M. B. de; RAMALHO, D. C. L.; LOAYZA, D. M. J.; LINS, D. R.; TELES, C. F. Scenarios for Ceará's electricity generation matrix in 2050: en. **Ciência e Natura**, [S. l.], v. 42, p. e63, 2022. DOI: 10.5902/2179460X42805. Disponível em: <https://periodicos.ufsm.br/cienciaenatura/article/view/42805>. Acesso em: 13 dec. 2024.

FUKUNAGA, K.; NARENDRA, P. M. A Branch and Bound Algorithm for Computing k-Nearest Neighbors. **IEEE Transactions on Computers**, v. C-24, n. 7, p. 750-753, jul. 1975. Disponível em: <https://doi.org/10.1109/t-c.1975.224297>. Acesso em: 13 dez. 2024.

GHOJOGH, Benyamin; CROWLEY, Mark. The theory behind overfitting, cross validation, regularization, bagging, and boosting: tutorial. **arXiv preprint arXiv:1905.12787**, 2019.

GOOGLE. Google Earth Website. [**Mapa de estudo**] Disponível em: <http://earth.google.com/>. Acesso em: 12 jul. 2023.

GUEDES, Kevin S. *et al.* Performance analysis of metaheuristic optimization algorithms in estimating the parameters of several wind speed distributions. **Applied Energy**, v. 268, p. 114952, jun. 2020. Disponível em: <https://doi.org/10.1016/j.apenergy.2020.114952>. Acesso em: 13 dez. 2024.

GUGLIANI, G. K. *et al.* New methods to assess wind resources in terms of wind speed, load, power and direction. **Renewable Energy**, v. 129, p. 168-182, dez. 2018. Disponível em: <https://doi.org/10.1016/j.renene.2018.05.088>. Acesso em: 13 dez. 2024.

GUPTA, Priya; SINGH, Rhythm. Combining a deep learning model with multivariate empirical mode decomposition for hourly global horizontal irradiance forecasting. **Renewable Energy**, fev. 2023. Disponível em: <https://doi.org/10.1016/j.renene.2023.02.052>. Acesso em: 13 dez. 2024.

GWEC COUNCIL, Global Wind Energy. Global offshore wind report 2020. **GWEC: Brussels, Belgium**, v. 19, p. 10-12, 2020.

HE, Junyi *et al.* Spatiotemporal analysis of offshore wind field characteristics and energy potential in Hong Kong. **Energy**, v. 201, p. 117622, jun. 2020. Disponível em: <https://doi.org/10.1016/j.energy.2020.117622>. Acesso em: 13 dez. 2024.

HOERL, Arthur E.; KENNARD, Robert W. Ridge Regression: Biased Estimation for Nonorthogonal Problems. **Technometrics**, v. 12, n. 1, p. 55-67, fev. 1970. Disponível em: <https://doi.org/10.1080/00401706.1970.10488634>. Acesso em: 13 dez. 2024.

HYNDMAN, Rob J.; KOEHLER, Anne B. Another look at measures of forecast accuracy. **International Journal of Forecasting**, v. 22, n. 4, p. 679-688, out. 2006. Disponível em: <https://doi.org/10.1016/j.ijforecast.2006.03.001>. Acesso em: 13 dez. 2024.

IEC – TURBINES—PART, Wind. 12-1: Power Performance Measurements of Electricity Producing Wind Turbines. **British Standard, IEC**, p. 61400-12, 2005.

INEICHEN, Pierre; PEREZ, Richard. A new air mass independent formulation for the Linke turbidity coefficient. **Solar Energy**, v. 73, n. 3, p. 151-157, set. 2002. Disponível em: [https://doi.org/10.1016/s0038-092x\(02\)00045-2](https://doi.org/10.1016/s0038-092x(02)00045-2). Acesso em: 13 dez. 2024.

INPE – INSTITUTO DE PESQUISAS ESPACIAIS. [SONDA - Sistema De Organização Nacional de Dados Ambientais]. 2012. Disponível em: <http://sonda.ccst.inpe.br>. Acesso em: 14 mar. 2024.

IPCC, CHANGE Climate et al. Mitigation of climate change. **Contribution of working group III to the fifth assessment report of the intergovernmental panel on climate change**, v. 1454, p. 147, 2014.

IRENA - International renewable energy agency. **Renewable Energy Statistics**. Abu Dhabi, v. 2020. Disponível em: <https://irena.org/>. Acesso em: 14 mar. 2024.

JAMES, Gareth *et al.* **An Introduction to Statistical Learning**. New York, NY: Springer New York, 2013. *E-book*. ISBN 9781461471370. Disponível em: <https://doi.org/10.1007/978-1-4614-7138-7>. Acesso em: 13 dez. 2024.

Jl, Lipeng *et al.* Short-Term Canyon Wind Speed Prediction Based on CNN—GRU Transfer Learning. **Atmosphere**, v. 13, n. 5, p. 813, 16 maio 2022. Disponível em: <https://doi.org/10.3390/atmos13050813>. Acesso em: 14 dez. 2024.

JIANG, Haiyan *et al.* Comparison of numerical methods and metaheuristic optimization algorithms for estimating parameters for wind energy potential assessment in low wind regions. **Renewable and Sustainable Energy Reviews**, v. 69, p. 1199-1217, mar. 2017. Disponível em: <https://doi.org/10.1016/j.rser.2016.11.241>. Acesso em: 14 dez. 2024.

JUNG, Christopher *et al.* Introducing a system of wind speed distributions for modeling properties of wind speed regimes around the world. **Energy Conversion and Management**, v. 144, p. 181-192, jul. 2017. Disponível em: <https://doi.org/10.1016/j.enconman.2017.04.044>. Acesso em: 14 dez. 2024.

KALLIO-MYERS, Viivi *et al.* Global horizontal irradiance forecast for Finland based on geostationary weather satellite data. **Solar Energy**, v. 198, p. 68-80, mar. 2020. Disponível em: <https://doi.org/10.1016/j.solener.2020.01.008>. Acesso em: 14 dez. 2024.

KANTAR, Yeliz Mert *et al.* Wind speed analysis using the Extended Generalized Lindley Distribution. **Renewable Energy**, v. 118, p. 1024-1030, abr. 2018. Disponível em: <https://doi.org/10.1016/j.renene.2017.09.053>. Acesso em: 14 dez. 2024.

KASTEN, Fritz; CZEPLAK, Gerhard. Solar and terrestrial radiation dependent on the amount and type of cloud. **Solar Energy**, v. 24, n. 2, p. 177-189, 1980. Disponível em: [https://doi.org/10.1016/0038-092x\(80\)90391-6](https://doi.org/10.1016/0038-092x(80)90391-6). Acesso em: 14 dez. 2024.

KISS, Péter; JÁNOSI, Imre M. Comprehensive empirical analysis of ERA-40 surface wind speed distribution over Europe. **Energy Conversion and Management**, v. 49, n. 8, p. 2142-2151, ago. 2008. Disponível em: <https://doi.org/10.1016/j.enconman.2008.02.003>. Acesso em: 14 dez. 2024.

KUHN, Max; JOHNSON, Kjell. **Applied Predictive Modeling**. New York, NY: Springer New York, 2013. *E-book*. ISBN 9781461468486. Disponível em: <https://doi.org/10.1007/978-1-4614-6849-3>. Acesso em: 14 dez. 2024.

KVÅLSETH, Tarald O. Cautionary Note about R². **The American Statistician**, v. 39, n. 4, p. 279-285, nov. 1985. Disponível em: <https://doi.org/10.1080/00031305.1985.10479448>. Acesso em: 14 dez. 2024.

KWASNEY, Bruno Cesar et al. Alternativa da energia eólica e solar: Estudo da Região Sul do Brasil entre 2006 a 2014. **Alternativa**, v. 38, n. 06, 2017.

LADENBURG, Jacob. Visual impact assessment of offshore wind farms and prior experience. **Applied Energy**, v. 86, n. 3, p. 380-387, mar. 2009. Disponível em: <https://doi.org/10.1016/j.apenergy.2008.05.005>. Acesso em: 14 dez. 2024.

LAHOUAR, A.; BEN HADJ SLAMA, J. Hour-ahead wind power forecast based on random forests. **Renewable Energy**, v. 109, p. 529-541, ago. 2017. Disponível em: <https://doi.org/10.1016/j.renene.2017.03.064>. Acesso em: 14 dez. 2024.

LAKKU, Naresh K. G.; BEHERA, Manasa R. Skill and Intercomparison of Global Climate Models in Simulating Wind Speed, and Future Changes in Wind Speed over South Asian Domain. **Atmosphere**, v. 13, n. 6, p. 864, 25 maio 2022. Disponível em: <https://doi.org/10.3390/atmos13060864>. Acesso em: 14 dez. 2024.

LANDBERG, L. *et al.* (2003). Wind resource estimation - an overview. **Wind Energy: An International Journal for Progress and Applications in Wind Power Conversion Technology**, 6(3), 261-271.

LAUBER, Volkmar. Political Economy of Renewable Energy. *In*: LAUBER, Volkmar. **International Encyclopedia of the Social & Behavioral Sciences**. [S. l.]: Elsevier, 2015. p. 367-373. ISBN 9780080970875. Disponível em: <https://doi.org/10.1016/b978-0-08-097086-8.91084-5>. Acesso em: 14 dez. 2024.

LI, Jiale; YU, Xiong (Bill). Onshore and offshore wind energy potential assessment near Lake Erie shoreline: A spatial and temporal analysis. **Energy**, v. 147, p. 1092-1107, mar. 2018. Disponível em: <https://doi.org/10.1016/j.energy.2018.01.118>. Acesso em: 14 dez. 2024.

LI, Yi *et al.* Comparative study of onshore and offshore wind characteristics and wind energy potentials: A case study for southeast coastal region of China. **Sustainable Energy Technologies and Assessments**, v. 39, p. 100711, jun. 2020. Disponível em: <https://doi.org/10.1016/j.seta.2020.100711>. Acesso em: 14 dez. 2024.

LINDLEY, D. V. Fiducial Distributions and Bayes' Theorem. **Journal of the Royal Statistical Society: Series B (Methodological)**, v. 20, n. 1, p. 102-107, jan. 1958. Disponível em: <https://doi.org/10.1111/j.2517-6161.1958.tb00278.x>. Acesso em: 14 dez. 2024.

LIU, Jingxuan *et al.* A Transformer-based multimodal-learning framework using sky images for ultra-short-term solar irradiance forecasting. **Applied Energy**, v. 342, p. 121160, jul. 2023. Disponível em: <https://doi.org/10.1016/j.apenergy.2023.121160>. Acesso em: 14 dez. 2024.

LIU, Zhuoyi; HARA, Ryoichi; KITA, Hiroyuki. Hybrid forecasting system based on data area division and deep learning neural network for short-term wind speed forecasting. **Energy Conversion and Management**, v. 238, p. 114136, jun. 2021. Disponível em: <https://doi.org/10.1016/j.enconman.2021.114136>. Acesso em: 14 dez. 2024.

MAKARIEVA, A. M. *et al.* Where do winds come from? A new theory on how water vapor condensation influences atmospheric pressure and dynamics. **Atmospheric Chemistry and Physics Discussions**, v. 10, n. 10, p. 24015-24052, 15 out. 2010. Disponível em: <https://doi.org/10.5194/acpd-10-24015-2010>. Acesso em: 14 dez. 2024.

MANWELL, James F.; ROGERS, Anthony L.; MCGOWAN, Jon G. **Wind Energy Explained: Theory, Design and Application**. [S. l.]: Wiley & Sons, Incorporated, John, 2010. 704 p. ISBN 9780470686287.

MARINHO, Felipe *et al.* Short-Term Solar Irradiance Forecasting Using CNN-1D, LSTM and CNN-LSTM Deep Neural Networks: A Case Study with the Folsom (USA) Dataset. **Journal of Solar Energy Engineering**, p. 1-11, 31 out. 2022. Disponível em: <https://doi.org/10.1115/1.4056122>. Acesso em: 14 dez. 2024.

MARQUEZ, Ricardo; COIMBRA, Carlos F. M. Proposed Metric for Evaluation of Solar Forecasting Models. **Journal of Solar Energy Engineering**, v. 135, n. 1, 23 out. 2012. Disponível em: <https://doi.org/10.1115/1.4007496>. Acesso em: 14 dez. 2024.

MEENAL, R. *et al.* Weather Forecasting for Renewable Energy System: A Review. **Archives of Computational Methods in Engineering**, 26 jan. 2022. Disponível em: <https://doi.org/10.1007/s11831-021-09695-3>. Acesso em: 14 dez. 2024.

MESA-JIMÉNEZ, J. J. *et al.* Long-term wind and solar energy generation forecasts, and optimisation of Power Purchase Agreements. **Energy Reports**, v. 9, p. 292-302, dez. 2023. Disponível em: <https://doi.org/10.1016/j.egy.2022.11.175>. Acesso em: 14 dez. 2024.

MOHAMMADI, Kasra; ALAVI, Omid; MCGOWAN, Jon G. Use of Birnbaum-Saunders distribution for estimating wind speed and wind power probability distributions: A review. **Energy Conversion and Management**, v. 143, p. 109-122, 2017.

MOKGONYANE, Tumisho Billson *et al.* Automatic Speaker Recognition System based on Machine Learning Algorithms. *In*: 2019 SOUTHERN AFRICAN UNIVERSITIES POWER ENGINEERING CONFERENCE/ROBOTICS AND MECHATRONICS/PATTERN RECOGNITION ASSOCIATION OF SOUTH AFRICA (SAUPEC/ROBMECH/PRASA), 2019, Bloemfontein, South Africa. **2019 Southern African Universities Power Engineering Conference/Robotics and Mechatronics/Pattern Recognition Association of South Africa**

(SAUPEC/RobMech/PRASA). [S. l.]: IEEE, 2019. ISBN 9781728103693. Disponível em: <https://doi.org/10.1109/robomech.2019.8704837>. Acesso em: 14 dez. 2024.

MORGAN, Eugene C. *et al.* Probability distributions for offshore wind speeds. **Energy Conversion and Management**, v. 52, n. 1, p. 15-26, jan. 2011. Disponível em: <https://doi.org/10.1016/j.enconman.2010.06.015>. Acesso em: 14 dez. 2024.

NADARAJAH, Saralees; BAKOUCH, Hassan S.; TAHMASBI, Rasool. A generalized Lindley distribution. **Sankhya B**, v. 73, n. 2, p. 331-359, nov. 2011. Disponível em: <https://doi.org/10.1007/s13571-011-0025-9>. Acesso em: 14 dez. 2024.

NAKAGAMI, MINORU. The m-Distribution—A General Formula of Intensity Distribution of Rapid Fading. *In*: NAKAGAMI, MINORU. **Statistical Methods in Radio Wave Propagation**. [S. l.]: Elsevier, 1960. p. 3-36. ISBN 9780080093062. Disponível em: <https://doi.org/10.1016/b978-0-08-009306-2.50005-4>. Acesso em: 14 dez. 2024.

NESHAT, Mehdi *et al.* A deep learning-based evolutionary model for short-term wind speed forecasting: A case study of the Lillgrund offshore wind farm. **Energy Conversion and Management**, v. 236, p. 114002, maio 2021. Disponível em: <https://doi.org/10.1016/j.enconman.2021.114002>. Acesso em: 14 dez. 2024.

NIKODINOSKA, Dragana; KÄSO, Mathias; MÜSGENS, Felix. Solar and wind power generation forecasts using elastic net in time-varying forecast combinations. **Applied Energy**, v. 306, p. 117983, jan. 2022. Disponível em: <https://doi.org/10.1016/j.apenergy.2021.117983>. Acesso em: 14 dez. 2024.

OLIVEIRA SANTOS, Victor *et al.* A New Graph-Based Deep Learning Model to Predict Flooding with Validation on a Case Study on the Humber River. **Water**, v. 15, n. 10, p. 1827, 10 maio 2023. Disponível em: <https://doi.org/10.3390/w15101827>. Acesso em: 14 dez. 2024.

OLIVEIRA SANTOS, Victor *et al.* Graph-Based Deep Learning Model for Forecasting Chloride Concentration in Urban Streams to Protect Salt-Vulnerable Areas. **Environments**, v. 10, n. 9, p. 157, 12 set. 2023. Disponível em: <https://doi.org/10.3390/environments10090157>. Acesso em: 14 dez. 2024.

OLIVEIRA SANTOS, Victor *et al.* Spatiotemporal Air Pollution Forecasting in Houston-TX: A Case Study for Ozone Using Deep Graph Neural Networks. **Atmosphere**, v. 14, n. 2, p. 308, 3 fev. 2023b. Disponível em: <https://doi.org/10.3390/atmos14020308>. Acesso em: 14 dez. 2024.

OLIVEIRA SANTOS, Victor *et al.* Spatiotemporal analysis of bidimensional wind speed forecasting: Development and thorough assessment of LSTM and ensemble graph neural networks on the Dutch database. **Energy**, p. 127852, maio 2023. Disponível em: <https://doi.org/10.1016/j.energy.2023.127852>. Acesso em: 14 dez. 2024.

OLIVEIRA, Jéssica Rayane Hernandez. Energia Solar Fotovoltaica. **Revista Ibero-Americana de Humanidades, Ciências e Educação**, v. 9, n. 3, p. 1945-1954, 2023.

OSMAN, Ahmed I. *et al.* Cost, environmental impact, and resilience of renewable energy under a changing climate: a review. **Environmental Chemistry Letters**, 28 out. 2022. Disponível em: <https://doi.org/10.1007/s10311-022-01532-8>. Acesso em: 14 dez. 2024.

OTTAVIANELI, Emilce; GONZÁLEZ, Facundo David Francisco; CADENA, Carlos Alberto. Hábitat y pobreza energética en zonas rurales aisladas en el noroeste Argentino. **Ciencia Latina Revista Científica Multidisciplinar**, v. 5, n. 5, p. 7997-8017, out. 2021b. Disponível em: https://doi.org/10.37811/cl_rcm.v5i5.886. Acesso em: 14 dez. 2024.

OUARDA, T. B. M. J.; CHARRON, C.; CHEBANA, F. Review of criteria for the selection of probability distributions for wind speed data and introduction of the moment and L-moment ratio diagram methods, with a case study. **Energy Conversion and Management**, v. 124, p. 247-265, set. 2016. Disponível em: <https://doi.org/10.1016/j.enconman.2016.07.012>. Acesso em: 14 dez. 2024.

PARSONS, David J. **Mobile Radio Propagation Channel**. [S. l.]: Wiley & Sons, Incorporated, John, 2000. ISBN 9781280554964.

PEDRO, Hugo T. C. *et al.* Assessment of machine learning techniques for deterministic and probabilistic intra-hour solar forecasts. **Renewable Energy**, v. 123, p. 191-203, ago. 2018. Disponível em: <https://doi.org/10.1016/j.renene.2018.02.006>. Acesso em: 14 dez. 2024.

PEEL, M. C.; FINLAYSON, B. L.; MCMAHON, T. A. Updated world map of the Köppen-Geiger climate classification. **Hydrology and Earth System Sciences Discussions**, v. 4, n. 2, p. 439-473, 1 mar. 2007. Disponível em: <https://doi.org/10.5194/hessd-4-439-2007>. Acesso em: 14 dez. 2024.

PENG, Zhiyun *et al.* A novel deep learning ensemble model with data denoising for short-term wind speed forecasting. **Energy Conversion and Management**, v. 207, p. 112524, mar. 2020. Disponível em: <https://doi.org/10.1016/j.enconman.2020.112524>. Acesso em: 14 dez. 2024.

R, Krishnamoorthy *et al.* An Assessment of Onshore and Offshore Wind Energy Potential in India Using Moth Flame Optimization. **Energies**, v. 13, n. 12, p. 3063, 13 jun. 2020. Disponível em: <https://doi.org/10.3390/en13123063>. Acesso em: 14 dez. 2024.

ROCHA, P. A. Costa *et al.* Estimation of daily, weekly and monthly global solar radiation using ANNs and a long data set: a case study of Fortaleza, in Brazilian Northeast region. **International Journal of Energy and Environmental Engineering**, v. 10, n. 3, p. 319-334, 11 jul. 2019. Disponível em: <https://doi.org/10.1007/s40095-019-0313-0>. Acesso em: 14 dez. 2024.

ROCHA, Paulo A. C.; SANTOS, Victor Oliveira. Global horizontal and direct normal solar irradiance modeling by the machine learning methods XGBoost and deep neural networks with CNN-LSTM layers: a case study using the GOES-16 satellite imagery. **International Journal of Energy and Environmental Engineering**, 19 abr. 2022. Disponível em: <https://doi.org/10.1007/s40095-022-00493-6>. Acesso em: 14 dez. 2024.

RSTUDIO. [Código computacional] 2020. Disponível em: <https://rstudio.com/about/>. Acesso em: 27 de mar. 2024.

SAMPAIO, Keila Regina Alves; BATISTA, Valmir. O atual cenário da produção de energia eólica no Brasil: Uma revisão de literatura. **Research, Society and Development**, v. 10, n. 1, p. e57710112107, 31 jan. 2021. Disponível em: <https://doi.org/10.33448/rsd-v10i1.12107>. Acesso em: 14 dez. 2024.

SANCHEZ-IBORRA, Ramon; CANO, Maria-Dolores; GARCIA-HARO, Joan. Performance evaluation of QoE in VoIP traffic under fading channels. *In: 2013 WORLD CONGRESS ON COMPUTER AND INFORMATION TECHNOLOGY (WCCIT)*, 2013, Sousse, Tunisia. **2013 World Congress on Computer and Information Technology (WCCIT)**. [S. l.]: IEEE, 2013. ISBN 9781479904624. Disponível em: <https://doi.org/10.1109/wccit.2013.6618721>. Acesso em: 14 dez. 2024.

SANTOS, Lindemberg Ferreira dos. **Parametrização e aplicação de métodos heurísticos no ajuste da curva de distribuição de Weibull em energia eólica**. 2018. 68 f. Dissertação (Mestrado em Engenharia Mecânica) - Centro de Tecnologia, Universidade Federal do Ceará, Fortaleza, 2018.

SCHWEGMANN, Sandra *et al.* Enabling Virtual Met Masts for wind energy applications through machine learning-methods. **Energy and AI**, p. 100209, out. 2022. Disponível em: <https://doi.org/10.1016/j.egyai.2022.100209>. Acesso em: 14 dez. 2024.

SERVAIN, Jacques *et al.* A Pilot Research Moored Array in the Tropical Atlantic (PIRATA). **Bulletin of the American Meteorological Society**, v. 79, n. 10, p. 2019-2031, out. 1998. Disponível em: [https://doi.org/10.1175/1520-0477\(1998\)079%3C2019:aprmia%3E2.0.co;2](https://doi.org/10.1175/1520-0477(1998)079%3C2019:aprmia%3E2.0.co;2). Acesso em: 14 dez. 2024.

SHIGA, Momotoshi *et al.* Characteristic phenotype of immortalized periodontal cells isolated from a Marfan syndrome type I patient. **Cell and Tissue Research**, v. 331, n. 2, p. 461-472, 30 nov. 2007. Disponível em: <https://doi.org/10.1007/s00441-007-0528-x>. Acesso em: 14 dez. 2024.

SHIKHOVTSEV, Artem Yu *et al.* Application of Neural Networks to Estimation and Prediction of Seeing at the Large Solar Telescope Site. **Publications of the Astronomical Society of the Pacific**, v. 135, n. 1043, p. 014503, 1 jan. 2023. Disponível em: <https://doi.org/10.1088/1538-3873/acb384>. Acesso em: 14 dez. 2024.

SIDHU, Manjit Singh. **Advanced Technology-Assisted Problem Solving in Engineering Education: Emerging Research and Opportunities**. [S. l.]: IGI Global, 2019. ISBN 9781799804666.

SMOLA, A. **Regression Estimation with Support Vector Learning Machines**. Master's Thesis, Technische Universit at Munchen, Munich, Germany, 1996.

STEVENS, M. J. M. et SMULDERS, P. T. **The estimation of the parameters of the Weibull wind speed distribution for wind energy utilization purposes**. *Wind engineering*, 1979, p. 132-145.

SU, Xinyue *et al.* Prediction of Short-Time Cloud Motion Using a Deep-Learning Model. **Atmosphere**, v. 11, n. 11, p. 1151, 26 out. 2020. Disponível em: <https://doi.org/10.3390/atmos11111151>. Acesso em: 14 dez. 2024.

SUN, Pengxiang; WANG, Juan; YAN, Zhenguo. Ultra-short-term wind speed prediction based on TCN-MCM-EKF. **Energy Reports**, v. 11, p. 2127-2140, jun. 2024. Disponível em: <https://doi.org/10.1016/j.egy.2024.01.058>. Acesso em: 14 dez. 2024.

TABRIZI, Sepideh Emami *et al.* Hourly road pavement surface temperature forecasting using deep learning models. **Journal of Hydrology**, v. 603, p. 126877, dez. 2021. Disponível em: <https://doi.org/10.1016/j.jhydrol.2021.126877>. Acesso em: 14 dez. 2024.

TEAM, R. Core. R A language and environment for statistical computing, R Foundation for Statistical. **Computing**, 2020. Disponível em: <https://www.r-project.org/>. Acesso em: 14 dez. 2024.

TIBSHIRANI, Robert. Regression Shrinkage and Selection Via the Lasso. **Journal of the Royal Statistical Society: Series B (Methodological)**, v. 58, n. 1, p. 267-288, jan. 1996. Disponível em: <https://doi.org/10.1111/j.2517-6161.1996.tb02080.x>. Acesso em: 14 dez. 2024.

USTA, Ilhan *et al.* A new estimation approach based on moments for estimating Weibull parameters in wind power applications. **Energy Conversion and Management**, v. 164, p. 570-578, maio 2018. Disponível em: <https://doi.org/10.1016/j.enconman.2018.03.033>. Acesso em: 14 dez. 2024.

VAPNIK, Vladimir Naumovich. **Statistical learning theory**. New York: Wiley, 1998. 736 p. ISBN 0471030031.

VASWANI, A. Attention is all you need. **Advances in Neural Information Processing Systems**, 2017.

VENKATA MAHESH, P.; MEYYAPPAN, S.; ALLA, RamaKoteswaraRao. Support Vector Regression Machine Learning based Maximum Power Point Tracking for Solar Photovoltaic systems. **International journal of electrical and computer engineering systems**, v. 14, n. 1, p. 100-108, 26 jan. 2023. Disponível em: <https://doi.org/10.32985/ijeces.14.1.11>. Acesso em: 14 dez. 2024.

VICHI, Flavio Maron; MANSOR, Maria Teresa Castilho. Energia, meio ambiente e economia: o Brasil no contexto mundial. **Química Nova**, v. 32, n. 3, p. 757-767, 2009. Disponível em: <https://doi.org/10.1590/s0100-40422009000300019>. Acesso em: 14 dez. 2024.

WAIS, Piotr. A review of Weibull functions in wind sector. **Renewable and Sustainable Energy Reviews**, v. 70, p. 1099-1107, abr. 2017. Disponível em: <https://doi.org/10.1016/j.rser.2016.12.014>. Acesso em: 14 dez. 2024.

WILLMOTT, CJ; MATSUURA, K. Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance. **Climate Research**, v. 30, p. 79-82, 2005. Disponível em: <https://doi.org/10.3354/cr030079>. Acesso em: 14 dez. 2024.

WU, Hao; LEVINSON, David. The ensemble approach to forecasting: A review and synthesis. **Transportation Research Part C: Emerging Technologies**, v. 132, p. 103357,

nov. 2021. Disponível em: <https://doi.org/10.1016/j.trc.2021.103357>. Acesso em: 14 dez. 2024.

WU, Qiang *et al.* Promoting wind energy for sustainable development by precise wind speed prediction based on graph neural networks. **Renewable Energy**, set. 2022. Disponível em: <https://doi.org/10.1016/j.renene.2022.09.036>. Acesso em: 14 dez. 2024.

YANG, Liwei *et al.* Intra-day global horizontal irradiance forecast using FY-4A clear sky index. **Sustainable Energy Technologies and Assessments**, v. 50, p. 101816, mar. 2022. Disponível em: <https://doi.org/10.1016/j.seta.2021.101816>. Acesso em: 14 dez. 2024.

YU, Jianxing *et al.* Assessment of Offshore Wind Characteristics and Wind Energy Potential in Bohai Bay, China. **Energies**, v. 12, n. 15, p. 2879, 26 jul. 2019. Disponível em: <https://doi.org/10.3390/en12152879>. Acesso em: 14 dez. 2024.

YUVAL, Janni; O’GORMAN, Paul A. Neural-Network Parameterization of Subgrid Momentum Transport in the Atmosphere. **Journal of Advances in Modeling Earth Systems**, v. 15, n. 4, abr. 2023. Disponível em: <https://doi.org/10.1029/2023ms003606>. Acesso em: 14 dez. 2024.

ZHANG, Yue *et al.* The Discharge Forecasting of Multiple Monitoring Station for Humber River by Hybrid LSTM Models. **Water**, v. 14, n. 11, p. 1794, 2 jun. 2022. Disponível em: <https://doi.org/10.3390/w14111794>. Acesso em: 14 dez. 2024.

ZHU, Qiaomu *et al.* Learning Temporal and Spatial Correlations Jointly: A Unified Framework for Wind Speed Prediction. **IEEE Transactions on Sustainable Energy**, v. 11, n. 1, p. 509-523, jan. 2020. Disponível em: <https://doi.org/10.1109/tste.2019.2897136>. Acesso em: 14 dez. 2024.

ZOU, Hui; HASTIE, Trevor. Regularization and variable selection via the elastic net. **Journal of the Royal Statistical Society: Series B (Statistical Methodology)**, v. 67, n. 2, p. 301-320, abr. 2005. Disponível em: <https://doi.org/10.1111/j.1467-9868.2005.00503.x>. Acesso em: 14 dez. 2024.