

CESAR SCHOOL

RAFAEL ABREU DE CARVALHO

**ESTRATÉGIAS DE INTERPOLAÇÃO ESPACIAL PARA ANÁLISE DE
MEDIDAS CLIMÁTICAS NO CONTEXTO DE ENERGIAS
RENOVÁVEIS**

**RECIFE
2025**

RAFAEL ABREU DE CARVALHO

ESTRATÉGIAS DE INTERPOLAÇÃO ESPACIAL PARA ANÁLISE DE MEDIDAS CLIMÁTICAS NO CONTEXTO DE ENERGIAS RENOVÁVEIS

Trabalho de conclusão de curso apresentado ao Curso de Graduação do Centro de Estudos e Sistemas Educacionais do Recife – CESAR SCHOOL, como requisito para a obtenção do título de Bacharel em Ciência da Computação.

Orientador: Diego de Freitas Bezerra

RECIFE
2025



RAFAEL ABREU DE CARVALHO

**ESTRATÉGIAS DE INTERPOLAÇÃO ESPACIAL PARA ANÁLISE DE MEDIDAS
CLIMÁTICAS NO CONTEXTO DE ENERGIAS RENOVÁVEIS**

Trabalho aprovado em Recife: **19/06/2025**

Diego de Freitas Bezerra
(CESAR SCHOOL)
Orientador

Ricardo Cesar de Andrade
(CESAR SCHOOL)
Avaliador Interno

Miguel Luiz Pessoa da Cruz Silva
(CESAR)
Avaliador Externo

RECIFE
2025

Agradecimentos

Toda a minha gratidão ao meu orientador, Professor Diego Bezerra. Seu suporte foi crucial em toda a jornada, desde a ajuda na definição do tema até os conselhos finais, que foram essenciais para a realização deste trabalho.

Resumo

A baixa disponibilidade e o custo para expansão da rede de estações meteorológicas para regiões remotas do Brasil impactam a avaliação da viabilidade de projetos de produção de energia renovável. Este trabalho investiga estratégias de interpolação espacial, incluindo IDW, Krigagem e Random Forest, aplicadas à análise de variáveis climáticas relevantes para projetos de energia renovável, com foco especial na produção de energia eólica e solar. Nesse sentido, foram avaliadas técnicas de interpolação espacial sobre dados de 12 estações do Instituto Nacional de Meteorologia, abrangendo variáveis como velocidade do vento, radiação solar e temperaturas máxima e mínima no Estado de Pernambuco. Essas técnicas foram avaliadas através da validação cruzada, com validação baseada na existência de usinas reais, onde foi possível constatar a superioridade do método IDW, sobretudo na estimativa da velocidade do vento. A pesquisa aponta a interpolação como ferramenta técnica e estratégica para o planejamento energético sustentável, contribuindo com dados precisos em regiões com baixa densidade de medições.

Palavras-chave: Modelos híbridos; Aprendizado Supervisionado; Análise geoespacial; Viabilidade energética.

Abstract

The limited availability and high cost of expanding the meteorological station network in remote regions of Brazil hinder the assessment of the feasibility of renewable energy projects. This study investigates spatial interpolation strategies, including IDW, Kriging, and Random Forest, applied to the analysis of climatic variables relevant to renewable energy projects, with a particular focus on wind and solar energy production. In this context, interpolation techniques were applied to data from 12 stations of the Brazilian National Institute of Meteorology, covering variables such as wind speed, solar radiation, and maximum and minimum temperatures in the state of Pernambuco. These techniques were evaluated through cross-validation and compared against data from actual power plants, revealing the superior performance of the IDW method, especially in estimating wind speed. The results highlight spatial interpolation as a technical and strategic tool for sustainable energy planning, providing accurate data in regions with low measurement density.

Key-words: Hybrid models; Supervised learning; Geospatial analysis; Energy feasibility.

Sumário

| | |
|---|-----------|
| 1 Introdução..... | 9 |
| 1.1 Motivação e contexto..... | 9 |
| 1.2 Problema de pesquisa..... | 10 |
| 1.3 Justificativa..... | 11 |
| 1.4 Objetivo geral..... | 12 |
| 1.4.1 Objetivos específicos..... | 12 |
| 2 Fundamentação teórica..... | 13 |
| 2.1 Energias renováveis e contexto climático..... | 13 |
| 2.2 Técnicas de interpolação espacial para dados climáticos..... | 15 |
| 2.3 Índice de Potencial de Produção de Energia Híbrida (IPPEH)..... | 19 |
| 2.4 Métricas para Avaliação dos Modelos..... | 20 |
| 3 Trabalhos relacionados..... | 22 |
| 4 Metodologia..... | 25 |
| 4.1 Dados..... | 26 |
| 4.2 Interpolação..... | 30 |
| 4.3 Seleção de estações para validação final..... | 31 |
| 4.4 Ambiente de Experimentação e Configuração dos Modelos..... | 32 |
| 5 Resultados..... | 34 |
| 5.1 Validação com usinas reais..... | 39 |
| 5.2 Discussão dos resultados..... | 42 |
| 5.3 Limitações e perspectivas..... | 43 |
| 6 Conclusão..... | 45 |
| Referências..... | 46 |
| Apêndice A – Interpolação em Diferentes Cidades..... | 48 |

Lista de ilustrações

| | |
|--|----|
| Figura 1 - Mapa exemplificando a interpolação de cidades no estado de Pernambuco..... | 16 |
| Figura 2 - Mapa com a distribuição das estações meteorológicas automáticas localizadas no estado de Pernambuco..... | 26 |
| Figura 3 - Histogramas das distribuições das variáveis (Radiação solar, velocidade do vento, temperatura mínima e temperatura máxima)..... | 29 |
| Figura 4 - Boxplot para análise de outliers das variáveis (Radiação solar, velocidade do vento, temperatura mínima e temperatura máxima)..... | 30 |
| Figura 5 - Mapa com a distribuição de projetos de energia renováveis implantados ou em implantação no estado de Pernambuco..... | 32 |
| Figura 6 - Gráfico de comportamento dos métodos na variável velocidade dos ventos..... | 35 |
| Figura 7 - Gráfico de comportamento dos métodos na variável radiação global..... | 36 |
| Figura 8 - Gráfico de comportamento dos métodos na variável temperatura máxima..... | 37 |
| Figura 9 - Gráfico de comportamento dos métodos na variável temperatura mínima..... | 37 |
| Figura 10 - Resultados do IPPEH para as cidades interpoladas..... | 41 |

Lista de tabelas

| | |
|---|-----------|
| Tabela 1 - Resultados das interpolações aplicadas à cidade de Ibimirim..... | 39 |
| Tabela 2 - Resultados das interpolações aplicadas ao conjunto de cidades selecionadas do estado de Pernambuco..... | 40 |
| Tabela A.1 - Resultados das Interpolações Aplicadas à Cidade de Arcoverde..... | 48 |
| Tabela A.2 - Resultados das Interpolações Aplicadas à Cidade de Cabrobó..... | 49 |
| Tabela A.3 - Resultados das Interpolações Aplicadas à Cidade de Caruaru..... | 50 |
| Tabela A.4 - Resultados das Interpolações Aplicadas à Cidade de Floresta..... | 51 |
| Tabela A.5 - Resultados das Interpolações Aplicadas à Cidade de Garanhuns..... | 52 |
| Tabela A.6 - Resultados das Interpolações Aplicadas à Cidade de Ouricuri..... | 53 |
| Tabela A.7 - Resultados das Interpolações Aplicadas à Cidade de Palmares..... | 54 |
| Tabela A.8 - Resultados das Interpolações Aplicadas à Cidade de Petrolina..... | 55 |
| Tabela A.9 - Resultados das Interpolações Aplicadas à Cidade de Salgueiro..... | 56 |
| Tabela A.10 - Resultados das Interpolações Aplicadas à Cidade de Serra Talhada..... | 57 |
| Tabela A.11 - Resultados das Interpolações Aplicadas à Cidade de Surubim..... | 58 |

1 INTRODUÇÃO

1.1 MOTIVAÇÃO E CONTEXTO

O contexto das energias renováveis tem se tornado cada vez mais relevante em virtude da busca por fontes sustentáveis de energia e da necessidade de mitigar os efeitos das mudanças climáticas. Em particular, a produção de energia solar e eólica se destacam como alternativas promissoras, especialmente em regiões como o Nordeste do Brasil, onde há grande potencial para a geração de energia a partir dessas fontes renováveis (FERREIRA et al., 2023).

No entanto, a avaliação precisa da viabilidade desses projetos depende diretamente de dados climáticos confiáveis. A velocidade dos ventos, por exemplo, é comumente obtida por meio de estações meteorológicas e desempenha um papel crucial nesse processo (SANTOS; VALENTIM, 2023). De forma análoga, a distribuição da radiação solar também precisa ser monitorada continuamente, uma vez que influencia diretamente o desempenho de sistemas fotovoltaicos. Contudo, a cobertura geográfica dessas estações é muitas vezes insuficiente, especialmente em áreas remotas ou de difícil acesso, gerando lacunas nos dados. Essas falhas de cobertura espacial tornam a interpolação espacial de dados climáticos uma ferramenta essencial para preencher essas lacunas e possibilitar uma análise mais completa.

A escolha das estratégias de interpolação pode impactar diretamente a qualidade dos resultados, especialmente em estudos relacionados ao potencial de produção de energia renovável, conforme revisado por Perin et al. (2015) para variáveis como temperatura e precipitação, e por Apaydin et al. (2004) em um contexto similar. Essa importância se estende à análise de dados para projetos de energia renovável, que dependem fortemente de estimativas precisas dessas variáveis climáticas.

1.2 PROBLEMA DE PESQUISA

O problema da interpolação espacial de medidas climáticas decorre da distribuição desigual e, muitas vezes, insuficiente das estações meteorológicas, especialmente em regiões de grande extensão geográfica ou em áreas remotas (SILVA JÚNIOR et al., 2019; APAYDIN et al., 2004). Essa limitação gera uma escassez de dados precisos para análises que dependem diretamente dessas informações, como a avaliação da viabilidade de projetos de energias renováveis.

Dados climáticos detalhados, como a velocidade do vento, são fundamentais para estimar a capacidade de geração de energia em projetos eólicos (SANTOS; VALENTIM, 2023), e, de forma análoga, dados de radiação solar são cruciais para projetos solares. No entanto, quando os dados são incompletos ou mal distribuídos, faz-se necessário o uso de técnicas de interpolação para estimar as variáveis em áreas onde não há medições diretas (PERIN et al., 2015).

No contexto das energias renováveis, os resultados da interpolação espacial inadequada podem levar a erros significativos na estimativa da disponibilidade dos recursos naturais. Isso pode impactar negativamente o planejamento e a implementação de tecnologias como painéis solares e turbinas eólicas, resultando em falhas no dimensionamento e na projeção de desempenho dos sistemas (FERREIRA et al., 2023). Portanto, a escolha da técnica de interpolação correta é crucial para garantir a acurácia das previsões e, consequentemente, o sucesso dos projetos.

Assim, como problema de pesquisa, busca-se identificar estratégias de interpolação espacial para garantir a precisão na análise de medidas climáticas utilizadas em projetos de energias renováveis. As questões centrais a serem exploradas incluem:

- Quais são as principais técnicas de interpolação espacial disponíveis para análise de dados climáticos?
- Quais fatores influenciam a escolha de uma técnica de interpolação em áreas com distribuição desigual de estações meteorológicas?

- Como diferentes modelos preditivos aplicados à interpolação espacial afetam a qualidade das estimativas para recursos eólicos e solares?
- Como as medidas interpoladas podem auxiliar nas estimativas de potencial de produção de energia em uma região?

Dessa forma, a pergunta norteadora da pesquisa é: “Quais são as melhores estratégias de interpolação espacial para análise de medidas climáticas no contexto de energias renováveis?”.

1.3 JUSTIFICATIVA

A relevância desta pesquisa reside no fato de que a geração de energia limpa e sustentável é uma prioridade global. Para que projetos solares e eólicos sejam viáveis e eficientes, é fundamental ter conhecimento de informações climáticas precisas que permitam um planejamento adequado (FERREIRA et al., 2023). No entanto, para expansão das soluções, é necessário que elas sejam economicamente viáveis e possam se apoiar em uma infraestrutura de monitoramento climático, que por sua vez, não são capazes de cobrir grandes extensões territoriais.

As técnicas de interpolação são essenciais para lidar com a ausência de dados meteorológicos em regiões mal atendidas por estações de medição, tornando-se, assim, uma ferramenta indispensável na análise de viabilidade energética (PERIN et al., 2015; APAYDIN et al., 2004). Além disso, a pesquisa que avalia a eficácia de diferentes técnicas de interpolação espacial, como a comparação entre métodos convencionais e de aprendizado de máquina para variáveis agroclimáticas (SILVA JÚNIOR et al., 2019), contribui para o avanço científico e técnico na modelagem de fenômenos ambientais, conhecimento este que pode ser estendido para otimizar estudos no setor de energias renováveis (FERREIRA et al., 2023; SANTOS; VALENTIM, 2023).

Uma melhor precisão na análise climática resultará em maior confiabilidade nos projetos de energias renováveis, otimizando investimentos e mitigando riscos de projeto. Assim, esta pesquisa é relevante tanto para a academia, que se beneficiará

de uma análise das metodologias de interpolação espacial, quanto para o mercado de energias renováveis, que poderá adotar técnicas mais adequadas e precisas na avaliação de seus projetos.

1.4 OBJETIVO GERAL

Avaliar estratégias de interpolação espacial aplicadas a medidas climáticas para análise de viabilidade de projetos de energias renováveis, em especial para energia solar e eólica.

1.4.1 Objetivos específicos

- Identificar as principais técnicas de interpolação espacial utilizadas na análise de dados climáticos;
- Avaliar o desempenho dessas técnicas de interpolação em áreas com distribuição desigual de estações meteorológicas;
- Analisar o potencial de uma região para geração de energia renovável com base em dados interpolados.

2 FUNDAMENTAÇÃO TEÓRICA

Neste Capítulo busca-se apresentar os conceitos e fundamentos necessários para compreender o desenvolvimento do trabalho. Com base na literatura, são apresentados tópicos que tratam da relevância das energias renováveis, da importância das variáveis climáticas e das técnicas de interpolação espacial, além do índice que mede o potencial de geração, que pode ser levado em conta para definir a viabilidade de novos projetos.

2.1 ENERGIAS RENOVÁVEIS E CONTEXTO CLIMÁTICO

A busca por fontes de energia sustentável tem se tornado uma prioridade global diante da crise climática e da necessidade de reduzir as emissões de gases de efeito estufa. O setor de energias renováveis, especialmente energia solar e eólica, destaca-se como uma solução viável e mais sustentável. Ferreira et al. (2023) destacam o papel dessas fontes de energia no cenário brasileiro, com ênfase na região Nordeste, devido ao elevado potencial solar e eólico dessa região. A disponibilidade de recursos naturais abundantes, como a radiação solar intensa e ventos constantes, posiciona o Brasil como um dos países mais promissores para a transição energética global.

A implementação de projetos de energia renovável depende de uma análise detalhada da viabilidade, considerando fatores técnicos, econômicos e ambientais. Um planejamento energético eficaz requer dados precisos que permitam identificar as regiões mais adequadas para a instalação de sistemas como painéis fotovoltaicos e turbinas eólicas (SANTOS; VALENTIM, 2023). No entanto, a coleta de dados climáticos em larga escala enfrenta desafios, como a distribuição desigual de estações meteorológicas, especialmente em áreas remotas (BEZERRA et al., 2019). Essa limitação torna indispensável o uso de técnicas que permitam a extrapolação dos dados geolocalizados disponíveis para regiões onde medições diretas não estão disponíveis. Além disso, o contexto climático é fortemente influenciado pelas mudanças globais, que afetam padrões de radiação solar e regimes de ventos em

diversas regiões do planeta. Isso reforça a necessidade de estudos contínuos que integrem modelos preditivos e medições constantes para acompanhar as mudanças nas condições ambientais e garantir o sucesso de novos empreendimentos energéticos.

Por sua vez, as variáveis climáticas desempenham um papel fundamental no planejamento e na operação de projetos de energia renovável. Parâmetros como a radiação solar, a velocidade e constância dos ventos, a temperatura e a precipitação são amplamente utilizados para estimar o potencial de geração e dimensionar adequadamente sistemas solares e eólicos. Segundo Ferreira et al. (2023), o monitoramento contínuo dessas variáveis é essencial para garantir uma produção energética segura e eficiente, considerando a natureza intermitente e sazonal desses recursos. No caso da energia eólica, compreender o comportamento estatístico do vento, como sua média e variabilidade, permite a escolha de locais com condições mais estáveis, o que impacta diretamente na qualidade e no aproveitamento do recurso. De forma análoga, a geração de energia solar está diretamente relacionada à irradiação solar, cuja variabilidade exige acompanhamento constante para reduzir perdas e otimizar a eficiência de sistemas fotovoltaicos.

Além de influenciarem diretamente o desempenho dos sistemas de geração, como painéis fotovoltaicos e turbinas eólicas, as variáveis climáticas exigem um monitoramento preciso e constante. A obtenção de dados confiáveis sobre radiação solar e velocidade do vento é essencial para o dimensionamento adequado desses sistemas, tanto em projetos *onshore* (em terra firme) quanto *offshore* (no mar). Contudo, essa tarefa ainda enfrenta obstáculos significativos em regiões de difícil acesso, onde a infraestrutura para medições meteorológicas é limitada ou inexistente.

A ausência de medições diretas leva ao uso de técnicas que estimem variáveis climáticas em locais sem estações meteorológicas. Perin et al. (2015) apontam que o uso de métodos de interpolação permite a construção de mapas climáticos detalhados. Esses mapas fornecem uma visão mais abrangente dos recursos

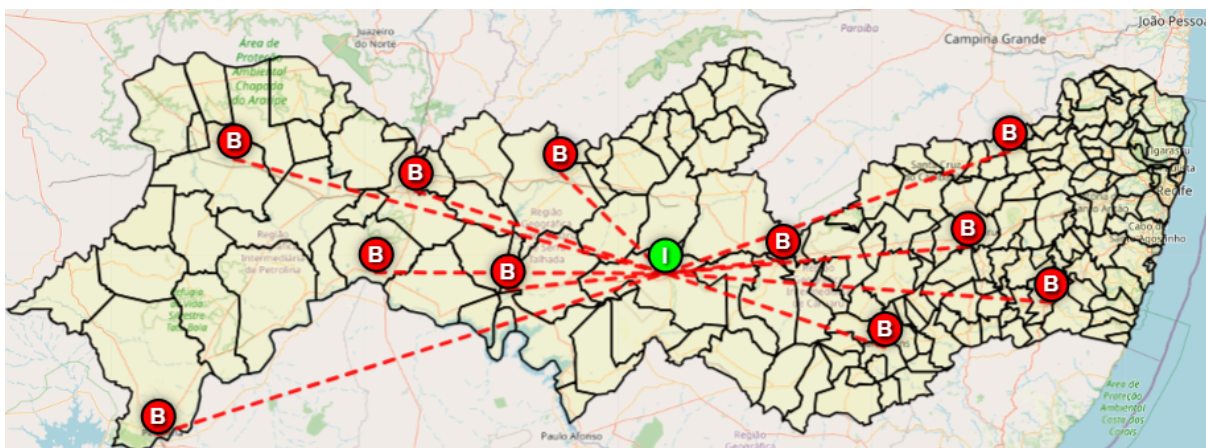
disponíveis, orientando decisões estratégicas e otimizando investimentos em energia renovável.

Outro fator relevante é a influência das características geográficas na distribuição das variáveis climáticas. Regiões de maior altitude, por exemplo, tendem a apresentar maior variabilidade nos padrões de vento e precipitação, o que exige o uso de modelos preditivos mais sofisticados. Da mesma forma, áreas costeiras ou interiores podem apresentar condições climáticas únicas, que impactam diretamente o desempenho de sistemas de geração de energia.

2.2 TÉCNICAS DE INTERPOLAÇÃO ESPACIAL PARA DADOS CLIMÁTICOS

A interpolação espacial é uma técnica estatística utilizada para estimar valores de variáveis em locais onde não há dados observados, com base nas informações coletadas em pontos próximos. Segundo Apaydin et al. (2004), essa abordagem é amplamente utilizada na climatologia para preencher lacunas em conjuntos de dados, especialmente em regiões com distribuição desigual de estações meteorológicas. A interpolação permite criar mapas detalhados de variáveis climáticas, como radiação solar e velocidade do vento, fornecendo informações essenciais para o planejamento de projetos de energia renovável.

Na figura 1, observa-se um exemplo da aplicação da interpolação espacial no estado de Pernambuco. Os pontos vermelhos (B) representam todas as cidades da base de dados que foram utilizadas como referência para a interpolação das variáveis climáticas, enquanto o ponto verde (I) indica a cidade cujo dado foi interpolado a partir das informações das demais localidades. As linhas tracejadas ilustram a relação espacial entre os pontos de coleta e o ponto de interpolação, evidenciando o funcionamento do processo de estimativa em áreas com ausência de medições diretas.



B - Representa todas as cidades na base de dados que são usadas para fazer a interpolação das variáveis; **I** - Representa a cidade cujo os dados foram interpolados

Figura 1 - Mapa exemplificando a interpolação de cidades no estado de Pernambuco.

Fonte: Próprio autor

Entre as principais técnicas de interpolação espacial estão a Krigagem, o método do inverso do quadrado da distância (IDW) e a literatura tem apresentado diversos algoritmos de aprendizado de máquina, como o Random Forest (SILVA JÚNIOR et al., 2019). A escolha da técnica depende de fatores como a densidade dos dados disponíveis, a variabilidade espacial das variáveis e as características geográficas da região.

A Krigagem é amplamente reconhecida por sua capacidade de fornecer estimativas precisas, mesmo em áreas com baixa densidade de medições. Ferreira et al. (2023) destacam que este método geoestatístico considera a estrutura espacial dos dados, utilizando variogramas para modelar a correlação entre pontos amostrados. Essa técnica é especialmente eficaz em regiões montanhosas ou com grande variabilidade topográfica, onde métodos mais simples, como o IDW, tendem a falhar.

No método de Krigagem, os dados interpolados são modelados como um processo gaussiano, baseado em uma função de covariância estimada a partir do variograma. A equação utilizada no estudo de Ferreira et al. (2023) é expressa da seguinte forma:

$$\hat{Y}(h) = c_0 + C \left[1 - \exp\left(-\frac{h}{a}\right) \right] \quad (2.1)$$

Onde:

- $\hat{Y}(h) \rightarrow$ valor interpolado para uma distância h ;
- $c_0 \rightarrow$ variância quando a distância entre pontos tende a zero;
- $C \rightarrow$ variância total ou o limite superior da função de variograma;
- $a \rightarrow$ distância a partir da qual as amostras não são mais espacialmente correlacionadas
- $h \rightarrow$ distância entre pontos amostrais

Esse modelo exponencial permite estimar valores em qualquer posição do campo analisado, com mínima variância e sem viés, considerando que a correlação espacial diminui com o aumento da distância entre os pontos.

O método IDW (Inverse Distance Weighted), por outro lado, é uma técnica mais simples e de fácil aplicação, que atribui pesos inversamente proporcionais à distância entre os pontos amostrados e o ponto a ser interpolado. Embora seja menos sofisticado que a Krigagem, pode ser uma alternativa eficaz em contextos com maior densidade de medições. O cálculo do IDW é feito com as seguintes expressões:

$$\hat{Y}(s_0) = \sum_{i=1}^N \lambda_i Z(s_i) \quad (2.2)$$

$$\lambda_i = \frac{d_{i0}^{-2}}{\sum_{i=1}^N d_{i0}^{-2}} \quad (2.3)$$

Onde:

- $\hat{Y}(s_0) \rightarrow$ valor interpolado no ponto s_0 ;
- $Z(s_i) \rightarrow$ valor conhecido no ponto s_i ;

- $\lambda_i \rightarrow$ peso associado ao ponto i , calculado com base no inverso do quadrado da distância;
- $d_{i0} \rightarrow$ distância entre o ponto conhecido s_i e o ponto de interpolação s_0
- $N \rightarrow$ número de vizinhos mais próximos considerados na interpolação

O estudo de Silva Júnior et al. (2019) demonstrou que algoritmos modernos como o Random Forest, mesmo utilizando apenas as coordenadas espaciais (latitude e longitude) como variáveis de entrada, aumentaram significativamente a precisão das estimativas em comparação com os métodos convencionais. Esse modelo de aprendizado de máquina baseia-se na construção de um conjunto de árvores de decisão (*ensemble*), cada uma treinada com uma amostra aleatória dos dados (*bootstrap*). Além disso, o Random Forest é uma técnica de aprendizado supervisionado, capaz de resolver tanto problemas de classificação quanto de regressão, o que o torna extremamente versátil para diferentes tipos de aplicações em ciência de dados (Breiman, 2001).

O Random Forest realiza a predição como a média das predições individuais das árvores, o que reduz a variância do modelo e melhora sua generalização, especialmente em contextos com dados climáticos escassos e espacialmente dispersos. A representação conceitual da predição no Random Forest é dada por:

$$\hat{Y}(x) = \frac{1}{M} \sum_{m=1}^M T_m(x) \quad (2.4)$$

Onde:

- $\hat{Y}(x) \rightarrow$ valor predito para a entrada x ;
- $M \rightarrow$ número total de árvores na floresta;
- $T_m(x) \rightarrow$ predição da m -ésima árvore de decisão para a entrada x ;

2.3 ÍNDICE DE POTENCIAL DE PRODUÇÃO DE ENERGIA HÍBRIDA (IPPEH)

O IPPEH é um índice apresentado por Ferreira et al. (2023), originalmente descrito como IP-PB (Índice de Potencial de Geração Híbrida da Paraíba), com o objetivo de sintetizar, em um único valor, o potencial de geração de energia a partir da combinação de duas variáveis fundamentais: a radiação solar e a velocidade do vento. Neste trabalho o índice foi chamado de IPPEH, buscando generalizar a avaliação do índice para além do estado da Paraíba. O processo de cálculo do IPPEH envolve, inicialmente, a normalização destas variáveis, seguida pela aplicação de uma função agregadora (FERREIRA et al., 2023).

O uso do IPPEH neste trabalho se deu pela semelhança entre os estados de Pernambuco e Paraíba no que diz respeito às características climáticas e topográficas, especialmente nas regiões do Agreste e Sertão, onde predomina o clima semiárido, a elevada incidência solar e ventos constantes em áreas de relevo elevado. Ao condensar as duas variáveis mais relevantes para sistemas solares e eólicos em um único valor, o IPPEH facilita a comparação entre municípios e contribui para decisões mais estratégicas na escolha de áreas para instalação de novos empreendimentos de energia renovável.

A normalização é realizada com o intuito de transformar os valores originais em uma escala padronizada entre 0 e 1, eliminando as diferenças de unidade de medida e facilitando a comparação e a combinação das variáveis. O cálculo da normalização de cada variável (radiação solar e velocidade do vento) é efetuado utilizando a seguinte expressão:

$$Z = \frac{(X - X_{min})}{(X_{max} - X_{min})} \quad (2.5)$$

Onde:

- $Z \rightarrow$ valor normalizado;
- $X \rightarrow$ valor interpolado;
- $X_{min} \rightarrow$ menor valor da variável na base de dados;

- $X_{max} \rightarrow$ maior valor da variável na base de dados;

Esse procedimento é aplicado individualmente para cada uma das variáveis, garantindo que ambas contribuam proporcionalmente para o cálculo final do índice. Após a obtenção dos valores normalizados da radiação solar e da velocidade do vento, procede-se ao cálculo do IPPEH, que é definido pela seguinte fórmula:

$$IPPEH = \sqrt{Z_{t_wind} + Z_{t_rad}} \quad (2.6)$$

Este método proporciona uma abordagem simples, mas eficaz, para a avaliação do potencial energético de uma determinada localidade, permitindo uma análise comparativa entre diferentes pontos geográficos com base em seus recursos naturais.

2.4 MÉTRICAS PARA AVALIAÇÃO DOS MODELOS

Para avaliar a precisão dos modelos de interpolação utilizados, foram aplicadas três métricas estatísticas amplamente reconhecidas: o Erro Quadrático Médio (MSE), a Raiz do Erro Quadrático Médio (RMSE) e o Erro Médio Absoluto (MAE). A seguir, são apresentadas as fórmulas e a interpretação de cada métrica:

O MSE quantifica a média dos quadrados dos erros, ou seja, a diferença entre os valores observados e os valores estimados. É uma métrica sensível a grandes erros, já que eleva as diferenças ao quadrado.

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2$$

O RMSE é a raiz quadrada do MSE. Ele mantém a unidade de medida original da variável, o que facilita a interpretação prática. Assim como o MSE, penaliza mais fortemente os grandes erros.

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

O MAE representa a média das diferenças absolutas entre os valores observados e os valores estimados. Por não elevar os erros ao quadrado, o MAE é menos sensível a outliers e proporciona uma visão mais equilibrada dos erros médios.

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i|$$

Onde:

- y_i é o valor observado,
- \hat{y}_i é o valor estimado pelo modelo,
- n é o número total de observações.

Essas métricas são fundamentais para comparar o desempenho dos diferentes modelos testados, permitindo identificar quais abordagens oferecem maior precisão nas estimativas de radiação solar e velocidade do vento.

3 TRABALHOS RELACIONADOS

Diversos estudos têm investigado métodos de interpolação espacial para a análise de variáveis climáticas, com o objetivo de melhorar a precisão dos dados utilizados em projetos de energias renováveis. A análise dessas variáveis, como radiação solar e velocidade do vento, é essencial para o planejamento de sistemas de energia solar e eólica, especialmente em regiões com distribuição irregular de estações meteorológicas. Tais lacunas de dados podem impactar significativamente a precisão das análises, tornando as técnicas de interpolação uma ferramenta indispensável para garantir resultados confiáveis.

Ferreira et al. (2023) propuseram um índice inovador para avaliar o potencial de energias renováveis, integrando dados de energia solar, eólica e híbrida, aplicados à região Nordeste do Brasil. O estudo destacou a escassez de estações meteorológicas na região, um problema comum em áreas remotas e de difícil acesso. Para superar essa limitação, foram utilizadas técnicas de interpolação espacial para integrar dados climáticos provenientes de diferentes fontes. Entre os métodos avaliados, a Krigagem foi escolhida devido à sua capacidade de gerar estimativas mais precisas em áreas com baixa densidade de medições. A pesquisa concluiu que o índice desenvolvido, combinado com a aplicação da Krigagem, foi eficaz para identificar regiões com maior potencial para a implantação de sistemas de energias renováveis, fornecendo subsídios valiosos para o planejamento energético na região.

No estudo de Perin et al. (2015), os autores realizaram uma revisão abrangente sobre os métodos de interpolação mais utilizados para variáveis climáticas, como temperatura e precipitação. O artigo destacou a popularidade de técnicas como a Krigagem ordinária e o método do inverso do quadrado da distância (IDW), amplamente aplicadas em estudos climáticos. Além disso, enfatizou a importância de ajustar os modelos conforme as características específicas da região de estudo, como a densidade das estações meteorológicas e a variabilidade espacial das variáveis climáticas. Uma das principais conclusões foi que o método mais eficiente varia conforme o contexto. Para a precipitação, o IDW foi mais eficaz para dados diários, a Krigagem para dados mensais e a Cokrigagem para dados anuais. Para a

temperatura, embora a Krigagem seja muito usada, o estudo apontou que o método da regressão apresentou o maior índice de eficiência, especialmente pela sua capacidade de incorporar a influência da altitude em regiões montanhosas. Essa flexibilidade metodológica destaca a importância de escolher a técnica mais adequada para cada variável e período específico da pesquisa.

O estudo de Silva Júnior et al. (2019) avaliou a aplicação de algoritmos de aprendizado de máquina para a interpolação espacial de dados de evapotranspiração no Nordeste do Brasil. Para isso, os autores compararam o desempenho de métodos convencionais, como a IDW e a Krigagem Ordinária, com estratégias baseadas em aprendizado de máquina, representadas pelo Random Forest. Os modelos foram desenvolvidos utilizando dados de evapotranspiração de estações meteorológicas automáticas do INMET, referentes a janeiro de 2017. O modelo Random Forest utilizou apenas as coordenadas (latitude e longitude) das estações como variáveis de entrada, e não dados adicionais, de topografia ou solo, e obteve uma precisão superior à das abordagens convencionais, chegando a reduzir o erro de estimativa (RMSE e MAE) aproximadamente pela metade.

Apaydin et al. (2004) investigaram diferentes métodos de interpolação espacial para variáveis climáticas na região do GAP, na Turquia. Os autores compararam diversas técnicas, incluindo Krigagem, IDW e Spline, avaliando-as em termos de precisão e eficiência para estimar variáveis como temperatura e precipitação. Os resultados mostraram que a Krigagem gerou as melhores estimativas em áreas com maior variabilidade topográfica, enquanto o IDW demonstrou melhor desempenho em regiões planas ou com menor variação espacial. O estudo destacou que a escolha da técnica deve considerar não apenas a topografia, mas também as características das variáveis analisadas e a densidade dos pontos de dados disponíveis. Essa abordagem evidencia a importância de personalizar as técnicas de interpolação para cada contexto de estudo.

O estudo de Santos e Valentim (2023) abordou a estimativa da velocidade do vento em regiões costeiras do Nordeste brasileiro, comparando dados de simulação com medições diretas onshore e offshore. A pesquisa enfatizou que a ausência de medições diretas pode limitar a avaliação do potencial energético, especialmente em

locais remotos. O trabalho reforça, portanto, que a simulação numérica constitui um método viável e estratégico para gerar informações confiáveis, sendo fundamental para delimitar áreas viáveis e otimizar o planejamento de parques eólicos.

No geral, os estudos analisados destacam a relevância das técnicas de interpolação espacial para melhorar a precisão das análises climáticas aplicadas a projetos de energias renováveis. Seja por meio de métodos tradicionais, como a Krigagem e o IDW, ou por abordagens modernas, como o aprendizado de máquina, essas técnicas têm se mostrado indispensáveis para lidar com os desafios de variabilidade geográfica e lacunas de dados. Além disso, a escolha criteriosa da técnica de interpolação, alinhada às características da região e das variáveis analisadas, desempenha um papel central na garantia de resultados confiáveis e na promoção do uso sustentável de recursos energéticos.

4 METODOLOGIA

Este trabalho é classificado como uma pesquisa aplicada, pois visa resolver um problema prático específico: identificar e avaliar estratégias de interpolação espacial para preencher lacunas em dados climáticos, essenciais para a análise de viabilidade de projetos de energias renováveis. A pesquisa busca gerar resultados que possam ser aplicados diretamente ao planejamento de instalações solares e eólicas, contribuindo para soluções sustentáveis e eficientes. A abordagem é quantitativa, pois envolve a análise de dados climáticos, como radiação solar e velocidade do vento, e a comparação de métodos de interpolação, utilizando métricas objetivas para avaliar a precisão. Os objetivos são descritivos, ao caracterizar as técnicas de interpolação espacial, e comparativos, ao avaliar o desempenho dessas técnicas em cenários de diferentes distribuições de estações meteorológicas. No que diz respeito aos objetivos, a pesquisa é classificada como descritiva e comparativa. Ela é descritiva porque procura detalhar as principais técnicas de interpolação espacial, explicando suas características e condições de uso. Ao mesmo tempo, é também comparativa, pois visa comparar o desempenho de métodos distintos de interpolação, como Krigagem, IDW e Random Forest, em cenários com distribuição desigual de estações meteorológicas, buscando identificar as técnicas que geram estimativas mais precisas para dados climáticos.

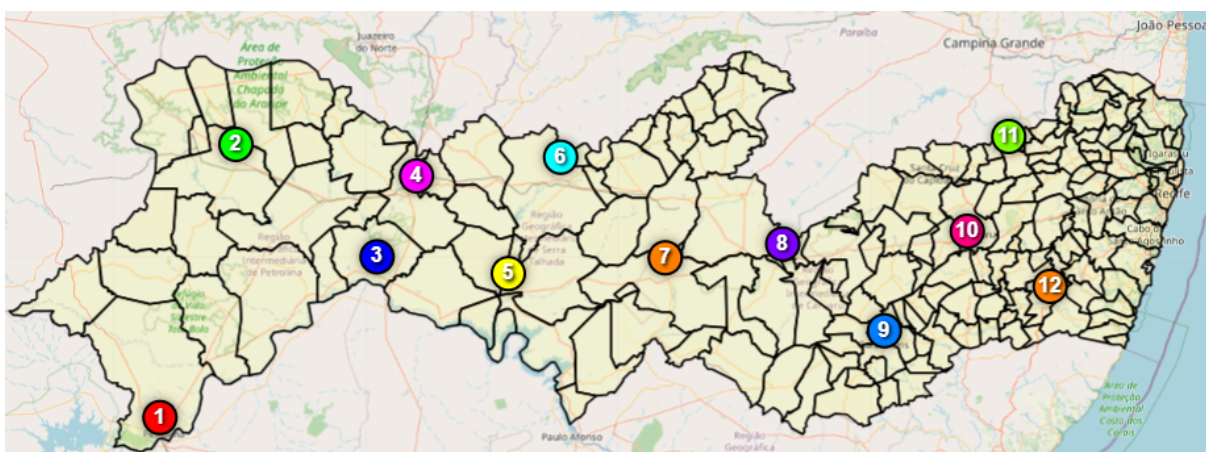
Os procedimentos técnicos envolvem levantamento bibliográfico sobre os estudos existentes de interpolação espacial aplicada a dados climáticos, para embasar teoricamente o trabalho. Além disso, inclui a coleta e organização de dados climáticos de fontes confiáveis, como bases públicas, por exemplo, o INMET, contendo medições de radiação solar e velocidade do vento. A aplicação das técnicas de interpolação espacial, como Krigagem, IDW e Random Forest, será realizada com o uso de ferramentas computacionais e métodos estatísticos apropriados. A avaliação quantitativa das técnicas será feita com o uso de métricas estatísticas como o Erro Quadrático Médio (MSE), a Raiz do Erro Quadrático Médio (RMSE) e o Erro Médio Absoluto (MAE), para comparar as estimativas interpoladas com dados reais ou simulados. A metodologia combina rigor quantitativo e

aplicabilidade prática, com o objetivo de fornecer *insights* claros e confiáveis para o planejamento de sistemas de energia renovável.

Como métrica adicional para verificar a efetividade das técnicas de interpolação na estimativa do potencial energético, foi utilizado neste trabalho o Índice de Potencial de Produção de Energia (IPPEH), proposto por Ferreira et al. (2023) e descrito nos pilares teóricos deste trabalho, que permite avaliar o potencial combinado de geração de energia solar e eólica. A seguir, as Subseções apresentam

4.1 DADOS

Os dados utilizados neste estudo foram coletados através do serviço do Instituto Nacional de Meteorologia (INMET), utilizando 12 estações meteorológicas automáticas localizadas no estado de Pernambuco, abrangendo os municípios de Petrolina, Ouricuri, Cabrobó, Salgueiro, Floresta, Serra Talhada, Ibimirim, Arcoverde, Garanhuns, Caruaru, Surubim e Palmares. A Figura 2 apresenta a distribuição geográfica das estações localizadas no estado de Pernambuco.



1- Petrolina; 2- Ouricuri; 3- Cabrobó; 4- Salgueiro; 5- Floresta; 6- Serra Talhada; 7- Ibimirim;
8- Arcoverde; 9- Garanhuns; 10- Caruaru; 11- Surubim; 12- Palmares.

Figura 2 - Mapa com a distribuição das estações meteorológicas automáticas localizadas no estado de Pernambuco.

Fonte: Próprio Autor

O serviço disponibiliza dados horários contendo 17 variáveis: Precipitação total; Pressão atmosférica ao nível da estação; Pressão atmosférica máxima; Pressão atmosférica mínima; Radiação global; Temperatura do ar; Temperatura do ponto de orvalho; Temperatura máxima; Temperatura mínima; Temperatura orvalho máxima; Temperatura orvalho mínima; Umidade relativa máxima; Umidade relativa mínima; Umidade relativa do ar; Direção do vento; Rajada do vento e Velocidade do vento.

Para cada cidade, a disponibilidade dos dados varia conforme o início da operação de cada estação. A mais antiga é a da cidade de Petrolina, que iniciou suas atividades em 20 de fevereiro de 2003. A mais recente, a da cidade de Salgueiro, teve suas operações iniciadas em 14 de setembro de 2017. Para este estudo, foram selecionadas as variáveis mais relevantes para o problema proposto:

- Velocidade do vento (m/s)
- Radiação global (kJ/m²)
- Temperatura máxima (°C)
- Temperatura mínima (°C)

Os dados foram coletados em formato CSV (*comma separated values*), e passaram por uma etapa de pré-processamento, que incluiu:

- A remoção de valores atípicos (outliers)
- O cálculo das médias anuais e históricas de cada variável por estação

Na etapa de remoção de outliers, foram adotadas duas abordagens complementares. A primeira consistiu na exclusão de exemplos com dados faltantes, nos quais uma variável apresentava informação enquanto outra não, resultando em um desbalanceamento na quantidade de dados disponíveis para cada variável.

A segunda abordagem envolveu a detecção e remoção de valores atípicos com base no desvio padrão, seguindo o seguinte procedimento:

- Passo 1: cálculo da média (μ) e do desvio padrão (σ) do conjunto de dados;

$$\mu = \frac{\sum_{i=1}^N X_i}{N} \quad (4.1) \quad \sigma = \sqrt{\frac{\sum_{i=1}^N (X_i - \mu)^2}{N}} \quad (4.2)$$

- Passo 2: identificação de valores que excedem o intervalo definido por $\mu \pm 3\sigma$, sendo estes classificados como outliers;
- Passo 3: exclusão dos valores atípicos identificados, a fim de assegurar maior qualidade e consistência aos dados utilizados nas análises subsequentes.

Esse tratamento visou padronizar e limpar os dados, aumentando a confiabilidade das análises subsequentes. É importante destacar também que os dados de temperatura máxima e temperatura mínima foram utilizados apenas para comparar o desempenho dos métodos de interpolação diante de diferentes variáveis climáticas. Os dados de temperatura não serão incorporados ao índice que analisa o potencial de geração de energia.

Além do pré-processamento descrito, foi realizada uma análise exploratória dos dados com o objetivo de compreender o comportamento estatístico das variáveis climáticas utilizadas neste trabalho. As variáveis analisadas foram: radiação global (KJ/m^2), velocidade do vento (m/s), temperatura máxima ($^{\circ}\text{C}$) e temperatura mínima ($^{\circ}\text{C}$).

Inicialmente, histogramas com curvas de densidade foram gerados para visualizar a distribuição dos dados (Figura 3). As temperaturas máxima e mínima apresentaram distribuições aproximadamente simétricas, com concentrações entre 24°C e 28°C , o que sugere estabilidade térmica entre os registros anuais. A velocidade do vento, por sua vez, apresentou distribuição quase normal, com média próxima de 3 m/s e pouca assimetria. Já a radiação global evidenciou uma distribuição assimétrica à direita (positiva), com valores concentrados entre 1400 e 1700 KJ/m^2 , mas com cauda longa indicando a presença de extremos superiores — comportamento típico de variáveis com forte dependência da sazonalidade solar.

Distribuição das Variáveis Climáticas

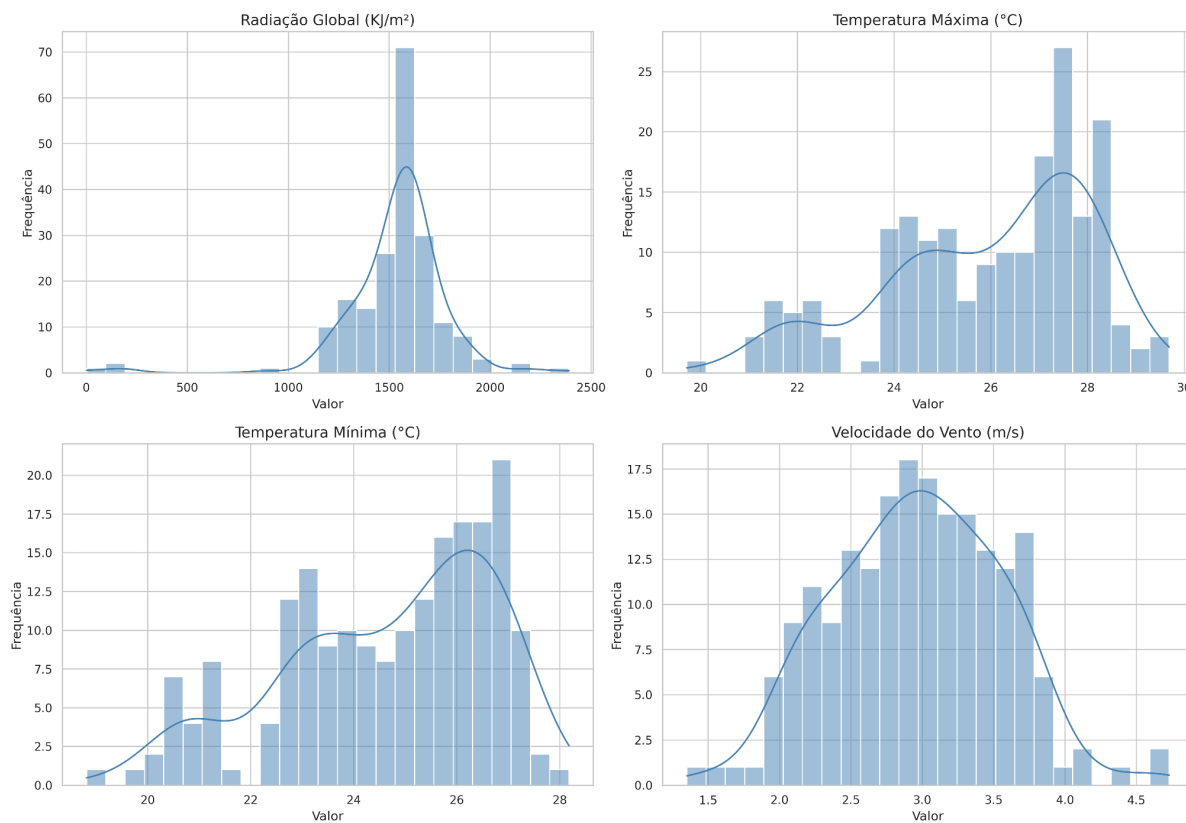


Figura 3 - Histogramas das distribuições das variáveis (Radiação solar, velocidade do vento, temperatura mínima e temperatura máxima)

Fonte: Próprio autor

Em seguida, como podemos visualizar na figura 4, foram analisados boxplots para identificação de valores atípicos (outliers). As variáveis radiação global e velocidade do vento apresentaram diversos pontos fora do intervalo interquartilício (IQR), reforçando a necessidade do processo de limpeza previamente descrito, como a exclusão de anos com medições anômalas (ex: 2005, 2006, 2007 e 2017). Temperaturas, por outro lado, exibiram dispersão controlada e sem ocorrência crítica de outliers.

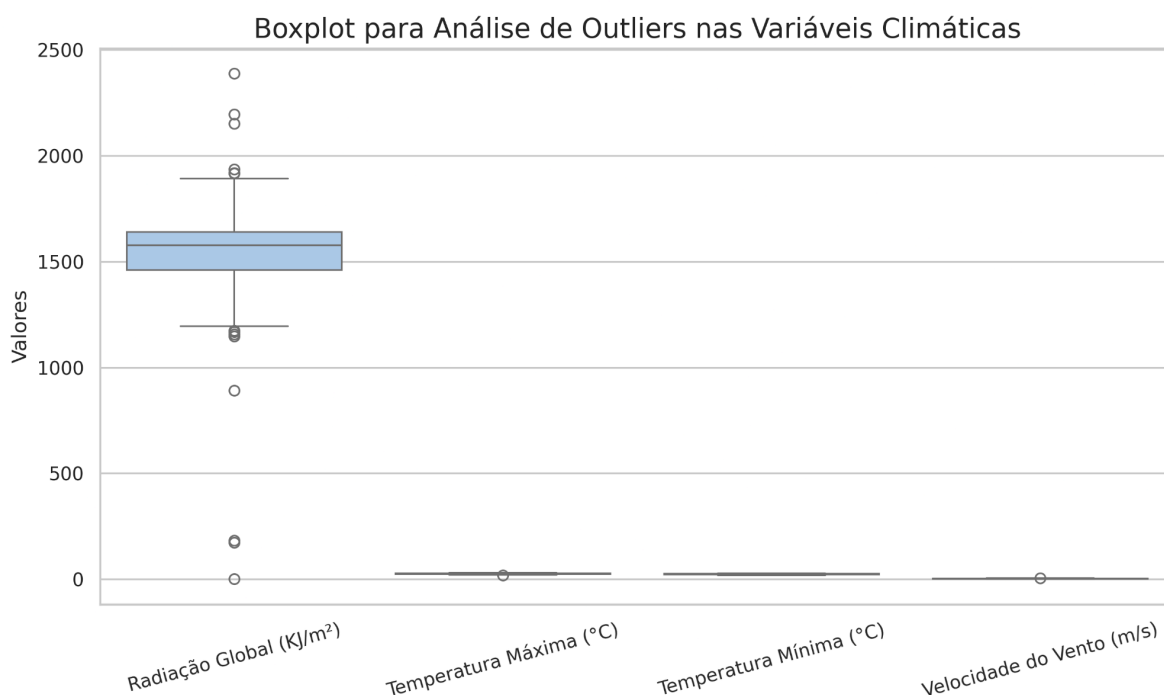


Figura 4 - Boxplot para análise de outliers das variáveis (Radiação solar, velocidade do vento, temperatura mínima e temperatura máxima)

Fonte: Próprio autor

Portanto, a EDA não apenas validou os procedimentos de pré-processamento adotados, como também forneceu subsídios técnicos para a escolha dos métodos de interpolação e a modelagem espacial das variáveis climáticas com maior robustez e coerência estatística.

4.2 INTERPOLAÇÃO

A interpolação espacial foi aplicada utilizando os métodos IDW, Krigagem e Random Forest. Para validar os modelos, adotou-se uma técnica de validação cruzada do tipo holdout espacial. Nesse processo, foi selecionada uma estação como ponto de teste, realizando a interpolação de suas variáveis com base apenas nas demais estações disponíveis. As estimativas obtidas pelos métodos de interpolação foram então comparadas com os valores reais observados. Com base na validação cruzada leave-one-out, foram calculadas métricas como o erro quadrático médio (MSE), para avaliar a média dos quadrados dos erros; a raiz do erro quadrático médio (RMSE), para identificar grandes discrepâncias de forma mais interpretável; e

o erro médio absoluto (MAE), para avaliar os erros médios absolutos das estimativas.

4.3 SELEÇÃO DE ESTAÇÕES PARA VALIDAÇÃO FINAL

Após a avaliação dos métodos de interpolação, foram selecionadas algumas localizações estratégicas onde já existem usinas eólicas e/ou solares em operação, conforme identificado no Atlas Eólico e Solar de Pernambuco, e aplicada a fórmula do índice IPPEH para validar seus resultados na análise do potencial de produção de energia solar e eólica. A presença dessas usinas foi considerada um indicador real da viabilidade energética das regiões, e os valores interpolados foram utilizados para estimar o potencial energético de uma determinada zona. Dado que o Atlas Eólico e Solar de Pernambuco não contempla atualizações recentes sobre os novos projetos de geração de energia, foram selecionadas três regiões ainda não mapeadas pelo atlas (Petrolina, São José do Belmonte e Flores), mas que receberam empreendimentos nos últimos anos e cuja implementação é amplamente conhecida^{1,2,3}.

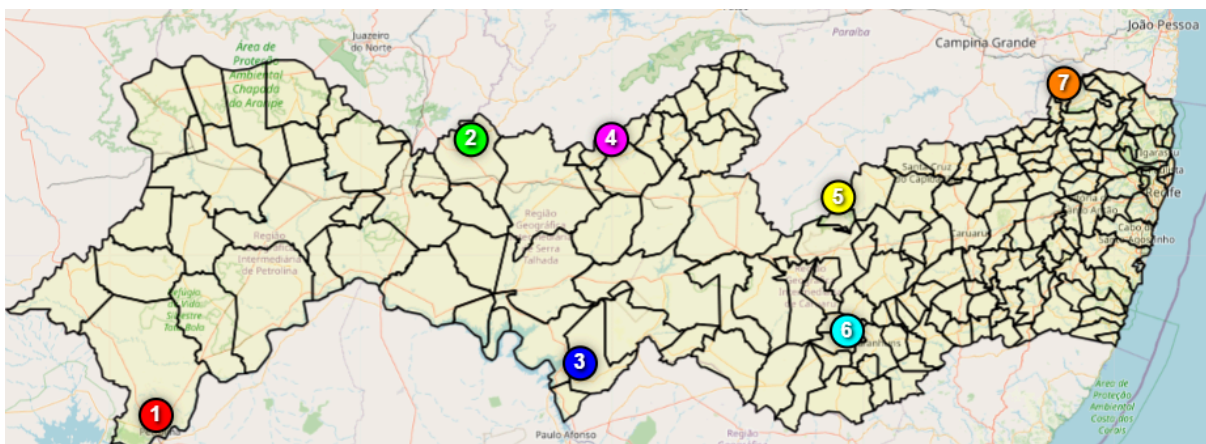
O atlas apresenta as cidades de Tacaratu, Poção, Paratama e Macaparana como aquelas que já possuem projetos em operação para a produção de energia eólica. Como o atlas não descreve nenhuma usina para produção de energia solar em operação, foram selecionadas consideradas as usinas de Petrolina, São José do Belmonte e Poção.

A distribuição destes pontos de validação é apresentada pela Figura 5.

¹ Disponível em: <<https://bit.ly/4mGwLkj>>. Acesso em: 29 de maio de 2025.

² Disponível em: <<https://bit.ly/4kJ0gA9>>. Acesso em: 29 de maio de 2025.

³ Disponível em: <<https://bit.ly/43MwN1x>>. Acesso em: 29 de maio de 2025.



1- Petrolina; 2- São José do Belmonte; 3- Tacaratu; 4- Flores; 5- Poção;
6- Paratama; 7- Macaparana.

Figura 5 - Mapa com a distribuição de projetos de energia renováveis implantados ou em implantação no estado de Pernambuco.

Fonte: Próprio autor

4.4 AMBIENTE DE EXPERIMENTAÇÃO E CONFIGURAÇÃO DOS MODELOS

A implementação das técnicas de interpolação espacial foi realizada em um ambiente Google Colab, utilizando a linguagem Python 3.10. A escolha desta plataforma se deu por sua capacidade de processamento em nuvem, integração nativa com o Google Drive para manipulação de arquivos e por facilitar o compartilhamento e a reprodutibilidade dos experimentos. Os artefatos gerados nesta etapa foram disponibilizados em um repositório do Github⁴.

As principais bibliotecas utilizadas foram:

- Pandas (v.1.5.3): Para tratamento e análise de dados estruturados.
- NumPy (v.1.22.4): Para operações matemáticas e manipulação de arrays.
- Math (nativa): Para funções matemáticas básicas.
- Scikit-learn (v.1.2.2): Para a implementação dos modelos Random Forest e Gaussian Process (Krigagem).

⁴ Disponível em: <https://github.com/RafaCarvalh0/Trabalho-de-Conclusao-de-Curso>

- Warnings (nativa): Para o tratamento de mensagens de alerta durante o ajuste dos modelos.

Para garantir a organização e a clareza do código, todos os scripts foram modularizados em funções específicas, incluindo os cálculos de interpolação (IDW, Krigagem e Random Forest), a avaliação das métricas de erro (MSE, RMSE e MAE) e o cálculo do índice IPPEH.

A escolha dos hiperparâmetros para os métodos de interpolação seguiu critérios práticos de estabilidade, desempenho e coerência com a literatura. Para o Random Forest foram utilizados $n_estimators = 100$, $max_depth = 5$ e $min_samples_split = 3$. Estes valores foram definidos para evitar o overfitting e garantir boa generalização.

O método do Inverso da Distância Ponderada (IDW) foi aplicado com um expoente de ponderação de 2. Este é um valor clássico, que atribui maior peso aos pontos de amostragem mais próximos (Bidgoli et al., 2023). Já no método da Krigagem foi adotado um kernel exponencial (RBF) combinado a um kernel constante da biblioteca `sklearn.gaussian_process`, com normalização dos dados (`StandardScaler`) e otimização automática do modelo ($n_restarts_optimizer = 10$).

Embora não tenha sido realizada uma busca sistemática por hiperparâmetros (como `grid search`), os valores adotados se mostraram eficazes, com base nos erros obtidos durante a validação cruzada `leave-one-out` e no desempenho consistente entre as variáveis climáticas. Os scripts e funções implementadas estão disponíveis e comentados no ambiente Google Colab, garantindo a reprodutibilidade dos experimentos.

5 RESULTADOS

Os resultados das interpolações indicam uma superioridade do método IDW na interpolação da radiação global, sobretudo pela sua maior constância e estabilidade nas estimativas, o que se mostra fundamental diante da alta volatilidade referente à variável de radiação solar, enquanto a Krigagem e o Random Forest apresentaram resultados bastante semelhantes, especialmente na interpolação da velocidade do vento.

Conforme detalhado no Apêndice A, o método IDW apresentou estimativas mais próximas dos valores observados para a radiação global. Na interpolação da velocidade do vento, embora ambos os métodos tivessem desempenhos próximos, a Krigagem se destacou pela constância e pelos menores erros. Seus resultados incluem um RMSE de apenas 0,05 em Ibimirim (Tabela 1) e valores baixos em Caruaru (0,12; Tabela A.3), Floresta (0,14; Tabela A.4), Garanhuns (0,24; Tabela A.5) e Cabrobó (0,24; Tabela A.2). Tais valores, que superam levemente o IDW e, em alguns casos, o Random Forest, reforçam a capacidade da Krigagem de modelar com maior precisão os gradientes espaciais, especialmente em áreas com estações mais espaçadas.

O Random Forest também demonstrou desempenho consistente, especialmente na interpolação das variáveis de temperatura máxima e mínima. Na modelagem da temperatura máxima, por exemplo, os resultados foram excelentes, com RMSE de 0,04 em Surubim (Tabela A.11), 0,07 em Cabrobó (Tabela A.2), 0,16 em Ibimirim (Tabela 1) e 0,30 em Petrolina (Tabela A.8). O mesmo padrão de performance satisfatória foi identificado na temperatura mínima. Os dados completos para cada método e localidade estão apresentados de forma detalhada nas tabelas de A.1 a A.11, contidas no Apêndice A.

Para facilitar a visualização comparativa do desempenho dos métodos de interpolação, as Figuras 4, 5, 6, e 7 apresentam os valores de RMSE obtidos para as variáveis escolhidas para realização deste trabalho. Os gráficos permitem observar o comportamento dos três métodos (IDW, Krigagem e Random Forest) em diferentes

idades, destacando visualmente os padrões de erro e a consistência de cada abordagem.

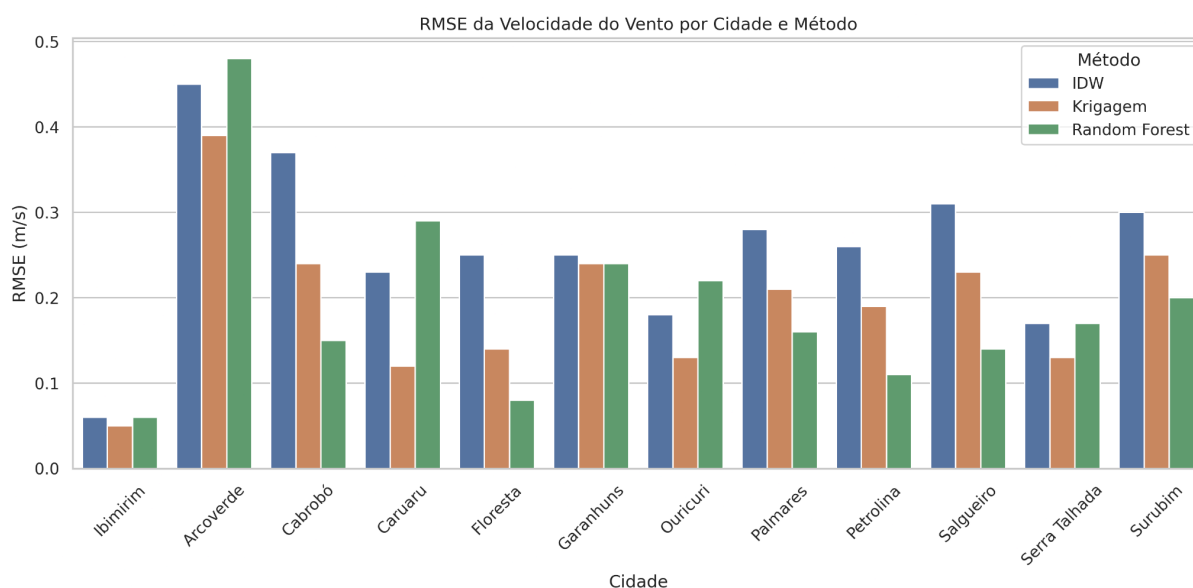


Figura 6 - Gráfico de comportamento dos métodos na variável velocidade dos ventos

Fonte: Próprio autor

A figura 6 revela que a Krigagem apresentou desempenho superior em diversas cidades, como Ibimirim e Caruaru, com RMSEs notavelmente baixos. Isso reforça sua capacidade de capturar os gradientes espaciais da variável de vento, especialmente em regiões com maior complexidade topográfica. O Random Forest e o IDW tiveram desempenhos semelhantes, mas levemente inferiores. Esse padrão evidencia a importância da modelagem da correlação espacial na estimativa da velocidade do vento, uma variável mais sensível a fatores geográficos como altitude e obstáculos naturais.

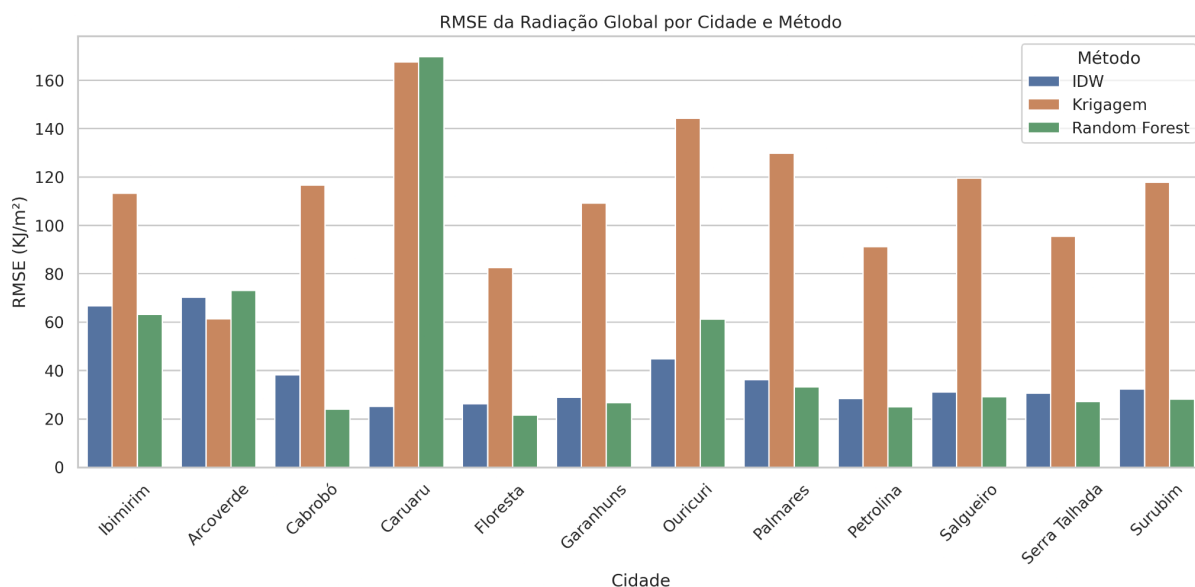


Figura 7 - Gráfico de comportamento dos métodos na variável radiação global

Fonte: Próprio autor

Na figura 7, podemos notar que o método IDW se destacou na interpolação da radiação global, apresentando os menores valores de RMSE na maioria das cidades. Isso pode ser atribuído à sua simplicidade e boa adaptação em áreas de relevo mais uniforme, onde a proximidade entre os pontos exerce maior influência. Já a Krigagem e o Random Forest oscilaram mais, com erros consideravelmente maiores em algumas cidades, o que evidencia a dificuldade desses métodos em capturar padrões de variabilidade associados à radiação, como nuvens ou microclimas locais. A variabilidade acentuada da radiação pode explicar por que mesmo técnicas mais robustas tiveram desempenho inferior aqui.

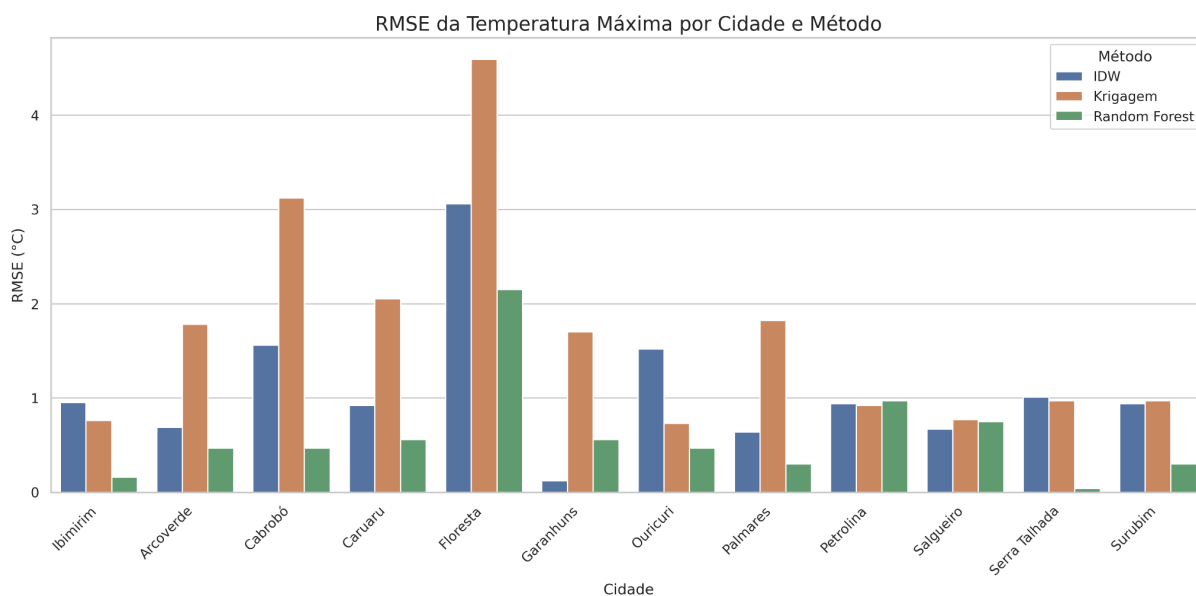


Figura 8 - Gráfico de comportamento dos métodos na variável temperatura máxima

Fonte: Próprio autor

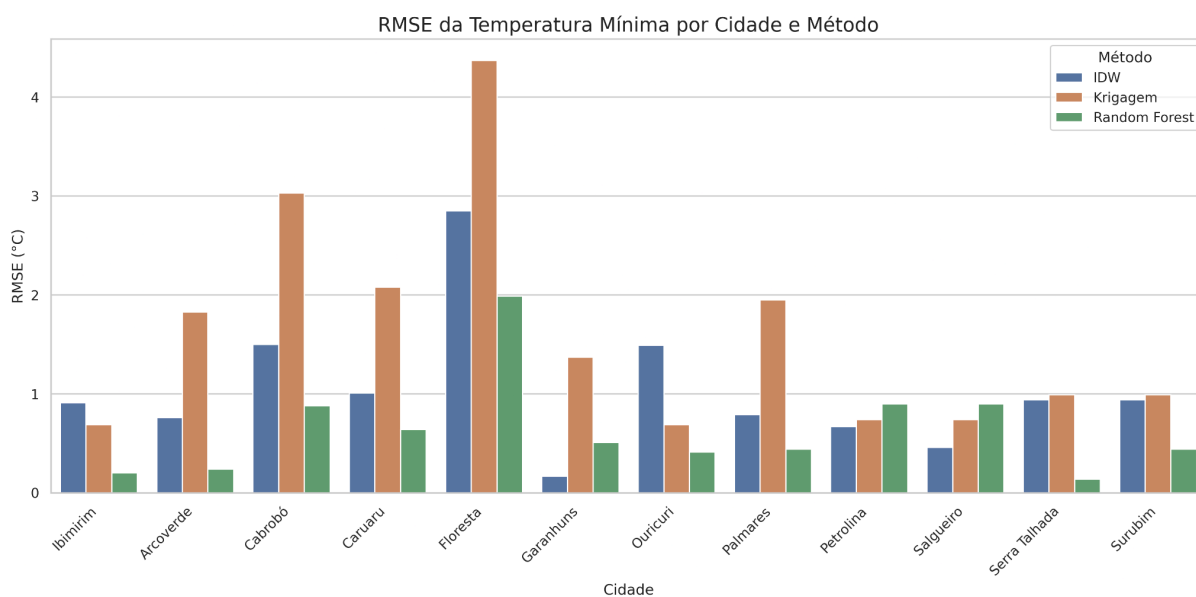


Figura 9 - Gráfico de comportamento dos métodos na variável temperatura mínima

Fonte: Próprio autor

Os gráficos de temperatura máxima e mínima, presentes nas figuras 8 e 9, evidenciam o excelente desempenho do modelo Random Forest, que obteve os

menores valores de RMSE de forma consistente em praticamente todas as cidades analisadas. Essa performance destaca sua capacidade de capturar padrões térmicos complexos, mesmo em regiões com variações sutis e não lineares. Por outro lado, a Krigagem demonstrou limitações notáveis, especialmente nas cidades de Floresta, Palmares, Cabrobó e Arcoverde, onde os erros foram significativamente mais elevados. Isso indica que, para variáveis térmicas, a abordagem geoestatística não foi eficaz em locais com baixa densidade de estações ou com forte influência de microclimas. O IDW, embora não tenha alcançado a precisão do Random Forest, mostrou-se mais estável e menos suscetível a grandes desvios do que a Krigagem, sobretudo na estimativa da temperatura mínima.

É importante ressaltar que a avaliação da precisão dos métodos foi realizada utilizando a validação cruzada leave-one-out (LOOCV). Nesse procedimento, cada localidade foi sucessivamente removida da base de dados, e os métodos foram aplicados para prever os valores correspondentes a essa localidade. Essa abordagem rigorosa permite testar a capacidade de generalização dos modelos para pontos não observados, simulando cenários reais em que se busca estimar valores para locais sem medições diretas.

| Variável | Método | Média Observada | Média Prevista | MSE | RMSE | MAE |
|-----------------------|---------------|-----------------|----------------|--------|------|------|
| TEMPERATURA MÁXIMA °C | IDW | 26.79 | 25.84 | 0.9025 | 0.95 | 0.95 |
| | KRIGAGEM | | 26.03 | 0.5776 | 0.76 | 0.76 |
| | RANDOM FOREST | | 26.95 | 0.0256 | 0.16 | 0.16 |
| TEMPERATURA MÍNIMA °C | IDW | 25.41 | 24.50 | 0.8281 | 0.91 | 0.91 |
| | KRIGAGEM | | 24.72 | 0.4761 | 0.69 | 0.69 |
| | RANDOM FOREST | | 25.61 | 0.04 | 0.2 | 0.2 |

| Variável | Método | Média Observada | Média Prevista | MSE | RMSE | MAE |
|-----------------------------------|---------------|-----------------|----------------|----------|--------|--------|
| RADIAÇÃO GLOBAL KJ/M ² | IDW | 1651.80 | 1585.05 | 4455.56 | 66.75 | 66.75 |
| | KRIGAGEM | | 1538.54 | 12827.82 | 113.26 | 113.26 |
| | RANDOM FOREST | | 1588.53 | 4003.1 | 63.27 | 63.27 |
| VELOCIDADE DO VENTO M/S | IDW | 2.89 | 2.95 | 0.0036 | 0.06 | 0.06 |
| | KRIGAGEM | | 2.94 | 0.0025 | 0.05 | 0.05 |
| | RANDOM FOREST | | 2.83 | 0.0036 | 0.06 | 0.06 |

Tabela 1 - Resultados das interpolações aplicadas à cidade de Ibimirim.

Fonte: Próprio autor

5.1 VALIDAÇÃO COM USINAS REAIS

A partir da Tabela 2, observa-se que, de modo geral, os valores estimados apresentaram coerência com as características climáticas regionais. Destaca-se, por exemplo, a cidade de Macaparana, que obteve estimativas de radiação global variando entre 1417,30 e 1579,07 KJ/m², além de uma velocidade do vento interpolada significativa, atingindo 3,26 m/s pelo modelo de Krigagem. De forma semelhante, Petrolina apresentou estimativas de radiação global entre 1436,66 e 1485,84 KJ/m², com destaque para a velocidade do vento, que alcançou 3,52 m/s no modelo IDW.

| Localização (Cidade) | Variável | IDW | Krigagem | Random Forest |
|----------------------|--------------------------------------|---------|----------|---------------|
| Flores | Temperatura Máxima (°C) | 26.70 | 26.95 | 26.98 |
| | Temperatura Mínima (°C) | 25.32 | 25.52 | 25.55 |
| | Radiação Global (KJ/m ²) | 1611.21 | 1600.15 | 1625.63 |
| | Velocidade do Vento (m/s) | 2.68 | 2.63 | 3.16 |

| | | | | |
|----------------------|---------------------------|---------|---------|---------|
| Macaparana | Temperatura Máxima (°C) | 25.02 | 25.37 | 25.20 |
| | Temperatura Mínima (°C) | 23.71 | 24.04 | 23.85 |
| | Radiação Global (KJ/m²) | 1523.75 | 1579.07 | 1417.30 |
| | Velocidade do Vento (m/s) | 3.20 | 3.26 | 2.95 |
| Paranatama | Temperatura Máxima (°C) | 22.39 | 22.35 | 22.98 |
| | Temperatura Mínima (°C) | 21.24 | 21.20 | 21.82 |
| | Radiação Global (KJ/m²) | 1460.04 | 1458.46 | 1441.24 |
| | Velocidade do Vento (m/s) | 2.76 | 2.73 | 2.87 |
| Petrolina | Temperatura Máxima (°C) | 27.76 | 27.66 | 27.70 |
| | Temperatura Mínima (°C) | 26.56 | 26.48 | 26.47 |
| | Radiação Global (KJ/m²) | 1436.66 | 1447.44 | 1485.84 |
| | Velocidade do Vento (m/s) | 3.52 | 3.47 | 3.34 |
| Poção | Temperatura Máxima (°C) | 24.38 | 23.60 | 23.20 |
| | Temperatura Mínima (°C) | 23.09 | 22.31 | 22.13 |
| | Radiação Global (KJ/m²) | 1522.81 | 1482.85 | 1500.08 |
| | Velocidade do Vento (m/s) | 03.03 | 2.97 | 2.90 |
| São José do Belmonte | Temperatura Máxima (°C) | 27.52 | 27.50 | 27.70 |
| | Temperatura Mínima (°C) | 26.09 | 26.07 | 26.30 |
| | Radiação Global (KJ/m²) | 1693.17 | 1677.03 | 1640.37 |
| | Velocidade do Vento (m/s) | 2.70 | 2.63 | 3.24 |
| Tacaratu | Temperatura Máxima (°C) | 26.79 | 25.78 | 27.11 |
| | Temperatura Mínima (°C) | 25.45 | 24.54 | 25.73 |
| | Radiação Global (KJ/m²) | 1613.95 | 1551.16 | 1542.49 |
| | Velocidade do Vento (m/s) | 2.94 | 2.99 | 2.75 |

Tabela 2 - Resultados das interpolações aplicadas ao conjunto de cidades selecionadas do estado de Pernambuco.

Fonte: Próprio autor

A partir dos dados da Tabela 2, a Figura 10 representa graficamente o IPPEH para cada localidade, mostrando a dispersão dos municípios de acordo com seus valores interpolados de velocidade do vento no eixo horizontal e irradiação solar no eixo vertical, o que confirma que eles se encontram no quadrante de bons índices para ambas as fontes energéticas.

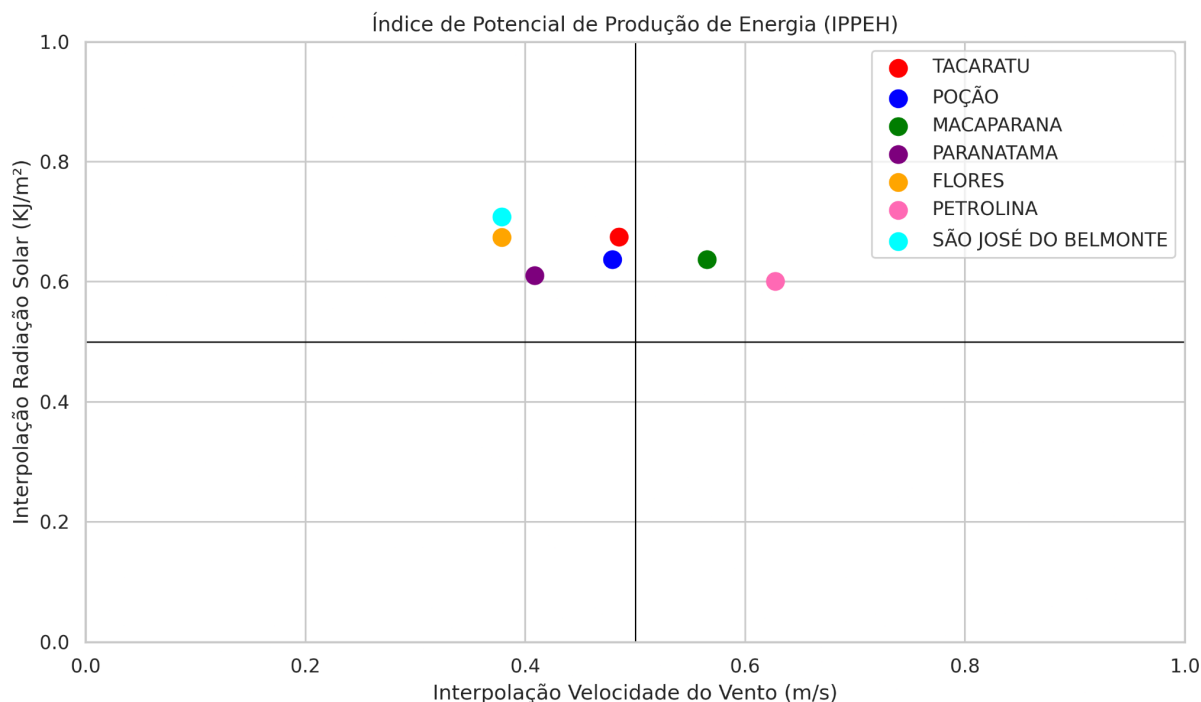


Figura 10 - Resultados do IPPEH para as cidades interpoladas.

Fonte: Próprio autor

Nesse gráfico (Figura 10), adota-se uma abordagem analítica inspirada na metodologia proposta por Ferreira et al. (2023), classificando as localidades conforme a combinação dos seus potenciais eólico e solar, organizando-as em quadrantes representativos: "Bom/Bom" (alto potencial para ambas as fontes), "Bom/Ruim" ou "Ruim/Bom" (alto potencial em uma fonte e baixo na outra) e "Ruim/Ruim" (baixo potencial em ambas). Isso permite uma análise integrada e comparativa das regiões, favorecendo a identificação de áreas mais adequadas para empreendimentos de geração híbrida.

Com base nessa classificação, os municípios de São José do Belmonte, Flores, Tacaratu, Paranatama e Poção foram posicionados no quadrante "Bom/Ruim", indicando um bom potencial para a geração de energia solar, mas um desempenho desfavorável para a energia eólica. Por outro lado, os municípios de Petrolina e Macaparana destacaram-se ao serem classificados no quadrante "Bom/Bom", evidenciando condições favoráveis para o aproveitamento simultâneo das duas fontes renováveis.

5.2 DISCUSSÃO DOS RESULTADOS

A partir dos resultados alcançados é possível observar que o método IDW superou os demais para interpolação dos dados em alguns cenários, possivelmente em função da distribuição relativamente equilibrada das estações meteorológicas e de certas áreas com relevo menos acidentado, uma vez que Pernambuco apresenta uma diversidade significativa de relevo, o que pode afetar a acurácia da interpolação espacial em determinadas regiões. Isso pode ser observado nos resultados para as cidades de Petrolina e Poção. Petrolina, localizada no Sertão do São Francisco, possui relevo predominantemente plano e aberto, inserido no Vale do São Francisco, o que favorece o método IDW por sua dependência da proximidade entre pontos e menor variabilidade altimétrica. Poção, situada no Agreste Pernambucano, está próxima ao Planalto da Borborema e com boa densidade de estações vizinhas.

Em áreas com topografia mais complexa, métodos como a Krigagem podem apresentar desempenho superior. Note que para estimativas de velocidade do vento a Krigagem apresentou melhores resultados para Macaparana e Tacaratu. Macaparana, na Zona da Mata Norte, está situada em meio à Serra do Mascarenhas, com vales e vegetação densa, o que requer técnicas que modelam a correlação espacial com maior precisão. Tacaratu, por sua vez, localiza-se no Sertão do Itaparica com fortes variações altimétricas em curtas distâncias.

O Random Forest foi superior em Flores, Paranatama e São José do Belmonte. Flores, no Sertão do Pajeú, está próxima à Serra Negra; Paranatama, no Agreste Central, situa-se junto à Serra do Pau Ferro; enquanto São José do Belmonte, no Sertão Central, é influenciado pelas altitudes da Serra do Catolé. Nessas localidades, múltiplos fatores como relevo, cobertura do solo e padrões de vento podem afetar as estimativas dos modelos. Embora essas características não estivessem explicitamente presentes nos dados, o Random Forest foi capaz de capturar variações não lineares com elevada acurácia.

Além disso, o uso dos dados do INMET foi positivo para Pernambuco, mas a expansão para outras regiões depende da disponibilidade e qualidade das estações locais. A ausência ou falhas nos dados atrapalham cálculos de médias e reduzem a precisão da interpolação. Foi possível observar que áreas com maior densidade de

estações próximas também apresentaram melhores resultados, indicando uma relação direta entre a quantidade de dados e a precisão das estimativas. Isso evidencia a importância de redes densas de monitoramento para otimizar as aplicações de interpolação espacial.

A validação com as usinas existentes buscou reforçar que é possível utilizar os modelos de interpolação espacial para identificar regiões de bom potencial energético. Apesar dos erros mais altos na interpolação da radiação global, os dados gerais indicam que a metodologia adotada consegue oferecer uma boa aproximação da realidade, podendo ser usada como ferramenta de apoio na escolha de locais para novos empreendimentos de energia renovável.

Por fim, a metodologia aplicada pode ser expandida como um serviço, a exemplo do serviço de estimativas de evapotranspiração apresentado por Júnior et al. (2019), que faz uso de técnicas de interpolação espacial para oferecer a produtores rurais estimativas de déficit hídrico para o plantio, auxiliando a tomada de decisão na irrigação. Isso permite que o setor público e privado avaliem o potencial energético da região para novos investimentos. Pode-se ainda expandir para potenciais clientes interessados em instalar um sistema de energia solar ou eólica, desde o investidor industrial ao agricultor que busca reduzir seus custos com eletricidade, irrigar suas plantações ou, eventualmente, suprir a demanda energética de veículos eletrificados.

5.3 LIMITAÇÕES E PERSPECTIVAS

Um dos principais desafios encontrados neste estudo foi a modelagem da variável radiação global, que apresentou resultados oscilantes. Em alguns casos, as estimativas se aproximaram satisfatoriamente dos valores observados, enquanto em outros, ocorreram erros elevados. Esse comportamento evidencia a complexidade inerente a essa variável, fortemente influenciada por fatores locais, como nebulosidade, altitude e características do relevo.

Outra limitação relevante refere-se à própria escala temporal e espacial dos dados utilizados. A escolha pelas médias anuais contribuiu para suavizar as variações sazonais que, por sua vez, são críticas para o dimensionamento e a operação de sistemas de geração de energia renovável, como os sistemas solar e eólico. Dessa forma, pode-se considerar a utilização de médias mensais ou diárias, permitindo uma caracterização mais detalhada dos padrões climáticos ao longo do ano. Além disso, a aplicação de técnicas de interpolação espaço-temporal pode fornecer resultados mais robustos e representativos, capturando a variabilidade tanto no espaço quanto no tempo.

É importante destacar também a significativa diversidade climática e geográfica presente dentro do estado de Pernambuco, que impacta diretamente nos resultados obtidos. O estado abrange desde o litoral úmido e quente, passando pela Zona da Mata e Agreste, até o Semi Árido do Sertão, caracterizado por clima mais seco e elevadas amplitudes térmicas. Essa heterogeneidade dificulta a interpolação espacial de variáveis climáticas, onde as condições atmosféricas podem mudar de maneira brusca em pequenas distâncias. Como estratégia, pode-se avaliar a redução do número de estações utilizadas, mediante o agrupamento de localidades com maior similaridade em termos de clima e relevo. Essa estratégia pode contribuir para a formação de zonas climáticas homogêneas, reduzindo a variabilidade interna do conjunto de dados e potencialmente aumentando a precisão das estimativas geradas pelos métodos de interpolação.

Por fim, uma perspectiva para superar algumas dessas limitações seria a integração de dados provenientes de estações automáticas localizadas em estados vizinhos. A ampliação do conjunto de dados, com informações adicionais das cidades de estados como Paraíba, Alagoas, Bahia e Ceará, pode contribuir significativamente para o aumento da densidade espacial da rede de monitoramento, melhorando a qualidade das interpolações, especialmente nas regiões fronteiriças de Pernambuco.

6 CONCLUSÃO

Este trabalho avaliou estratégias de interpolação espacial aplicadas à análise de variáveis climáticas relevantes para projetos de energia renovável, com foco nas fontes solar e eólica. Foram exploradas as técnicas IDW, Krigagem e Random Forest, aplicadas a dados de 12 estações meteorológicas do estado de Pernambuco.

Os resultados indicaram que o método IDW apresentou melhor desempenho na interpolação da radiação solar, sobretudo em regiões de relevo plano e com distribuição equilibrada de estações. A Krigagem demonstrou maior precisão na estimativa da velocidade do vento, especialmente em áreas com maior complexidade topográfica. O Random Forest obteve bons resultados para as variáveis de temperatura máxima e mínima, evidenciando capacidade de generalização mesmo em cenários com múltiplos fatores ambientais.

A utilização do índice IPPEH permitiu uma avaliação integrada do potencial energético híbrido das localidades, e a validação com usinas reais reforçou a aplicabilidade dos métodos propostos. A análise evidenciou que a interpolação espacial é uma ferramenta técnica e estratégica para a identificação de regiões com potencial para novos empreendimentos de energia renovável, sobretudo em áreas com baixa densidade de medições.

Como limitações, destacam-se a oscilação nas estimativas da radiação global e a adoção de médias anuais, que reduzem a sensibilidade às variações sazonais. Sugere-se, como trabalhos futuros, o uso de resoluções temporais mais finas, a incorporação de técnicas de interpolação espaço-temporal e a ampliação da base de dados com estações de estados vizinhos, de modo a aumentar a robustez e a aplicabilidade dos modelos desenvolvidos.

REFERÊNCIAS

- 1- FERREIRA, M. M. et al. A new index to evaluate renewable energy potential: a case study on solar, wind and hybrid generation in Northeast Brazil. *Energy Reports*, v. 9, p. 2051–2064, 2023.
- 2- PERIN, E. B. et al. Interpolação das variáveis climáticas. Temperatura do ar e precipitação: revisão dos métodos mais eficientes. *Revista Brasileira de Meteorologia*, v. 30, n. 2, p. 123–138, 2015.
- 3- SILVA JÚNIOR, J. C. et al. Random forest techniques for spatial interpolation of evapotranspiration data from Brazilian's Northeast. *Journal of Applied Remote Sensing*, v. 13, n. 3, p. 034518, 2019.
- 4- APAYDIN, H.; SONMEZ, F. K.; YILDIRIM, Y. E. Spatial interpolation techniques for climate data in the GAP region in Turkey. *Climate Research*, v. 28, p. 31–40, 2004.
- 5- SANTOS, A. S.; VALENTIM, A. M. Estimativa da velocidade do vento para estudo eólico offshore e onshore. *Revista Brasileira de Energia*, v. 29, n. 1, p. 45–56, 2023.
- 6- INSTITUTO NACIONAL DE METEOROLOGIA. Mapas Meteorológicos. Disponível em: <https://mapas.inmet.gov.br/#>. Acesso em: 6 jun. 2025.
- 7- GOVERNO DO ESTADO DE PERNAMBUCO. Atlas Eólico e Solar de Pernambuco. Disponível em: <http://www.atlaseolicosolar.pe.gov.br/>. Acesso em: 6 jun. 2025.
- 8- BEZERRA, D. de F. et al. Availability analysis of the Brazilian's national weather measurement system. *Revista Brasileira de Computação Aplicada*, v. 11, n. 3, p. 146–154, 2019. DOI: <https://doi.org/10.5335/rbca.v11i3.10026>.
- 9- JÚNIOR, W. M.; VALERIANO, T. T. B.; DE SOUZA ROLIM, G. EVAPO: a smartphone application to estimate potential evapotranspiration using cloud gridded meteorological data from NASA-POWER system. *Computers and Electronics in Agriculture*, v. 156, p. 187–192, 2019.

10- BREIMAN, Leo. Random forests. *Machine Learning*, v. 45, n. 1, p. 5–32, 2001. DOI: 10.1023/A:1010933404324.

11- BIDGOLI, A. M.; SADEGHIAN, M. S.; SAREMI, A.; HAJIKANDI, H. Analysis of annual and monthly precipitation trends based on long-term data (Case study: Zohreh–Jirahi Basin in Iran). *Journal of Earth System Science*, v. 132, n. 3, p. 144, 2023.

APÊNDICE A – INTERPOLAÇÃO EM DIFERENTES CIDADES

| Variável | Método | Média Observada | Média Prevista | MSE | RMSE | MAE |
|-----------------------------------|---------------|-----------------|----------------|----------|-------|-------|
| TEMPERATURA MÁXIMA °C | IDW | 24.46 | 25.15 | 0.4761 | 0.69 | 0.69 |
| | KRIGAGEM | | 26.24 | 3.1684 | 1.78 | 1.78 |
| | RANDOM FOREST | | 23.99 | 0.221 | 0.47 | 0.47 |
| TEMPERATURA MÍNIMA °C | IDW | 23.10 | 23.86 | 0.5776 | 0.76 | 0.76 |
| | KRIGAGEM | | 24.93 | 3.3489 | 1.83 | 1.83 |
| | RANDOM FOREST | | 22.86 | 0.0576 | 0.24 | 0.24 |
| RADIAÇÃO GLOBAL KJ/M ² | IDW | 1604.26 | 1533.92 | 4947.71 | 70.34 | 70.34 |
| | KRIGAGEM | | 1542.86 | 3769.96 | 61.4 | 61.4 |
| | RANDOM FOREST | | 1531.14 | 5346.534 | 73.12 | 73.12 |
| VELOCIDADE DO VENTO M/S | IDW | 3.29 | 2.84 | 0.2025 | 0.45 | 0.45 |
| | KRIGAGEM | | 2.90 | 0.1521 | 0.39 | 0.39 |
| | RANDOM FOREST | | 2.81 | 0.2304 | 0.48 | 0.48 |

Tabela A.1 - Resultados das Interpolações Aplicadas à Cidade de Arcoverde.

Fonte: Próprio autor

| Variável | Método | Média Observada | Média Prevista | MSE | RMSE | MAE |
|-----------------------------------|---------------|-----------------|----------------|----------|--------|--------|
| TEMPERATURA MÁXIMA °C | IDW | 27.88 | 27.64 | 0.0576 | 0.24 | 0.24 |
| | KRIGAGEM | | 25.93 | 3.8025 | 1.95 | 1.95 |
| | RANDOM FOREST | | 27.81 | 0.0049 | 0.07 | 0.07 |
| TEMPERATURA MÍNIMA °C | IDW | 26.65 | 26.21 | 0.1936 | 0.44 | 0.44 |
| | KRIGAGEM | | 24.60 | 4.2025 | 2.05 | 2.05 |
| | RANDOM FOREST | | 26.47 | 0.0324 | 0.18 | 0.18 |
| RADIAÇÃO GLOBAL KJ/M ² | IDW | 1655.02 | 1693.19 | 1456.95 | 38.17 | 38.17 |
| | KRIGAGEM | | 1538.24 | 13637.56 | 116.78 | 116.78 |
| | RANDOM FOREST | | 1679.07 | 578.4025 | 24.05 | 24.05 |
| VELOCIDADE DO VENTO M/S | IDW | 3.15 | 2.78 | 0.1369 | 0.37 | 0.37 |
| | KRIGAGEM | | 2.91 | 0.0576 | 0.24 | 0.24 |
| | RANDOM FOREST | | 3.00 | 0.0225 | 0.15 | 0.15 |

Tabela A.2 - Resultados das Interpolações Aplicadas à Cidade de Cabrobó.

Fonte: Próprio autor

| Variável | Método | Média Observada | Média Prevista | MSE | RMSE | MAE |
|-----------------------------------|---------------|-----------------|----------------|----------|--------|--------|
| TEMPERATURA MÁXIMA °C | IDW | 23.23 | 24.79 | 2.4336 | 1.56 | 1.56 |
| | KRIGAGEM | | 26.35 | 9.7344 | 3.12 | 3.12 |
| | RANDOM FOREST | | 23.70 | 0.2209 | 0.47 | 0.47 |
| TEMPERATURA MÍNIMA °C | IDW | 22 | 23.50 | 2.25 | 1.5 | 1.5 |
| | KRIGAGEM | | 25.03 | 9.18 | 3.03 | 3.03 |
| | RANDOM FOREST | | 22.88 | 0.7744 | 0.88 | 0.88 |
| RADIAÇÃO GLOBAL KJ/M ² | IDW | 1394.24 | 1419.38 | 632.02 | 25.14 | 25.14 |
| | KRIGAGEM | | 1561.95 | 28126.64 | 167.71 | 167.71 |
| | RANDOM FOREST | | 1563.94 | 28798.1 | 169.7 | 169.7 |
| VELOCIDADE DO VENTO M/S | IDW | 3.04 | 2.81 | 0.0529 | 0.23 | 0.23 |
| | KRIGAGEM | | 2.92 | 0.0144 | 0.12 | 0.12 |
| | RANDOM FOREST | | 2.75 | 0.0841 | 0.29 | 0.29 |

Tabela A.3 - Resultados das Interpolações Aplicadas à Cidade de Caruaru.

Fonte: Próprio autor

| Variável | Método | Média Observada | Média Prevista | MSE | RMSE | MAE |
|-----------------------------------|---------------|-----------------|----------------|--------|-------|-------|
| TEMPERATURA MÁXIMA °C | IDW | 27.97 | 27.05 | 0.8464 | 0.92 | 0.92 |
| | KRIGAGEM | | 25.92 | 4.2025 | 2.05 | 2.05 |
| | RANDOM FOREST | | 27.41 | 0.3136 | 0.56 | 0.56 |
| TEMPERATURA MÍNIMA °C | IDW | 26.68 | 25.67 | 1.02 | 1.01 | 1.01 |
| | KRIGAGEM | | 24.60 | 4.3264 | 2.08 | 2.08 |
| | RANDOM FOREST | | 26.04 | 0.41 | 0.64 | 0.64 |
| RADIAÇÃO GLOBAL KJ/M ² | IDW | 1623.68 | 1649.96 | 690.63 | 26.28 | 26.28 |
| | KRIGAGEM | | 1541.09 | 6821.1 | 82.59 | 82.59 |
| | RANDOM FOREST | | 1602.00 | 470.02 | 21.68 | 21.68 |
| VELOCIDADE DO VENTO M/S | IDW | 3.06 | 2.81 | 0.0625 | 0.25 | 0.25 |
| | KRIGAGEM | | 2.92 | 0.0196 | 0.14 | 0.14 |
| | RANDOM FOREST | | 2.98 | 0.0064 | 0.08 | 0.08 |

Tabela A.4 - Resultados das Interpolações Aplicadas à Cidade de Floresta.

Fonte: Próprio autor

| Variável | Método | Média Observada | Média Prevista | MSE | RMSE | MAE |
|-----------------------------------|---------------|-----------------|----------------|---------|--------|--------|
| TEMPERATURA MÁXIMA °C | IDW | 21.88 | 24.94 | 9.3636 | 3.06 | 3.06 |
| | KRIGAGEM | | 26.47 | 21.0681 | 4.59 | 4.59 |
| | RANDOM FOREST | | 24.03 | 4.6225 | 2.15 | 2.15 |
| TEMPERATURA MÍNIMA °C | IDW | 20.77 | 23.62 | 8.1225 | 2.85 | 2.85 |
| | KRIGAGEM | | 25.14 | 19.0969 | 4.37 | 4.37 |
| | RANDOM FOREST | | 22.76 | 3.9601 | 1.99 | 1.99 |
| RADIAÇÃO GLOBAL KJ/M ² | IDW | 1447.85 | 1476.85 | 841 | 29 | 29 |
| | KRIGAGEM | | 1557.08 | 11931.2 | 109.23 | 109.23 |
| | RANDOM FOREST | | 1421.13 | 713.96 | 26.72 | 26.72 |
| VELOCIDADE DO VENTO M/S | IDW | 2.71 | 2.96 | 0.0625 | 0.25 | 0.25 |
| | KRIGAGEM | | 2.95 | 0.0576 | 0.24 | 0.24 |
| | RANDOM FOREST | | 2.95 | 0.0576 | 0.24 | 0.24 |

Tabela A.5 - Resultados das Interpolações Aplicadas à Cidade de Garanhuns.

Fonte: Próprio autor

| Variável | Método | Média Observada | Média Prevista | MSE | RMSE | MAE |
|-----------------------------------|---------------|-----------------|----------------|--------|-------|-------|
| TEMPERATURA MÁXIMA °C | IDW | 27.35 | 27.47 | 0.0144 | 0.12 | 0.12 |
| | KRIGAGEM | | 29.05 | 2.89 | 1.7 | 1.7 |
| | RANDOM FOREST | | 27.91 | 0.3136 | 0.56 | 0.56 |
| TEMPERATURA MÍNIMA °C | IDW | 25.95 | 26.12 | 0.0289 | 0.17 | 0.17 |
| | KRIGAGEM | | 27.32 | 1.8769 | 1.37 | 1.37 |
| | RANDOM FOREST | | 26.46 | 0.2601 | 0.51 | 0.51 |
| RADIAÇÃO GLOBAL KJ/M ² | IDW | 1619.69 | 1649.93 | 914.45 | 30.24 | 30.24 |
| | KRIGAGEM | | 1541.45 | 6121.5 | 78.24 | 78.24 |
| | RANDOM FOREST | | 1696.17 | 5849.2 | 76.48 | 76.48 |
| VELOCIDADE DO VENTO M/S | IDW | 2.91 | 2.93 | 0.0004 | 0.02 | 0.02 |
| | KRIGAGEM | | 2.31 | 0.36 | 0.6 | 0.6 |
| | RANDOM FOREST | | 2.78 | 0.0169 | 0.13 | 0.13 |

Tabela A.6 - Resultados das Interpolações Aplicadas à Cidade de Ouricuri.

Fonte: Próprio autor

| Variável | Método | Média Observada | Média Prevista | MSE | RMSE | MAE |
|-----------------------------------|---------------|-----------------|----------------|----------|--------|--------|
| TEMPERATURA MÁXIMA °C | IDW | 25.42 | 23.90 | 2.31 | 1.52 | 1.52 |
| | KRIGAGEM | | 26.15 | 0.533 | 0.73 | 0.73 |
| | RANDOM FOREST | | 24.95 | 0.2209 | 0.47 | 0.47 |
| TEMPERATURA MÍNIMA °C | IDW | 24.14 | 22.65 | 2.22 | 1.49 | 1.49 |
| | KRIGAGEM | | 24.83 | 0.4761 | 0.69 | 0.69 |
| | RANDOM FOREST | | 23.73 | 0.1681 | 0.41 | 0.41 |
| RADIAÇÃO GLOBAL KJ/M ² | IDW | 1118.10 | 1481.15 | 131805.3 | 363.05 | 363.05 |
| | KRIGAGEM | | 1587.05 | 219914.1 | 468.95 | 468.95 |
| | RANDOM FOREST | | 1605.13 | 237198.2 | 487.03 | 487.03 |
| VELOCIDADE DO VENTO M/S | IDW | 2.10 | 3.08 | 0.9604 | 0.98 | 0.98 |
| | KRIGAGEM | | 3.01 | 0.8281 | 0.91 | 0.91 |
| | RANDOM FOREST | | 3.11 | 1.0201 | 1.01 | 1.01 |

Tabela A.7 - Resultados das Interpolações Aplicadas à Cidade de Palmares.

Fonte: Próprio autor

| Variável | Método | Média Observada | Média Prevista | MSE | RMSE | MAE |
|-----------------------------------|---------------|-----------------|----------------|----------|--------|--------|
| TEMPERATURA MÁXIMA °C | IDW | 27.76 | 27.12 | 0.41 | 0.64 | 0.64 |
| | KRIGAGEM | | 25.94 | 3.31 | 1.82 | 1.82 |
| | RANDOM FOREST | | 27.46 | 0.09 | 0.3 | 0.3 |
| TEMPERATURA MÍNIMA °C | IDW | 26.56 | 25.77 | 0.624 | 0.79 | 0.79 |
| | KRIGAGEM | | 24.61 | 3.8 | 1.95 | 1.95 |
| | RANDOM FOREST | | 26.12 | 0.19 | 0.44 | 0.44 |
| RADIAÇÃO GLOBAL KJ/M ² | IDW | 1436.66 | 1634.64 | 39196.08 | 197.98 | 197.98 |
| | KRIGAGEM | | 1558.09 | 14745.25 | 121.43 | 121.43 |
| | RANDOM FOREST | | 1594.47 | 24904 | 157.81 | 157.81 |
| VELOCIDADE DO VENTO M/S | IDW | 3.51 | 2.91 | 0.36 | 0.6 | 0.6 |
| | KRIGAGEM | | 2.88 | 0.3969 | 0.63 | 0.63 |
| | RANDOM FOREST | | 2.88 | 0.3969 | 0.63 | 0.63 |

Tabela A.8 - Resultados das Interpolações Aplicadas à Cidade de Petrolina.

Fonte: Próprio autor

| Variável | Método | Média Observada | Média Prevista | MSE | RMSE | MAE |
|-----------------------------------|---------------|-----------------|----------------|----------|--------|--------|
| TEMPERATURA MÁXIMA °C | IDW | 28.30 | 27.36 | 0.88 | 0.94 | 0.94 |
| | KRIGAGEM | | 27.38 | 0.84 | 0.92 | 0.92 |
| | RANDOM FOREST | | 27.33 | 0.94 | 0.97 | 0.97 |
| TEMPERATURA MÍNIMA °C | IDW | 26.73 | 26.06 | 0.45 | 0.67 | 0.67 |
| | KRIGAGEM | | 25.99 | 0.54 | 0.74 | 0.74 |
| | RANDOM FOREST | | 25.83 | 0.81 | 0.9 | 0.9 |
| RADIAÇÃO GLOBAL KJ/M ² | IDW | 1824.18 | 1619.35 | 41955.33 | 204.83 | 204.83 |
| | KRIGAGEM | | 1522.86 | 90793.74 | 301.32 | 301.32 |
| | RANDOM FOREST | | 1614.73 | 43869.30 | 209.45 | 209.45 |
| VELOCIDADE DO VENTO M/S | IDW | 2.54 | 3.00 | 0.2116 | 0.46 | 0.46 |
| | KRIGAGEM | | 2.97 | 0.1849 | 0.43 | 0.43 |
| | RANDOM FOREST | | 2.86 | 0.1024 | 0.32 | 0.32 |

Tabela A.9 - Resultados das Interpolações Aplicadas à Cidade de Salgueiro.

Fonte: Próprio autor

| Variável | Método | Média Observada | Média Prevista | MSE | RMSE | MAE |
|-----------------------------------|---------------|-----------------|----------------|---------|-------|-------|
| TEMPERATURA MÁXIMA °C | IDW | 26.80 | 27.08 | 0.079 | 0.28 | 0.28 |
| | KRIGAGEM | | 26.03 | 0.60 | 0.77 | 0.77 |
| | RANDOM FOREST | | 27.55 | 0.5625 | 0.75 | 0.75 |
| TEMPERATURA MÍNIMA °C | IDW | 25.39 | 25.72 | 0.11 | 0.33 | 0.33 |
| | KRIGAGEM | | 24.72 | 0.45 | 0.67 | 0.67 |
| | RANDOM FOREST | | 25.99 | 0.36 | 0.6 | 0.6 |
| RADIAÇÃO GLOBAL KJ/M ² | IDW | 1601.80 | 1647.89 | 2124.28 | 46.09 | 46.09 |
| | KRIGAGEM | | 1543.08 | 3448.03 | 58.72 | 58.72 |
| | RANDOM FOREST | | 1668.40 | 4435.56 | 66.6 | 66.6 |
| VELOCIDADE DO VENTO M/S | IDW | 2.46 | 2.93 | 0.2209 | 0.47 | 0.47 |
| | KRIGAGEM | | 2.98 | 0.2704 | 0.52 | 0.52 |
| | RANDOM FOREST | | 2.91 | 0.2025 | 0.45 | 0.45 |

Tabela A.10 - Resultados das Interpolações Aplicadas à Cidade de Serra Talhada.

Fonte: Próprio autor

| Variável | Método | Média Observada | Média Prevista | MSE | RMSE | MAE |
|-----------------------------------|---------------|-----------------|----------------|----------|--------|--------|
| TEMPERATURA MÁXIMA °C | IDW | 25.20 | 24.19 | 1.02 | 1.01 | 1.01 |
| | KRIGAGEM | | 26.17 | 0.94 | 0.97 | 0.97 |
| | RANDOM FOREST | | 25.24 | 0.0016 | 0.04 | 0.04 |
| TEMPERATURA MÍNIMA °C | IDW | 23.87 | 22.93 | 0.8836 | 0.94 | 0.94 |
| | KRIGAGEM | | 24.86 | 0.9801 | 0.99 | 0.99 |
| | RANDOM FOREST | | 23.73 | 0.0196 | 0.14 | 0.14 |
| RADIAÇÃO GLOBAL KJ/M ² | IDW | 1598.37 | 1389.62 | 43576.56 | 208.75 | 208.75 |
| | KRIGAGEM | | 1543.39 | 3022.8 | 54.98 | 54.98 |
| | RANDOM FOREST | | 1496.68 | 10340.85 | 101.69 | 101.69 |
| VELOCIDADE DO VENTO M/S | IDW | 3.47 | 2.81 | 0.4356 | 0.66 | 0.66 |
| | KRIGAGEM | | 2.88 | 0.3481 | 0.59 | 0.59 |
| | RANDOM FOREST | | 2.71 | 0.5776 | 0.76 | 0.76 |

Tabela A.11 - Resultados das Interpolações Aplicadas à Cidade de Surubim.

Fonte: Próprio autor