

Jackson Freitas de Amorim

**Avaliação de desempenho de técnicas de interpolação da
precipitação para o Nordeste brasileiro**

Recife

20 de Fevereiro de 2020



UNIVERSIDADE FEDERAL RURAL DE PERNAMBUCO
PRÓ-REITORIA DE PESQUISA E PÓS-GRADUAÇÃO
PROGRAMA DE PÓS-GRADUAÇÃO EM BIOMETRIA E ESTATÍSTICA APLICADA

**Avaliação de desempenho de técnicas de interpolação da precipitação para o
Nordeste brasileiro**

Dissertação julgada adequada para obtenção do título de Mestre em Biometria e Estatística Aplicada, defendida e aprovada por unanimidade em 20/02/2020 pela comissão examinadora.

Área de concentração: Biometria e Estatística Aplicada

Orientador: Ph.D. Rômulo Simões Cezar Menezes

Coorientador: Dr. Antonio Samuel Alves da Silva

Recife

20 de Fevereiro de 2020

Dados Internacionais de Catalogação na Publicação
Universidade Federal Rural de Pernambuco
Sistema Integrado de Bibliotecas
Gerada automaticamente, mediante os dados fornecidos pelo(a) autor(a)

- 524a Amorim, Jackson Freitas de
Avaliação de desempenho de técnicas de interpolação da precipitação para o Nordeste brasileiro / Jackson Freitas de Amorim. - 2020.
49 f. : il.
- Orientador: Ph D Romulo Simoes Cezar Menezes.
Coorientador: Dr Antonio Samuel Alves da Silva.
Inclui referências.
- Dissertação (Mestrado) - Universidade Federal Rural de Pernambuco, Programa de Pós-Graduação em Biometria e Estatística Aplicada, Recife, 2022.
1. Métodos de interpolação. 2. Cross validation. 3. Precipitação pluviométrica. 4. Geoestatística. 5. Séries temporais.
I. Menezes, Ph D Romulo Simoes Cezar, orient. II. Silva, Dr Antonio Samuel Alves da, coorient. III. Título

CDD 519.5

**UNIVERSIDADE FEDERAL RURAL DE PERNAMBUCO
PRÓ-REITORIA DE PESQUISA E PÓS-GRADUAÇÃO
PROGRAMA DE PÓS-GRADUAÇÃO EM BIOMETRIA E ESTATÍSTICA APLICADA**

**Avaliação de desempenho de técnicas de interpolação da precipitação para o
Nordeste brasileiro**

Jackson Freitas de Amorim

Dissertação julgada adequada para obtenção do título de Mestre em Biometria e Estatística Aplicada, defendida e aprovada por unanimidade em 20/02/2020 pela comissão examinadora.

Orientador:

**Ph.D. Rômulo Simões Cezar Menezes
Universidade Federal de Pernambuco -
UFPE**

Banca examinadora:

**Dr. Antonio Samuel Alves da Silva
Universidade Federal Rural de
Pernambuco – UFRPE**

**Dr. Josimar Gurgel Fernandes
Instituto Agrônômico de Pernambuco
– IPA**

*Dedico aos meus filhos Helena e Thales,
a minha esposa Beatriz que tem sido minha companheira e
as minhas tias Ana e Lia por tudo que sou.*

Agradecimentos

Ao Triuno Deus, criador do universo, ao qual sem Ele, nada do que existe viria a existir.

Aos professores Rômulo Simões Cezar Menezes e Antonio Samuel Alves da Silva, pelas orientações, paciência, inspirações e todo conhecimento adquirido nesta jornada.

À Universidade Federal Rural de Pernambuco e ao Programa de Pós-Graduação em Biometria e Estatística Aplicada e todo seu corpo docente e técnico.

Ao Departamento de Energia Nuclear da Universidade Federal de Pernambuco pela receptividade.

À Agência Nacional de Águas por disponibilizar os dados de precipitação utilizados neste trabalho.

Ao Instituto Agrônomo de Pernambuco nas pessoas dos pesquisadores Me. Manoel Américo Carvalho da Fonseca e ao Dr. José Nildo Tabosa por possibilitarem a realização do curso.

Aos colegas de trabalho na pessoa do pesquisador Domingos Eduardo, pelo apoio e incentivo.

Aos amigos de caminhada acadêmica, Luciano Pereira, Jucarlos Freitas, Laura Alves, Leika Tenório, Ademir, Mickaelle Almeida, Joelma Mayara, Denise Stéphane, Bárbara, Fábio Sandro, Edgo Jackson, Taciana Araújo, Henrique, Domingos, Eduardo, Kerolly, Gutemberg e todos que de forma direta e indireta contribuíram para o sucesso do curso.

A todos os familiares, em especial as minhas tias, que nunca mediram esforços para que meus sonhos fossem concretizados.

À minha esposa e filha, Beatriz Amorim e Helena Amorim, que pacientemente acompanhou, encorajou e orou em todos momentos desta jornada.

À Igreja Evangélica Congregacional em Macaxeira, em nome do Reverendo Ubirani Alves e todo o Conselho Eclesiástico, que sempre estiveram em orações ao meu favor.

*“Não há um único centímetro quadrado,
em todos os domínios de nossa existência,
sobre os quais Cristo, que é soberano sobre tudo,
não clame: É meu!”*
Abraham Kuyper (1898)

Resumo

Características como alta variabilidade temporal e espacial, além de baixa precipitação (abaixo de 800mm/ano) contribuem para que a região Nordeste brasileira enfrente problemas com disponibilidade de água. O regime pluviométrico da região afeta a economia, o meio ambiente e diversos setores da sociedade, constituindo um dos elementos meteorológicos de grande importância. Informações pontuais, limitadas e escassas levam à necessidade da utilização de métodos eficientes de interpolação de dados georreferenciados que possam ser empregados para preencher as lacunas de informações. Neste sentido, através das medidas de avaliação Erro Médio (EM), Erro Absoluto Médio (EAM), Erro Quadrático Médio (EQM), Raiz do Erro Quadrático Médio ($REQM$), Eficiência de Nash-Sutcliffe (NSE), D de Willmott (D) e Coeficiente de Determinação (R^2), este trabalho avalia o desempenho de modelos distintos para interpolação dos dados de precipitação no Nordeste brasileiro. Foram utilizadas séries de precipitações de 3513 estações meteorológicas cedidas pela Agência Nacional de Águas (ANA). Os métodos avaliados foram o *Inverse Distance Weighting* (para diferentes valores de $p = 0, 5; \dots; 5, 0$), o *Modified Shepard*, o *Nearest Neighbour*, o *Natural Neighbour*, a *Kriging* com diversos variogramas (*Linear*, *Power*, *Exponential*, *Gaussian*, *Quadratic*, *Rational Quadratic*, *Wave*, *Logarithmic*, *Spherical*, *Pentasppherical*, *Cubic*), o *Radial Basis Function* com as funções bases (*Multiquadric*, *Inverse Multiquadric*, *Multilog*, *Natural Cubic Spline*, *Thin Plate Spline*) e o *Trend Surface Analysis* com graus de superfície ($n = 1, \dots, 5$). Os resultados indicam que o melhor conjunto de medidas de avaliação foi obtido pelo *Modified Shepard*, com $EM = -0,03$, $EAM = 30,64$ e $REQM = 57,69$ e valores de eficiência $NSE = 0,76$, $D = 0,93$ e $R^2 = 0,76$. Por outro lado, dentre os sete métodos avaliados, o que apresentou o pior desempenho foi o *Radial Basis Function* utilizando a função base *Thin Plate Spline*, com destaque para os valores de $NSE = -7,79$, $D = 0,30$ e $R^2 = 0,07$. Os resultados mostraram que foi possível obter excelentes estimativas para precipitação mensal no Nordeste brasileiro a partir de método de interpolação em regiões onde possuam uma rede bem distribuída de estações meteorológicas. Ferramenta de acesso e visualização da precipitação para o Nordeste brasileiro pode ser desenvolvida e disponibilizada em site de instituições de pesquisas e ensino a fim de servir de suporte para trabalhos futuros bem como instrumento de auxílio para políticas públicas.

Palavras-chave: Métodos de interpolação. Cross validation. Precipitação pluviométrica. Geoestatística. Séries temporais.

Abstract

Characteristics such as high temporal and spatial variability, as well as low rainfall (below 800mm/year) contribute to the Northeast Brazilian region facing problems with water availability. The region's rainfall regime affects the economy, the environment and various sectors of society, constituting one of the most important meteorological elements. Punctual, limited and scarce information leads to the need to use efficient methods of interpolation of georeferenced data that can be used to fill the information gaps. In this sense, through the measurement measures Average Error (EM), Average Absolute Error (EAM), Mean Square Error (EQM), Root Mean Square Error ($REQM$), Nash-Sutecliffe Efficiency (NSE), Willmott D (D) and Coefficient of Determination (R^2), this work evaluates the performance of different models for interpolation of precipitation data in Northeast Brazil. Precipitation series from 3513 meteorological stations provided by the National Water Agency (ANA) were used. The evaluated methods were Inverse Distance Weighting (for different values of $p = 0.5; \dots; 5.0$), Modified Shepard, Nearest Neighbor, Natural Neighbor, Kriging with several variograms (Linear, Power, Exponential, Gaussian, Quadratic, Rational Quadratic, Wave, Logarithmic, Spherical, Pentaspherical, Cubic), the Radial Basis Function with the base functions (Multiquadric, Inverse Multiquadric, Multilog, Natural Cubic Spline, Thin Plate Spline) and Trend Surface Analysis with degrees of surface ($n = 1, \dots, 5$). The results indicate that the best set of evaluation measures was obtained by the Modified Shepard, with $EM = -0.03$, $EAM = 30.64$ and $REQM = 57.69$ and efficiency values $NSE = 0.76$, $D = 0,93$ and $R^2 = 0.76$. On the other hand, among the seven evaluated methods, the one with the worst performance was the Radial Basis Function using the Thin Plate Spline base function, with emphasis on the values of $NSE = -7.79$, $D = 0.30$ and $R^2 = 0.07$. The results showed that it was possible to obtain excellent estimates for monthly precipitation in Northeast Brazil using the interpolation method in regions where they have a well distributed network of meteorological stations. A tool for accessing and visualizing precipitation for the Brazilian Northeast can be developed and made available on the website of research and teaching institutions in order to serve as a support for future work as well as an aid tool for public policies.

Key-words: Interpolation methods. Cross validation. Rainfall. Geostatistics. Time series.

Lista de Figuras

Figura 1 – Diagrama de Voronoi e suas respectivas áreas A_i para 9 pontos amostrais	13
Figura 2 – Localização geográfica do Nordeste brasileiro.	15
Figura 3 – Delimitação do Semiárido brasileiro.	16
Figura 4 – Distribuição espacial das estações meteorológicas sobre o Nordeste brasileiro para o período de 1950 à 2017.	18
Figura 5 – Levantamento percentual de dados ausentes de precipitação para o Estados de Alagoas no período de 1988 à 2017.	19
Figura 6 – Número mensal de estações.	19
Figura 7 – Validação Cruzada: (A) seleção da estação para a ser “retirada” e (B) estação removida e dados interpolados considerando os demais.	20
Figura 8 – Erro Médio mensal de 1950 à 2017 para o <i>Modified Shepard</i>	27
Figura 9 – Erro Médio Absoluto mensal de 1950 à 2017 para o <i>Modified Shepard</i>	28
Figura 10 – Erro Quadrático Médio mensal de 1950 à 2017 para o <i>Modified Shepard</i>	28
Figura 11 – Raiz do Erro Quadrático Médio mensal de 1950 à 2017 para o <i>Modified Shepard</i>	28
Figura 12 – Coeficiente de Nash-Sutcliffe mensal de 1950 à 2017 para o <i>Modified Shepard</i>	29
Figura 13 – Coeficiente de determinação R^2 mensal de 1950 à 2017 para o <i>Modified Shepard</i>	29
Figura 14 – D de Willmott’s mensal de 1950 à 2017 para o <i>Modified Shepard</i>	29
Figura 15 – Boxplot das medidas de avaliação (a) <i>EM</i> e (b) <i>EMA</i> para o período de Jan-1950 a Dez-2017.	30
Figura 16 – Boxplot das medidas de avaliação (a) <i>EQM</i> e (b) <i>REQM</i> para o período de Jan-1950 a Dez-2017.	31
Figura 17 – Boxplot das medidas de avaliação (a) <i>NSE</i> e (b) <i>D</i> de Willmott’s para o período de Jan-1950 a Dez-2017.	31
Figura 18 – Boxplot da medida de avaliação R^2 para o período de Jan-1950 a Dez-2017. para o <i>Modified Shepard</i>	31

Lista de tabelas

Tabela 1 – Diferentes <i>Radial Basis Function</i> (<i>RBF's</i>)	10
Tabela 2 – Medidas de avaliação de desempenho para os 7 métodos de interpolação e suas variações.	21
Tabela 3 – Valores obtidos através das medidas de avaliação de desempenho para os sete métodos de interpolação e suas variações.	25

Lista de abreviaturas e siglas

<i>EM</i>	Erro Médio
<i>EAM</i>	Erro Absoluto Médio
<i>EQM</i>	Erro Quadrático Médio
<i>REQM</i>	Raiz do Erro Quadrático Médio
<i>NSE</i>	<i>Eficiência de Nash-Sutcliffe</i>
<i>D</i>	<i>D de Willmott</i>
<i>R²</i>	Coefficiente de Determinação
<i>IDW</i>	<i>Inverse Distance Weighting</i>
<i>TSA</i>	<i>Trend Surface Analysis</i>
<i>RBF</i>	<i>Radial Basis Function</i>
ANA	Agência Nacional de Águas
RHN	Rede Hidrometeorológica Nacional
ZCIT	Zona de Convergência Intertropical
SIG	Sistemas de Informação Geográfica
<i>OK</i>	<i>Ordinary Kriging</i>

Sumário

1	Introdução	1
2	Objetivos	3
2.1	Geral	3
2.2	Específicos	3
3	Revisão de Literatura	4
3.1	Precipitação	4
3.1.1	Mecanismos de formação das precipitações	5
3.1.2	Classificações das Precipitações	5
3.1.3	Precipitação na Região Nordeste do Brasil	7
3.2	Técnicas de interpolação espacial	8
3.2.1	Inverse Distance Weighting (<i>IDW</i>)	8
3.2.2	Trend Surface Analysis (<i>TSA</i>)	9
3.2.3	Radial Basis Function (<i>RBF</i>)	9
3.2.4	Kriging	11
3.2.5	Shepard Modified	11
3.2.6	Nearest Neighbour	12
3.2.7	Natural Neighbour	13
4	Metodologia	15
4.1	Caracterização da área	15
4.2	Dados	17
4.3	Avaliação dos métodos de interpolação	20
5	Resultados e discussão	24
6	Conclusão	32
	Referências Bibliográficas	33

1 Introdução

O Nordeste brasileiro possui extensão territorial de aproximadamente $1.550.000km^2$, dos quais 62% da sua área corresponde ao polígono das secas, uma região semiárida de $940.000km^2$, abrangendo nove Estados: Pernambuco – PE, Alagoas – AL, Paraíba – PB, Rio Grande do Norte – RN, Bahia - BA, Piauí – PI, Ceará – CE, Sergipe – SE e Maranhão – MA. Por sua baixa precipitação (abaixo de $800mm/ano$) e pela alta variabilidade temporal e espacial da precipitação, esta região enfrenta um grande problema de disponibilidade de água (SANTOS et al., 2017).

As precipitações são responsáveis pelo retorno da maior parte da água doce ao planeta, e executa um papel importante no ciclo hidrológico (MARUYAMA; KAWACHI; SINGH, 2005). Entretanto, o padrão de precipitação da região Nordeste do Brasil acarreta frequentes secas, que causam transtornos à população. A seca hidrológica está associada a redução do nível médio de água dos recursos hídricos superficiais e subsuperficiais como lagos, reservatórios, aquíferos e córregos. Vários métodos foram desenvolvidos para monitorar e quantificar a intensidade e os impactos da seca e da variável hidrológica precipitação (BRITO et al., 2018).

A região Nordeste foi afetada por uma seca iniciada em 2011, intensificada em 2012 que persistiu até 2017, período este considerado o mais grave nas últimas décadas, provocando um grande impacto na região, principalmente nos municípios inseridos no perímetro do semiárido, afetando aproximadamente 9 milhões de pessoas (BRITO et al., 2018). Além disso, é importante mencionar que a escassez hídrica também afeta negativamente a fauna e flora da região. Fica evidente, portanto, a importância do conhecimento do regime de chuvas da região, com o intuito de melhorar as estratégias de adaptação aos momentos de escassez hídrica.

As agências estaduais de monitoramento de águas e climas são responsáveis pela alimentação do banco de dados da Rede Hidrometeorológica Nacional RHN. Essas informações nem sempre são fornecidas com frequência, apresentando assim limitações referentes a disponibilidade de dados de precipitação, densidade da rede pluviométrica, bem como em relação ao período de observações. Desta forma, ferramentas que possam reunir, sistematizar, espacializar e disponibilizar esses dados de modo ágil, podem ser úteis para subsidiar estudos nas áreas ambientais, agrometeorologia, hidrologia, assim como, contribuir em questões de políticas públicas.

Com a finalidade de contornar os problemas de precipitações pontuais, bem como sua espacialização, encontram-se na literatura diversos estudos, tais como o de MIRANDA et al. (2017) que faz uma avaliação de dez métodos de interpolação sobre dados meteorológicos para a cidade de Petrolina-PE, em que o método *Natural Neighbor* (SIBSON, 1981) foi o que apresentou melhor resultado. Afim de verificar a aplicação de técnicas de interpolação para espacialização da chuva na região de Calha, Norte – PA (FILHO et al., 2019) utilizou a *Krigagem* (KRIGE, 1951) e o *Inverse Distance Weighted (IDW)* (SHEPARD, 1968), sendo o método da *Krigagem* o mais indicado para a análise da distribuição espacial e temporal da precipitação em comparação ao *IDW*.

O presente trabalho tem como objetivo avaliar o desempenho de técnicas de interpolação para precipitação no Nordeste brasileiro. Para isto serão avaliados 7 métodos de interpolação e suas variações: *IDW* ($p = 0, 5; \dots; 5, 0$) (SHEPARD, 1968), *Modified Shepard* (FRANKE; NIELSON, 1980), *Nearest Neighbour* (THIESSEN, 1911), *Natural Neighbour* (SIBSON, 1981), *Kriging* (*Linear, Power, Exponential, Gaussian, Quadratic, Rational Quadratic, Wave, Logarithmic, Spherical, Pentaspherical, Cubic*) (KRIGE, 1951), *Radial Basis Function* (*Multiquadric, Inverse Multiquadric, Multilog, Natural Cubic Spline, Thin Plate Spline*) (FORNBERG; LARSSON; WRIGHT, 2006), *Trend Surface Analysis* ($n=1; \dots; 5$) (WEBSTER; OLIVER et al., 1990).

Com essa abordagem, espera-se identificar o método de interpolação que apresente o melhor conjunto de medidas de avaliação (erro médio, erro absoluto médio, raiz do erro quadrático médio eficiência de Nash-Sutecliffe, D de Willmott e coeficiente de determinação) afim de preencher os dados faltantes, bem como em trabalhos futuros, sistematizar uma base de dados que possa ser disponibilizada em sites de instituições de pesquisa e ensino por meio de aplicativo de fácil acesso e visualização da precipitação para o Nordeste brasileiro.

2 Objetivos

2.1 Geral

Avaliar o desempenho de técnicas de interpolação para estimar a precipitação no Nordeste brasileiro.

2.2 Específicos

1. Sistematizar uma base de dados de precipitação para Região Nordeste do Brasil;
2. Avaliar o desempenho dos diferentes métodos de interpolação por meio de medidas de erro e eficiência de modelos.

3 Revisão de Literatura

3.1 Precipitação

A determinação prévia da variação dos elementos meteorológicos ao longo do ano, possibilita um melhor planejamento das mais diversas atividades. A precipitação é um elemento meteorológico de grande importância diretamente relacionada aos mais diversos setores da sociedade, de forma que o regime pluviométrico afeta a economia (NOGUEIRA; SANTANA, 2018), o meio ambiente e a sociedade como um todo (SILVA et al., 2017).

Precipitação é a água proveniente do vapor d'água da atmosfera, que chega a superfície terrestre sob a forma de: chuva, granizo, neve, orvalho, entre outros. Para as condições climáticas do Brasil, a chuva é a mais significativa em termos de volume. A umidade atmosférica é o elemento básico para a formação das precipitações.

Na agricultura, o conhecimento antecipado das condições edafoclimáticas são significativos para a obtenção de resultados satisfatórios, dado que esses fatores são determinantes para o sucesso nos cultivos. Segundo Silva et al. (2007), é importante caracterizar a precipitação local para o planejamento de atividades agrícolas, sendo imprescindível também no dimensionamento de reservatórios de água, na elaboração de projetos de proteção e conservação de solos, e em atividades de lazer esportivas.

A chuva é um dos elementos climáticos fundamentais na produção agrícola, devido ao seu caráter aleatório, aumentando consequentemente os riscos na programação das atividades do setor agrícola (GASPARIN et al., 2018). A precipitação pluviométrica tem grande importância na caracterização do clima de uma região (ALBUQUERQUE et al., 2019), interferindo nas alternâncias de rendimento das culturas e na construção civil (MEDEIROS et al., 2017).

O fenômeno da precipitação é o elemento alimentador da fase terrestre do ciclo hidrológico sendo fator importante para nos processos de escoamento superficial direto, infiltração, evaporação, transpiração, recarga de aquíferos, vazão básica dos rios e outros (VIEIRA et al., 2014).

Quando se faz um estudo de planejamento de longo prazo do uso de uma ou mais bacias hidrográficas, a precipitação é um dado básico, pois não sofre influências diretas de alterações antrópicas provocadas no meio (MONTAÑO; SOUZA, 2016). As alterações do

uso do solo, por exemplo, sobre as vazões escoadas, poderão ser avaliadas por modelos matemáticos que transformam as chuvas em vazões, e que consideram as variações de infiltração em função da área impermeável da bacia (OLIVEIRA, 2018). Nos projetos de drenagem, de construção de reservatórios de regularização (barragens) e outros, os dados de precipitação serão necessários para o dimensionamento das obras e conduzirão a resultados mais seguros quanto melhor for sua definição (NOLLI, 2017).

3.1.1 Mecanismos de formação das precipitações

A formação da precipitação segue o seguinte processo: o ar úmido das camadas baixas da atmosfera é aquecido por condução, torna-se mais leve que o ar das vizinhanças e sofre uma ascensão adiabática (resfriamento do ar por diminuição de pressão). A ascensão do ar provoca resfriamento, fazendo-o atingir o ponto de saturação. A partir desse nível, há condensação do vapor d'água na forma de gotículas mantidas em suspensão (como nuvens ou nevoeiros) até possuírem massa suficiente para vencer a resistência do ar ocasionando assim a precipitação (MIRANDA; OLIVEIRA; SILVA, 2010).

Os processos de crescimento das gotas mais importantes são os de coalescência e de difusão do vapor. Na coalescência, as pequenas gotas das nuvens aumentam seu tamanho através da colisão (contato com outras gotas), provocada pelo deslocamento de gotículas de água por meio de movimentos turbulentos do ar, forças elétricas e movimento Browniano. Quando as gotas atingem tamanho suficiente para vencer a resistência do ar, elas se deslocam em direção ao solo; nesse movimento de queda, as gotas menores serão alcançadas e incorporadas às maiores aumentando, portanto, seu tamanho (TAYOB et al., 2014).

Já no processo de difusão do vapor, o ar após o nível de condensação continua evoluindo, provocando difusão do vapor supersaturado e sua conseqüente condensação em torno das gotículas que aumentam de tamanho (STEINKE, 2016).

3.1.2 Classificações das Precipitações

Conforme o mecanismo fundamental pelo qual se produz a ascensão do ar úmido, as precipitações podem ser classificadas da seguinte forma: convectivas, orográficas, frontais ou ciclônicas (VILLELA; MATTOS, 1975).

Segundo Villela e Mattos (1975), a chuva convectiva está relacionada com o deslocamento de camadas de ar provocada pelo aquecimento do ar úmido localizado na vizinhança do solo. Através desse processo, pode-se criar camadas de ar que se mantêm

em equilíbrio instável. Perturbado o equilíbrio, forma-se uma brusca ascensão local do ar menos denso que atingirá seu nível de condensação com formação de nuvens, e muitas vezes, precipitações. Esse tipo de chuva é característica das regiões equatoriais, onde os ventos são fracos e os movimentos de ar são essencialmente verticais, podendo ocorrer nas regiões temperadas por ocasião do verão (tempestades violentas). São comumente chuvas de grande intensidade e pequena duração, restritas a pequenas áreas. Precipitações deste tipo podem provocar importantes inundações em pequenas bacias.

As chuvas orográficas ocorrem quando os ventos quentes e úmidos, geralmente no sentido do oceano para o continente, encontra uma barreira montanhosa, elevam-se e se resfriam adiabaticamente havendo condensação do vapor, formação de nuvens e ocorrência de chuvas (SANTOS et al., 2018). São chuvas de pequena intensidade e grande duração, que cobrem pequenas áreas. Quando os ventos conseguem ultrapassar a barreira montanhosa, do lado oposto é projetada uma sombra pluviométrica, dando lugar a áreas secas ou semiáridas causadas pelo ar seco, já que a umidade foi descarregada na encosta oposta.

Segundo COLLISCHONN e DORNELLES (2013), a formação de chuvas frontais ou ciclônicas são originárias da interação de massas de ar quente e frias. Nas regiões de convergência na atmosfera, o ar quente e úmido é violentamente impulsionado para cima, resultando no seu resfriamento e na condensação do vapor de água, de forma a produzir chuvas. São chuvas de grande duração, atingindo grandes áreas com intensidade média. Essas precipitações podem vir acompanhadas por ventos fortes com circulação ciclônica e podem produzir cheias em grandes bacias.

São observadas diferentes formas de precipitação na natureza, como: chuvisco, chuva, neve, saraiva, granizo, orvalho e geada. O chuvisco (neblina ou garoa), caracteriza-se como precipitação muito fina e de baixa intensidade; enquanto a chuva é ocorrência da precipitação na forma líquida, podendo também proceder de maneira congelada, constituída por gotas de água sobrefundida que congelam instantaneamente ao chocarem contra o solo, formando uma capa de gelo no mesmo. Já a neve é caracterizada como precipitação em forma de cristais de gelo, que durante a queda coalescem formando blocos de dimensões variáveis. A saraiva consiste na precipitação de pequenas pedras de gelo arredondadas com diâmetro de cerca de 5mm; enquanto o granizo ocorre na situação que as pedras de gelo, redondas ou de forma irregular, ultrapassam o diâmetro maior ou igual a 5mm. Além disso, o orvalho é consequentemente do resfriamento noturno, ocasionando redução da temperatura até ponto de orvalho, promovendo a condensação do vapor d'água presente no ar atmosférico. Sendo assim, em noites claras e calmas, os objetos expostos ao ar amanhecem cobertos por gotículas d'água. Ademais, geada é a deposição de cristais de gelo (fenômeno semelhante ao da formação de orvalho), ocorrendo quando a temperatura

é inferior a 0°C (YNOUE et al., 2017).

3.1.3 Precipitação na Região Nordeste do Brasil

A região Nordeste é caracterizada por diferentes regimes de chuvas. Esses regimes estão relacionados com vários sistemas atmosféricos de várias escalas. No Nordeste, os processos relacionados a precipitação estão inteiramente ligados a fenômenos de escala global, que essencialmente controla, intensificando ou inibindo, as chuvas sobre a região (MOLION; BERNARDO, 2002).

Segundo Nobre et al. (1986), a região é considerada anômala, pois em contraste a outras regiões nesta faixa latitudinal, apresenta um clima semiárido, devido aos valores relativamente baixos de precipitação sobre a maior parte da região. Isto ocorre por conta da circulação atmosférica sobre esta região, que é fortemente modulada e modificada pelos padrões termodinâmicos sobre as bacias dos oceanos Pacífico e Atlântico Tropicais influenciando o comportamento da precipitação pluvial sobre a região (FERREIRA; MELLO, 2005).

Para Silva, Oliveira e Oliveira (2017), os fatores determinantes para a distribuição dos elementos climáticos no Nordeste estão associados a sua posição geográfica, ao relevo e aos sistemas de pressão atuantes na região. Assim, os mecanismos dinâmicos que produzem chuvas nesta região podem ser classificados em: mecanismos de grande escala, esponsáveis por cerca de 30% a 80% da precipitação observada dependendo do local; e mecanismos de meso e microescalas, que contemplam os totais observados. Dentre os mecanismos de grande escala, destacam-se os sistemas frontais e a Zona de Convergência Intertropical (ZCIT).

Ainda segundo Nobre et al. (1986), é provável que o fator mais importante na determinação de quão abundante ou deficiente serão as chuvas de uma dada estação chuvosa na parte seca do Nordeste, é a posição latitudinal da ZCIT sobre o Atlântico Equatorial, mais especificamente sobre a porção ocidental próximo à costa da América do Sul. Em anos de seca, a ZCIT usualmente não cruza o equador na sua migração sazonal para o Sul, portanto não atinge o Nordeste.

Assim, ao norte da região Nordeste, a estação chuvosa geralmente está compreendida entre os meses de março a maio, ao sul e sudeste, normalmente as chuvas ocorrem no período de dezembro a fevereiro, e ao leste, a chuva ocorre geralmente entre maio a julho. Portanto, a variação temporal e espacial da precipitação anual sobre esta região é alta, pois a precipitação anual varia desde 300mm, na região mais seca, a 2.000mm, na costa leste (SILVA; OLIVEIRA, 2017; MENEGHETTI; FERREIRA, 2009).

3.2 Técnicas de interpolação espacial

Dados de precipitação são obtidos de estações meteorológicas pontuais. No entanto, estudos exigem dados espaciais e contínuos sobre a região de interesse, para subsidiar tomadas de decisões mais eficazes. Neste sentido, ferramentas e técnicas tem sido utilizadas para contornar o problema, sendo as técnicas de interpolação alinhada aos sistemas de informação geográfica (SIG), uma ferramenta poderosa para este fim.

A interpolação é um processo onde se busca estimar valores por meio de técnicas matemáticas. Sua forma geral pode ser expressa da seguinte maneira:

$$F(\mathbf{r}) = \sum_{k=1}^N Q(\mathbf{r}_k) f(\mathbf{r}_k) \quad (3.1)$$

em que $F(\mathbf{r})$ é o valor interpolado em \mathbf{r} , $f(\mathbf{r}_k)$ é o valor observado no ponto $\mathbf{r}_k \equiv (x_k, y_k)$, $Q(\mathbf{r}_k)$ são pesos atribuídos a cada uma das localidades e N é o número de pontos observados na vizinhança da área interpolada.

Dentre as diversas técnicas de interpolação utilizadas na literatura destacam-se: *Inverse Distance Weighting (IDW)*, *Trend Surface Analysis (TSA)*, *Radial Basis Function (RBF)*, *Kriging*, *Shepard Modified*, *Nearest Neighbour* e o *Natural Neighbour*.

3.2.1 Inverse Distance Weighting (IDW)

O *IDW* é um método determinístico proposto por [Shepard \(1968\)](#). O método é baseado na suposição de que o valor em um ponto desconhecido pode ser estimado pela média ponderada dos valores observados dentro de uma vizinhança de busca circular, cujo raio pode ser definido pelo intervalo de um número fixo de dados mais próximos. Isto indica que pontos amostrais mais próximos aos valores interpolados possuem maior peso do que os valores mais distantes, ou seja, cada ponto possui uma influência sobre a coordenada a ser interpolada diminuindo a medida que a distância aumenta. Sua expressão é descrita da seguinte forma:

$$F(\mathbf{r}) = \sum_{k=1}^N W(\mathbf{r}_k) f(\mathbf{r}_k) \quad (3.2)$$

em que o valor interpolado $F(\mathbf{r})$ é uma combinação linear dos pesos $W(\mathbf{r}_k)$, que são atribuídos a cada ponto coordenado, e dos valores $f(\mathbf{r}_k)$ em \mathbf{r} , sendo $W(\mathbf{r}_k)$ definido como:

$$W(\mathbf{r}_k) = \frac{d_k(\mathbf{r})^{-p}}{\sum_{i=1}^N d_k(\mathbf{r})^{-p}} \quad (3.3)$$

sendo $d_k(\mathbf{r}) \equiv \sqrt{(x - x_k)^2 + (y - y_k)^2}$ a distância euclidiana entre o local interpolado (\mathbf{r}) e cada um dos locais observados (\mathbf{r}_k); p o parâmetro de potência, em que determina a taxa de decaimento da influência das observações com a distância (CHENG et al., 2017).

3.2.2 Trend Surface Analysis (*TSA*)

O método *TSA* é considerado uma técnica de interpolação polinomial global, ajustando um polinômio à toda superfície. Também denominada regressão polinomial, no qual o método ajusta uma superfície lisa aos pontos originais georreferenciados. A superfície da *TSA* muda gradualmente de região para região sobre a área de interesse capturando a tendência global nos dados. Comumente, o ajuste de superfície polinomial global não é usado como um meio de interpolação, mas como uma forma de análise de tendências.

Segundo SILVA et al. (2019), o ajuste polinomial da superfície se dá por meio da regressão múltipla entre os valores originais georreferenciadas sendo descrito da seguinte forma:

$$F(\mathbf{r}) = \sum_{p=0}^m \sum_{\iota=0}^p \sum_{j=0}^{p-\iota} \psi_{\iota,j,p} x^\iota y^j + \xi \quad (3.4)$$

em que m indica a ordem do polinômio, $\psi_{\iota,j,p}$ são coeficientes, e ξ o erro. Os coeficientes são determinados utilizando o ajuste dos mínimos quadrados, minimizando a soma dos desvios quadrados da superfície da tendência, dada pela equação:

$$S = \sum_{k=1}^N [F(\mathbf{r}_k) - f(\mathbf{r}_k)]^2 \quad (3.5)$$

3.2.3 Radial Basis Function (*RBF*)

O *RBF* é um método de interpolação determinístico amplamente difundido na criação de superfícies mais suaves com a menor curvatura da superfície, todavia o método exige que a superfície passe através de cada ponto medido. Ademais, pode prever valores acima do valor máximo e abaixo do valor mínimo medido (LI et al., 2019). O *RBF* é um termo geral utilizado para designar várias funções de interpolação, tais como: *Inverse Multiquadric*, *Multiquadric*, *Multilog*, *Natural Cubic Spline* e *Thin Plate Spline* que foram utilizadas neste trabalho.

A versão mais simples deste método pode ser vista como uma função linear ponderada da distância do ponto observado ao ponto estimado, sendo expresso da seguinte

forma (SANTOS; SOUZA; FORTES, 2009):

$$\mathbf{F}(\mathbf{r}_j) = \sum_{j=1}^N \omega_j \phi(\mathbf{r}_j) \quad (3.6)$$

em que $\mathbf{F}(\mathbf{r}_j)$ é o valor interpolado, \mathbf{r}_j representa a distância euclidiana entre o valor a ser estimado e observado, $\phi(\mathbf{r}_j)$ é a função base radial selecionada e ω_j são coeficientes a serem determinados de forma que, $\mathbf{F}(\mathbf{r}_j) = t_j$, $j = 1, \dots, N$, isto é, como a solução para o sistema linear:

$$\begin{bmatrix} \beta \end{bmatrix} \begin{bmatrix} \omega_1 \\ \vdots \\ \omega_N \end{bmatrix} = \begin{bmatrix} t_1 \\ \vdots \\ t_N \end{bmatrix} \quad (3.7)$$

em que as entradas da matriz β são

$$\beta_{i,j} = \phi(\mathbf{r}_j), \quad i = 1, \dots, N; \quad j = 1, \dots, N \quad (3.8)$$

As funções de base radial $\phi(\mathbf{r}_j)$ mais utilizadas na literatura (Tabela 1), são expressas como:

Tabela 1 – Diferentes *Radial Basis Function (RBF's)*

Definição	Medida
<i>Inverse Multiquadric</i>	$\phi(\mathbf{r}_j) = \frac{1}{\sqrt{\mathbf{r}_j^2 + k^2}}$
<i>Multiquadric</i>	$\phi(\mathbf{r}_j) = \sqrt{\mathbf{r}_j^2 + k^2}$
<i>Multilog</i>	$\phi(\mathbf{r}_j) = \ln(k^2 + \mathbf{r}_j^2)$
<i>Natural Cubic Spline</i>	$\phi(\mathbf{r}_j) = (k^2 + \mathbf{r}_j^2)^{\frac{3}{2}}$
<i>Thin Plate Spline</i>	$\phi(\mathbf{r}_j) = (\mathbf{r}_j + k^2) \log(\mathbf{r}_j + k^2)$

em que o parâmetro de suavização k é definido:

$$k^2 = \frac{D}{25N} \quad (3.9)$$

sendo D o comprimento diagonal da grade e N o número de pontos contidos na grade.

3.2.4 Kriging

O método *Kriging* foi desenvolvido por Krige (1951) e leva em consideração a correlação espacial entre os pontos medidos. A técnica é frequentemente usada em geociência para estimar densidades minerais. Ela possui variações (tipos de família) tais como: *Simple Kriging*, *Ordinary Kriging (OK)* (adotado neste trabalho), *Universal Kriging*, *Indicator Kriging*, *Probability Kriging* e *Disjunctive Kriging*. Esses tipos de família foram utilizados em diferentes trabalhos, verificados em Li e Heap (2011), Wang et al. (2014), Jakob e Young (2016), Xiao et al. (2016), Javari (2017), Cheng et al. (2017).

O método de interpolação *OK* pressupõe que a média é constante, porém desconhecida, concentrando-se nos componentes espaciais e utilizando apenas os pontos de amostragem na vizinhança local para a estimativa do valor previsto. Para isto, o *OK* tem como base calcular um modelo empírico de semivariograma estimado para os dados, que pode ser definido pela seguinte expressão:

$$\hat{\theta}(x) = \frac{1}{2N(x)} \sum_{(i=1)}^{N(x)} [g(\mathbf{r}_i) - g(\mathbf{r}_i + x)]^2 \quad (3.10)$$

em que $N(x)$ corresponde ao número de pares amostrais $[g(\mathbf{r}_i) - g(\mathbf{r}_i + x)]$ separados por uma distância x (SILVA et al., 2019; PENG et al., 2017). Para o ajuste do modelo empírico $\hat{\theta}(x)$, são aplicados os seguintes variogramas teóricos: *Cubic*, *Linear*, *Power*, *Exponential*, *Gaussian*, *Quadratic*, *Rational Quadratic*, *Wave*, *Logarithmic*, *Spherical* e o *Pentaspherical*.

3.2.5 Shepard Modified

O método foi proposto por Franke e Nielson (1980) como uma modificação do método *IDW*, com objetivo de eliminar o efeito de mira. Esse processo é possível através da implementação do método de interpolação dos mínimos quadrados locais em cada ponto de observação \mathbf{r}_i (nó) para gerar as chamadas funções nodais $Q(\mathbf{r}_i)$, que são utilizadas para interpolar todos os outros pontos, levando em consideração apenas os nós vizinhos. Para os cálculos dos valores interpolados $F(\mathbf{r})$, aplica-se a seguinte equação:

$$F(\mathbf{r}) = \frac{\sum_{i=1}^N W_k(\mathbf{r})Q_k(\mathbf{r})}{\sum_{i=1}^N W_k(\mathbf{r})} \quad (3.11)$$

em que $W_k(\mathbf{r})$ é a função peso e definida pela seguinte expressão:

$$W_k(\mathbf{r}) = \left\{ \frac{[R_w - d_k(\mathbf{r})]_+}{R_w d_k(\mathbf{r})} \right\}^2 \quad (3.12)$$

com

$$[R_w - d_k(\mathbf{r})]_+ = \begin{cases} R_w - d_k(\mathbf{r}), & d_k(\mathbf{r}) < R_w \\ 0, & d_k(\mathbf{r}) \geq R_w \end{cases} \quad (3.13)$$

limitando a influência no valor interpolado para observações próximos dentro de um raio R_w . Logo, $d_k(\mathbf{r})$ corresponde a distância euclidiana entre o ponto interpolado (x, y) e o k -ésimo ponto amostral (x_k, y_k) ; $R_w > 0$ é o raio de influência sobre o ponto (x_k, y_k) . A função nodal $Q_k(\mathbf{r})$ é uma função quadrática determinada por:

$$Q_k(\mathbf{r}) = \bar{c}_{k1}(x - x_k)^2 + \bar{c}_{k2}(x - x_k)(y - y_k) + \bar{c}_{k3}(y - y_k)^2 + \bar{c}_{k4}(x - x_k) + \bar{c}_{k5}(y - y_k) + f(\mathbf{r}_k) \quad (3.14)$$

em que os coeficientes \bar{c}_{ki} , $i = 1, \dots, 5$ minimiza

$$\sum_{\substack{i=1 \\ j \neq k}}^N w_i(\mathbf{r}_k) [c_{k1}(x - x_k)^2 + \dots + c_{k5}(y - y_k) + f(\mathbf{r}_k) - f(\mathbf{r}_i)]^2 \quad (3.15)$$

com pesos

$$w_i(\mathbf{r}) = \left\{ \frac{[R_q - d_i(\mathbf{r})]_+}{R_q d_i(\mathbf{r})} \right\}^2 \quad (3.16)$$

limitando assim a regressão dos mínimos quadrados ao raio de influência R_q sobre o nó \mathbf{r}_k . $[R_q - d_i(\mathbf{r})]_+$ é definida de maneira análoga a [Equação 3.13](#). Para determinar os raios R_w e R_q , [Franke e Nielson \(1980\)](#) estabelecem a seguinte relação:

$$R_w = \frac{D}{2} \sqrt{\frac{N_w}{N}} \quad \text{e} \quad R_q = \frac{D}{2} \sqrt{\frac{N_q}{N}} \quad (3.17)$$

no qual $D = \max_k d_k(x, y)$ e N é o número total de amostras. Os números de pontos amostrais N_q e N_w foram considerados dentro dos círculos de raios R_q e R_k , com valores 19 e 13, respectivamente, utilizados por [SILVA et al. \(2019\)](#).

3.2.6 Nearest Neighbour

Proposto por [Thiessen \(1911\)](#), o método *Nearest Neighbour* também é conhecido como *Thiessen polygon*. O método é um dos mais antigos e simples, na qual cada interseção da grade recebe o valor do ponto mais próximo no conjunto de entradas. Equivalente ao método *Natural Neighbour*, o *Nearest Neighbour* tem como pressuposto que os pontos mais próximos aos pontos amostrais detêm as mesmas características. Podemos ver na [Figura 1](#) um exemplo de um diagrama de Voronoi ([VORONOI, 1908](#)) que fora obtido a partir de 9 pontos amostrais $\{(x_1, y_1), (x_2, y_2), \dots, (x_9, y_9)\}$ com suas respectivas áreas A_1, A_2, \dots, A_9 .

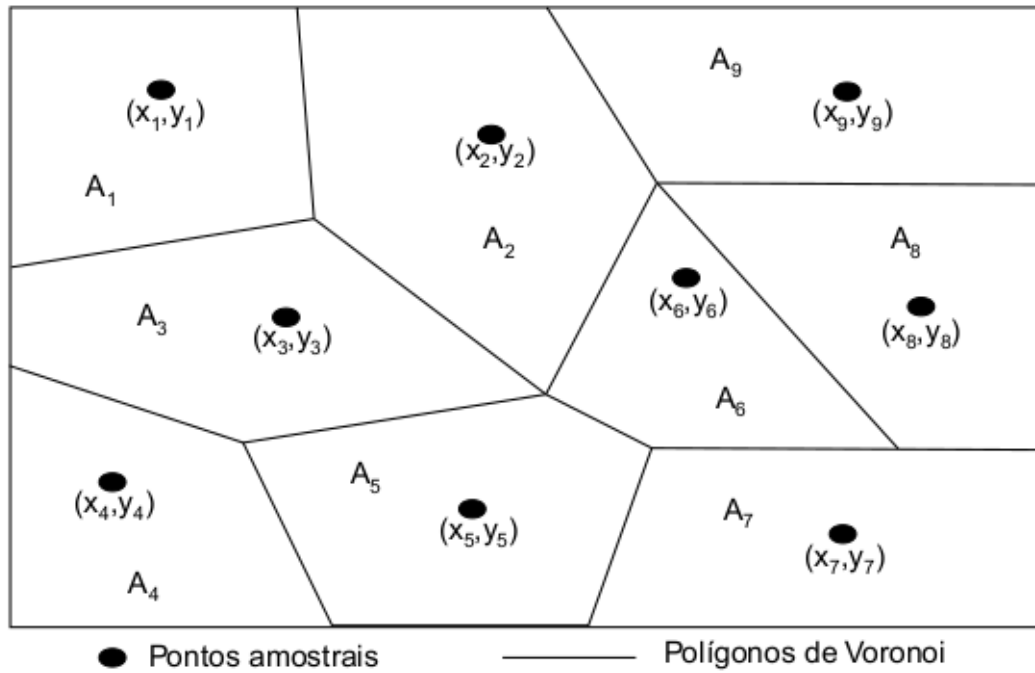


Figura 1 – Diagrama de Voronoi e suas respectivas áreas A_i para 9 pontos amostrais

No processo de interpolação as áreas geradas pelos polígonos são utilizadas como pesos. Dessa forma, podendo ser expressa de maneira:

$$F(x, y) = \sum_{k=1}^n W_k(x, y) F(x_k, y_k) \quad (3.18)$$

em que $F(x, y)$ é o valor interpolado em cada polígono, e $F(x_k, y_k)$ sendo o valor no k -ésimo ponto amostral. O peso é considerado na interpolação como função da área A_k associada em cada um dos polígonos e pode ser descrita desta forma:

$$W(r) = \frac{A_k}{\sum_{k=1}^n A_k} \quad (3.19)$$

Apesar da simplicidade do método, os valores dentro dos polígonos de Voronoi são constantes, mas sofrem abrupta mudança quando passa para outro polígono (SILVA et al., 2019).

3.2.7 Natural Neighbour

O método de interpolação *Natural Neighbour* foi proposto por Sibson (1981), que consiste em criar pesos para cada um dos pontos de entrada com base em sua suposta “área de influência”. Essas áreas são determinadas pela geração de polígonos de Voronoi

(VORONOI, 1908) em torno de cada ponto de entrada, partionando o plano em polígonos convexos de modo que cada polígono contenha exatamente um dos N pontos nos quais as as observações foram feitas. Em princípio, toda interseção de grade criada estaria em um desses polígonos, podendo receber o valor do ponto em torno do qual o polígono foi criado, resultando numa superfície que se assemelha a um degrau. Esse é o tipo de resultado obtido da interpolação do vizinho mais próximo.

Uma abordagem mais eficaz desse método tem como resultado final uma superfície lisa com descontinuidades nos pontos de entrada. Portanto, a primeira etapa do processo é criar uma triangulação Delaunay (AURENHAMMER, 1991; DELAUNAY et al., 1934) dos pontos de dados de entrada $j = 1, 2, \dots$ como um estágio preliminar na criação dos polígonos de Voronoi.

O segundo estágio é gerar um conjunto de polígonos de Voronoi para a região de estudo. Cada um dos pontos, j , no conjunto de dados de origem possui seu próprio polígono de Voronoi, de área A_j . Para determinar o valor estimado em um ponto de amostra P , o ponto P é temporariamente adicionado ao conjunto e os polígonos de Voronoi são recalculados. A adição do ponto P resulta em um novo polígono de Voronoi e na redefinição dos polígonos dos pontos geradores em sua vizinhança imediata. Logo, o polígono criado possui área A_p . De fato, esse novo ponto cede parte da área de influência de cada um dos pontos próximos, então a área total do polígono de Voronoi de P é dada por:

$$A_p = \sum_{i=1}^k A_{ip} \quad (3.20)$$

e, portanto, a proporção “cedida” de cada um dos pontos originais é:

$$\lambda_i = \frac{A_{ip}}{\sum_{i=1}^k A_{ip}} \quad (3.21)$$

Essas proporções são os pesos usados para calcular o valor estimado em P com base na equação de ponderação linear padrão:

$$Z_p = \sum_{i=1}^k \lambda_i Z_i \quad (3.22)$$

Se o ponto P coincidir com um dos pontos existentes, sua área de sobreposição com esse ponto seria 100%, portanto, seu peso seria 1. Se o polígono de Voronoi para P não se sobrepõe a uma região, o peso associado a essa região é 0. Uma das principais vantagens desse método de interpolação é que ele não requer tomada de decisão em relação ao número de pontos a serem usados, o raio, a direção de pesquisa e/ou quaisquer outros parâmetros (SILVA et al., 2019).

4 Metodologia

4.1 Caracterização da área

O Nordeste brasileiro, [Figura 2](#), está situado entre as coordenadas 1° e 18°30' latitude **S** e 34°30' e 40°20' longitude **W**, com extensão territorial de aproximadamente 1.550.000km². O seu relevo composto por planícies, vales, regiões da Serra da Borborema e Chapada Diamantina.

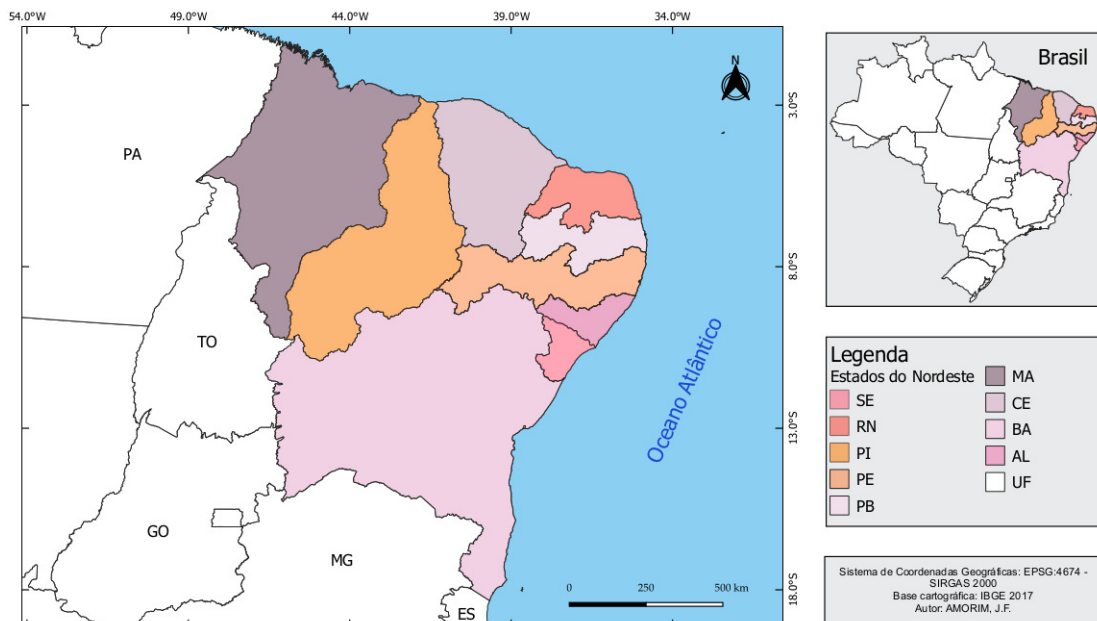


Figura 2 – Localização geográfica do Nordeste brasileiro.

A vegetação característica do Nordeste são remanescentes da Mata Atlântica, Mata dos cocais, Cerrado, Caatinga, Matas Ciliares e Vegetações Litorâneas. As principais bacias hidrográficas desta região são: Bacia do São Francisco, Bacia do Parnaíba, Bacia do Atlântico Nordeste Oriental, Bacia do Atlântico Nordeste Ocidental e Bacia do Atlântico Leste.

Segundo [Cavalcanti \(2016\)](#), a região apresenta três tipos de clima: litorâneo úmido, abrangendo todo o litoral nordestino; tropical, que compreende algumas áreas do Estado da Bahia, Ceará, Maranhão e Piauí; e tropical semiárido, envolvendo todo sertão nordestino. O mesmo autor propõe que as diferenças de clima no Nordeste Brasileiro são ocasionadas pela

atuação de diversos mecanismos físicos que interagem e são responsáveis pela distribuição de chuva nessa região. O semiárido brasileiro possui uma área correspondente ao polígono das secas de $1.128.697\text{km}^2$, que abrange dez Estados (Pernambuco - PE, Alagoas - AL, Paraíba - PB, Rio Grande do Norte - RN, Bahia - BA, Piauí - PI, Ceará - CE, Sergipe - SE e Maranhão - MA e parte Norte do Estado de Minas Gerais), que podem ser vistos na Figura 3.

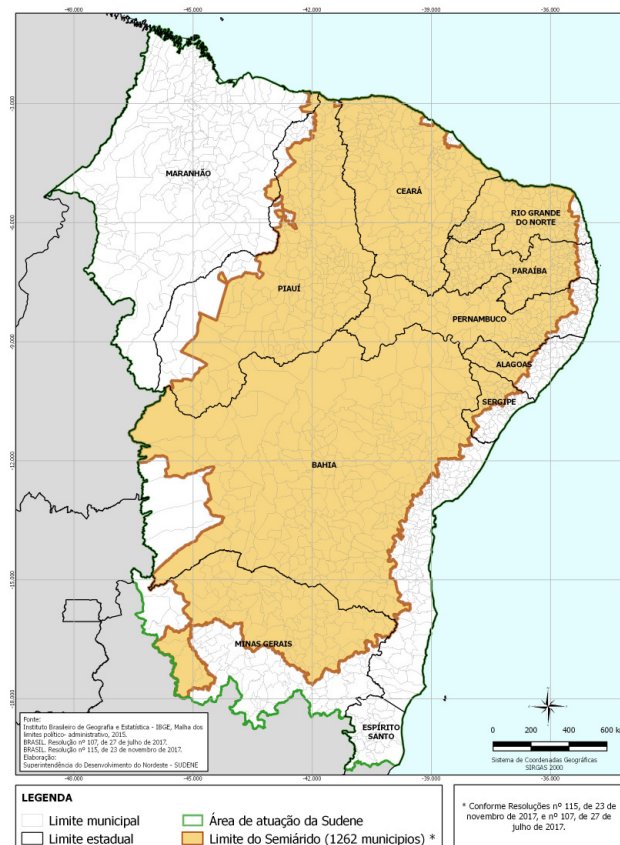


Figura 3 – Delimitação do Semiárido brasileiro.

Por sua baixa precipitação, inferior a 800mm/ano , a região Nordeste enfrenta um grande problema de disponibilidade de água (SANTOS et al., 2017) e uma alta variabilidade temporal e espacial da precipitação. Por conta do seu clima, com irregularidade na distribuição de chuva, a região semiárida é considerada crítica e ainda pouco explorada (OLIVEIRA, 2015). Sendo assim, o esgotamento de recursos hídricos é considerado um dos principais problemas sócio-ambientais do mundo.

4.2 Dados

Os dados utilizados neste trabalho são séries históricas de precipitação diária cedidas pela Agência Nacional de Água (ANA). A ANA é responsável pela coordenação da Rede Hidrometeorológica Nacional (RHN), um sistema que hoje abriga 4.641¹ pontos de monitoramento no país divididos em estações que monitoram parâmetros relacionados aos rios (1.874), como níveis, vazões, qualidade da água e transporte de sedimentos, e outros pontos que monitoram principalmente as chuvas (2.767). A RHN possui 8.207 estações (pluviométricas e fluviométricas). Para este estudo, foram utilizadas apenas estações pluviométricas, totalizando 3467 estações compreendendo o período de 1900 à 2018.

Foi estabelecido um período para análise, no qual a série foi truncada em 1950 até 2017, tendo em vista que este período apresenta o maior número de estações. Para o período em análise (Jan–1950 à Dez–2017), os dados foram coletados em 3.430 estações meteorológicas convencionais e automatizadas quem em seguida, foram calculada a precipitação mensal, obtendo assim uma nova série em escala mensal.

No intuito de obter uma melhor acurácia para a região de fronteira com os estados do Pará, Tocantins, Goiás e Minas Gerais, foram acrescentadas ao banco de dados 83 estações pluviométricas para minimizar o efeito de borda. Este efeito significa que a precisão do método de interpolação diminui na borda devido ao fato de que as estações meteorológicas nesta região não são cercadas por outras estações ([HARZHEIM; WARNECKE, 2010](#)).

As estações estão georreferenciadas, de modo que é possível visualizar a distribuição espacial no Nordeste brasileiro ([Figura 4](#)).

¹ Os números estão relacionados ao inventário da ANA em 03/04/2018, podendo ser alterado mediante incorporação de novas estações ou de sua extinção.

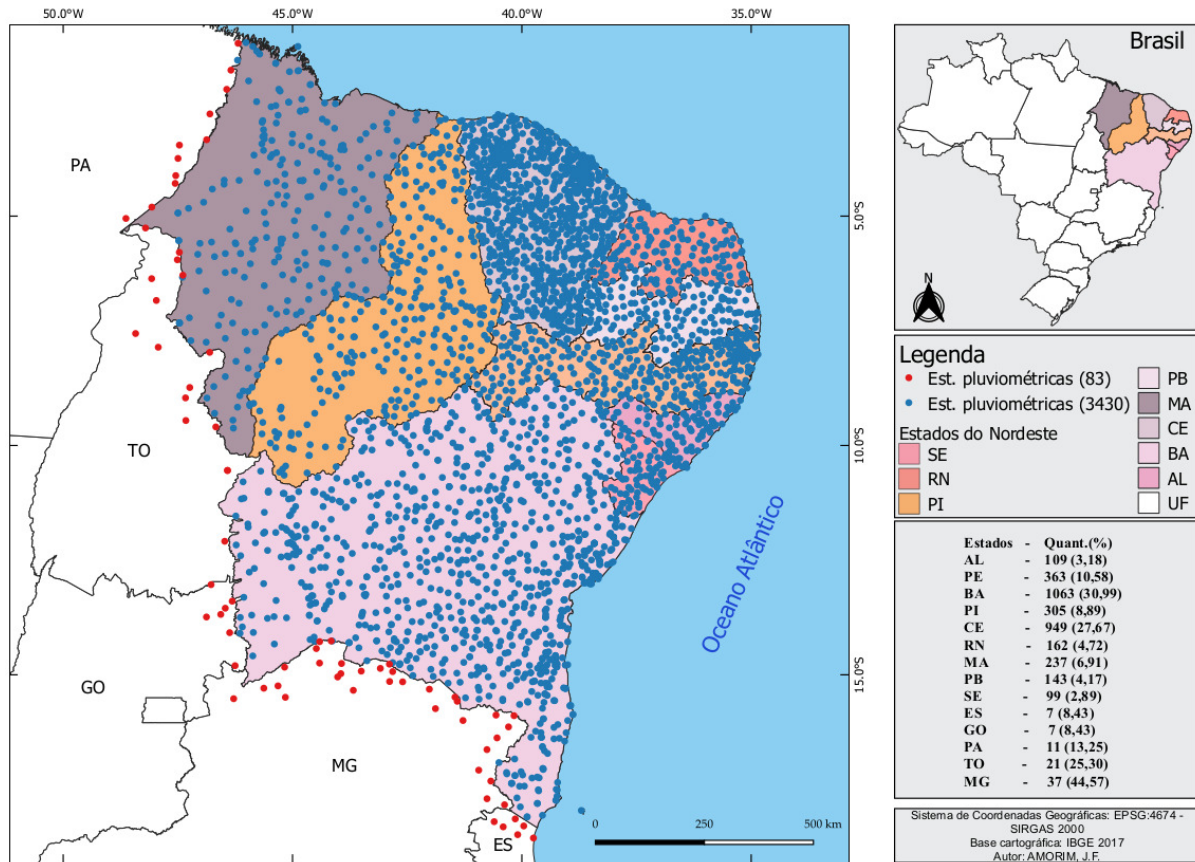


Figura 4 – Distribuição espacial das estações meteorológicas sobre o Nordeste brasileiro para o período de 1950 à 2017.

A série (escala mensal) possui dados faltantes. A Figura 5 para o estado de Alagoas ilustra o quanto de informações ausentes possuem algumas estações meteorológicas para o período de 1988 à 2017. A escala percentual varia de 0% de dados ausentes na cor azul a 15% na cor vermelho. As lacunas em branco indicam que não há informações sobre o período nas respectivas estações. Esta situação apresentada no Estado de Alagoas, estende-se para os demais estados da região Nordeste.

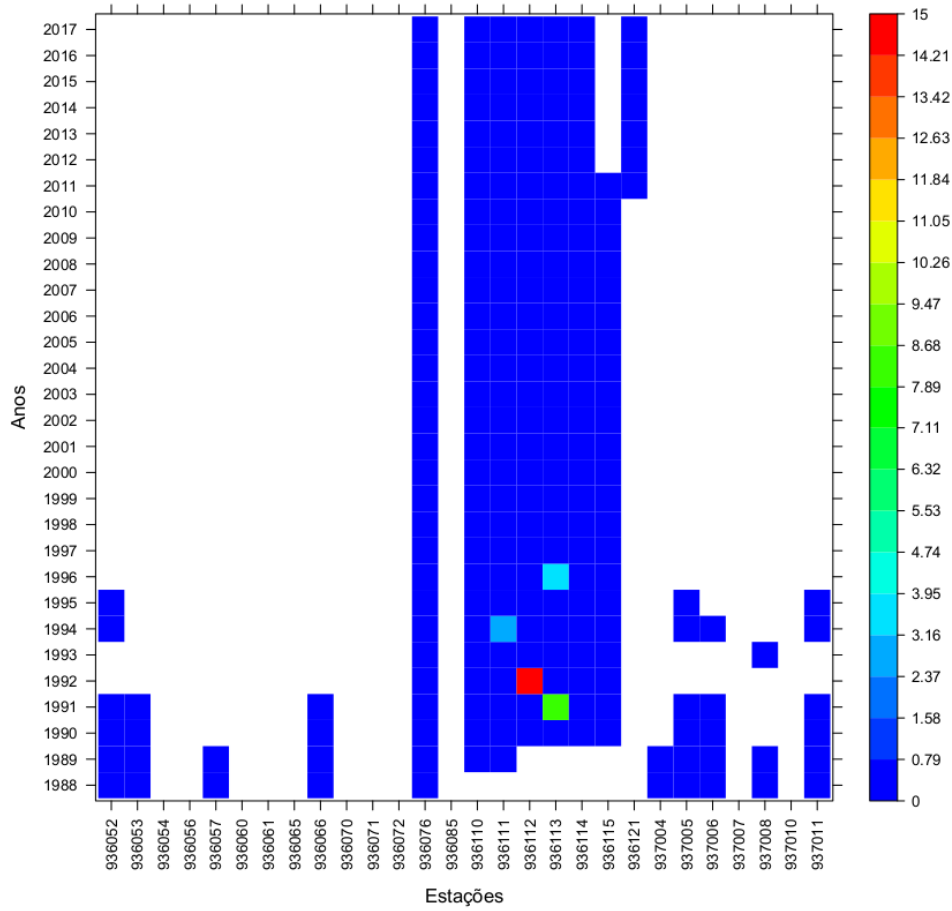


Figura 5 – Levantamento percentual de dados ausentes de precipitação para o Estados de Alagoas no período de 1988 à 2017.

Posteriormente, foi realizado levantamento com relação ao número de estações a cada mês/ano. Percebe-se que há uma variação no quantitativo de estações mês/ano (Figura 6). Esta amplitude de estações tem número mínimo na ordem de 487, e máximo na ordem de 2.081.

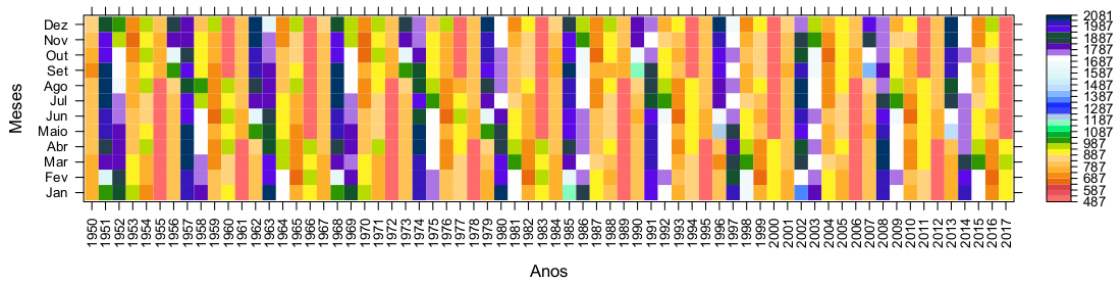


Figura 6 – Número mensal de estações.

4.3 Avaliação dos métodos de interpolação

Ao interpolar uma variável desconhecida, as superfícies geradas (mapas) pelos métodos são semelhantes. Por isso, é necessário estabelecer medidas que viabilizem comparar a precisão dos métodos utilizando o procedimento de validação cruzada – *cross validation* (WONG, 2015). O procedimento consiste na remoção de uma determinada estação meteorológica (Figura 7) do conjunto de dados, para realizar a estimativa neste ponto removido, empregando cada um dos métodos de interpolação em análise, que após a estimação, os dados que foram removidos serão realocados. Esse processo ocorre até que todas as coordenadas tenham valores observados e estimados (LUNDGREN; SILVA; FERREIRA, 2018).

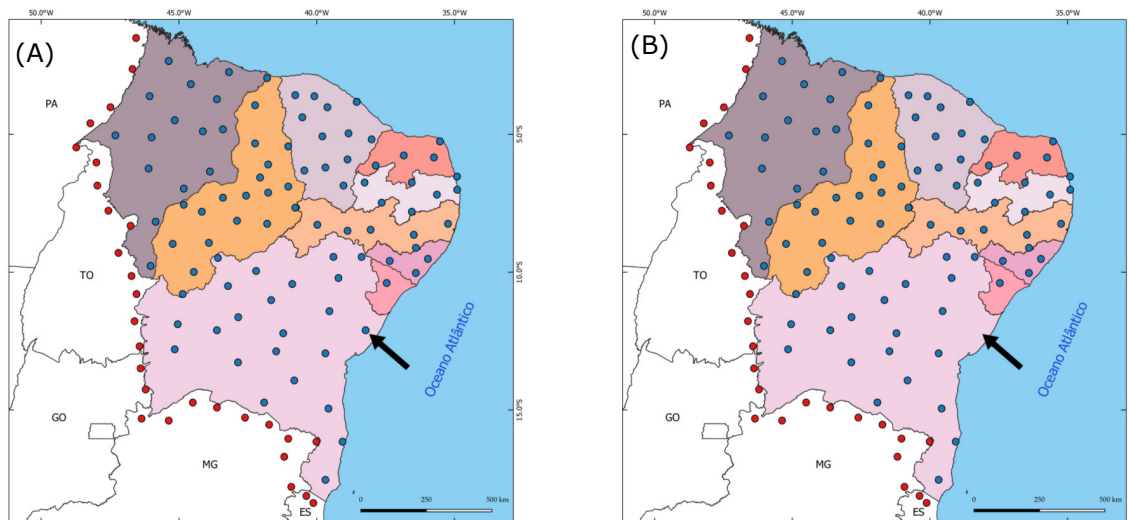


Figura 7 – Validação Cruzada: (A) seleção da estação para a ser “retirada” e (B) estação removida e dados interpolados considerando os demais.

A análise dos métodos de interpolação foram realizado em dois grupos. O primeiro denominado global, correspondendo a união de todos os dados observados de todas as estações meteorológicas em um único arquivo de Jan/1950 à Dez/2017. O segundo grupo designado temporal, constando os dados organizados por cada mês de cada ano.

Foram aplicadas sete medidas de avaliação de desempenho distintas (Tabela 2), utilizadas nos estudos de Xavier, King e Scanlon (2016), Farias, Francisco e Senna (2017), Cheng et al. (2017), SILVA et al. (2019); com intuito de analisar a eficácia das técnicas de interpolação, comparando dados estimados e observados. Para as análises estatísticas, foi utilizado o software SURFER® (GOLDEN SOFTWARE, LLC., 2015) aplicando o pacote hydroGOF.

Tabela 2 – Medidas de avaliação de desempenho para os 7 métodos de interpolação e suas variações.

Definição	Notação	Medida ²
Erro médio	EM	$\frac{1}{N} \sum_{k=1}^N [F(\mathbf{r}_k) - f(\mathbf{r}_k)]$
Erro absoluto médio	EAM	$\frac{1}{N} \sum_{k=1}^N F(\mathbf{r}_k) - f(\mathbf{r}_k) $
Erro quadrático médio	EQM	$\frac{1}{N} \sum_{k=1}^N [F(\mathbf{r}_k) - f(\mathbf{r}_k)]^2$
Raiz do erro quadrático médio	$REQM$	$\sqrt{\frac{1}{N} \sum_{k=1}^N [F(\mathbf{r}_k) - f(\mathbf{r}_k)]^2}$
Eficiência de Nash-Sutecliffe	NSE	$1 - \frac{\sum_{k=1}^N [F(\mathbf{r}_k) - f(\mathbf{r}_k)]^2}{\sum_{k=1}^N [f(\mathbf{r}_k) - \bar{f}(\mathbf{r}_k)]^2}$
D de Willmott	D	$1 - \frac{\sum_{k=1}^N [F(\mathbf{r}_k) - f(\mathbf{r}_k)]^2}{\sum_{k=1}^N F(\mathbf{r}_k) - \bar{f}(\mathbf{r}_k) + f(\mathbf{r}_k) - \bar{f}(\mathbf{r}_k) }$
Coefficiente de determinação	R^2	$\left\{ \frac{\sum_{k=1}^N [F(\mathbf{r}_k) - \bar{f}(\mathbf{r}_k)][f(\mathbf{r}_k) - \bar{f}(\mathbf{r}_k)]}{\sqrt{\sum_{k=1}^N [F(\mathbf{r}_k) - \bar{f}(\mathbf{r}_k)]^2} \sqrt{\sum_{k=1}^N [f(\mathbf{r}_k) - \bar{f}(\mathbf{r}_k)]^2}} \right\}^2$

² N : número de observações; $f(\mathbf{r}_k)$: valor observado no ponto $\mathbf{r}_k \equiv (x_k, y_k)$; $F(\mathbf{r}_k)$: valor estimado no ponto r_k e $\bar{f}(\mathbf{r}_k)$: média dos valores observados.

Segundo Lima et al. (2017), o Erro Médio (EM) ou viés, indica o erro sistemático, subestimação ou superestimação ($EM = 0$ ideal), da técnica em relação aos dados observados. Ainda segundo (HALLAK; FILHO, 2011), o EM não apresenta informações dos erros individuais, não sendo recomendado como única medida de avaliação. Além disso, afetado pelo somatório de mesma magnitude dos erros positivos e negativos.

Para contornar o problema apresentado no EM , utiliza-se o Erro Médio Absoluto (EMA). O EMA possui a vantagem de ser menos afetado por pontos com valores especificamente extremos, conhecidos por “pontos singulares” ou *outliers*. Por isso, é considerado uma medida precisa e robusta para determinar a eficácia da técnica de interpolação em reproduzir a realidade (SILVA et al., 2019).

Outra medida de avaliação bastante utilizada na verificação da acurácia dos métodos de interpolação é o Erro Quadrático Médio (EQM). Porém, é sensível aos grandes erros por aumentar as diferenças individuais ao quadrado.

A Raiz do Erro Quadrático Médio ($REQM$) determina o quadrado do desvio entre os valores observados e os estimados e, portanto, é mais sensível a erros maiores

(XAVIER; KING; SCANLON, 2016). O EM , EMA e o $REQM$ medem a precisão dos dados estimados, de forma que, quando os dados estimados e observados são semelhantes (EM , EMA e o $REQM$ estão próximos de zero), indicam uma interpolação mais precisa. Ao extrair a raiz quadrada, a $REQM$ apresenta valores nas mesmas dimensões dos dados observados, sendo comumente utilizada para expressar a acurácia dos modelos de interpolação analisados.

Associada ao EM , EMA , EQM e o $REQM$, medidas como Eficiência de Nash-Sutcliffe (NSE), D de Willmott (D) e o Coeficiente de Determinação (R^2) também são amplamente difundidos para quantificar a incerteza da predição afim de avaliar o desempenho de técnicas de interpolação. Proposto por Nash e Sutcliffe (1970), a NSE pode assumir valores no intervalo $(-\infty, 1]$. Uma eficiência menor que 0 ($NSE < 0$) indica que a média é um melhor preditor do que o modelo, enquanto que, uma eficiência igual a 0 ($NSE = 0$) indica que os dados estimados pelo modelo são tão precisas quanto a média dos dados observados. Para uma eficiência ser considerada ótima, o seu valor necessita estar o mais próximo do valor 1.

Proposto por Willmott (1981), o índice de concordância ou D de Willmott (D), determina a precisão do método de interpolação aplicado e avalia o grau de afastamento entre os dados estimados em relação aos observados. O D pode assumir valores no intervalo $[0, 1]$, em que $D = 0$ indica nenhuma concordância, enquanto que $D = 1$ indica concordância perfeita (SILVA et al., 2019).

Ademais, a qualidade do ajuste do modelo de interpolação pode ser avaliado por meio do Coeficiente de Determinação (R^2), que consiste no quadrado do coeficiente de correlação de Pearson. Basicamente, o R^2 indica o quanto o modelo foi capaz de explicar os dados coletados, ou seja, é uma medida da proporção da variabilidade em uma variável que é explicada pela variabilidade da outra. Os valores do R^2 estão no intervalo $[0, 1]$ e quanto maior é o seu valor, mais explicativo o modelo de interpolação utilizado (QUININO; REIS; BESSEGATO, 1991).

Com base nos dados originalmente obtidos pela ANA (escala diária) foi obtida uma nova série em escala mensal, considerando cada uma das estações meteorológicas pertencente a um mês específico e seu respectivo ano, totalizando 816 meses analisados, ou seja, 12 meses (Janeiro a Dezembro) em 68 anos (1950 a 2017). A técnica *cross validation* foi aplicada a cada mês para analisar as técnicas de interpolação descritas na seção 4.3, com objetivo de comparar e identificar qual dentre os 34 métodos de interpolação apresenta o melhor conjunto de medidas de avaliação (Tabela 2) para todos os meses analisados.

As interpolações foram executadas no programa SURFER® (GOLDEN SOFTWARE, LLC., 2015), considerando cada um dos 7 métodos e suas diferentes variações (34) totali-

zando 27744 interpolações, concernente a escolha de parâmetros, variogramas, função base e grau da superfície.

Para o processo de interpolação pelo método *Inverse Distance Weighting* neste trabalho foram utilizados 10 valores diferentes de p , ou seja, valores de ponderação com intervalos de $p = 0, 5; 1, 0; 1, 5; \dots; 5, 0$. Na técnica de interpolação pelo método *Modified Shepard*, os números de pontos amostrais N_Q e N_W foram considerados dentro dos círculos de raios R_Q e R_W , nos quais seus valores foram respectivamente 19 e 13, metodologia utilizada por [SILVA et al. \(2019\)](#).

Na interpolação por meio da técnica *Radial Basis Function*, as funções *Inverse Multiquadric*, *Multiquadric*, *Multilog*, *Natural Cubic Spline* e *Thin Plate Spline* foram utilizadas com valores de $R^2 = D^2/(25N)$ nos 816 meses, em que D é o comprimento da diagonal da grade dos dados a serem interpolados e N sendo o número total das estações pluviométricas ([SILVA et al., 2019](#)).

No método de interpolação *Ordinary Kriging* os seguintes variogramas foram considerados: *Cubic*, *Linear*, *Power*, *Exponencial*, *Gaussian*, *Quadratic*, *Rational Quadratic*, *Wave*, *Logarithmic*, *Spherical* e *Pentospherical*. Os variogramas foram ajustados previamente para os 816 meses utilizando o método dos mínimos quadrados. Para o método *Trend Surface Analysis*, superfícies de graus ($n = 1, 2, 3, 4$ e 5) foram utilizadas para interpolações para cada um dos 816 meses.

5 Resultados e discussão

Como mencionado, as técnicas de interpolação produzem como resultados imagens que por sua vez são semelhantes, tornando uma tarefa extremamente difícil comparar e identificar qual método obtem melhores estimativas. Com a finalidade de identificar o método que apresentou o melhor desempenho, primeiramente as técnicas foram avaliadas para o conjunto de dados de precipitação mensal para o Nordeste brasileiro. Após a identificação do método que reuniu o melhor conjunto de resultados para as medidas de avaliação EM , EMA , EQM , $REQM$, NSE , D e R^2 , foi analisado temporalmente o desempenho da técnica selecionada, ou seja, para cada mês de cada ano foram obtidas medidas de avaliação. Para isto, o procedimento de *validação cruzada* foi utilizado nos 816 meses para o período de 1950 à 2017, obtendo valores estimados para cada uma das 34 técnicas de interpolação com suas variações.

O desempenho dos diferentes métodos de interpolação foram avaliados utilizando medidas de avaliação descrita na [Tabela 2](#), em que os resultados dos valores estimados foram comparados com os dados observados. Assim, a combinação dessas medidas de avaliação permitiu identificar qual dentre os 34 métodos e suas variações apresentaram melhores desempenhos ([Tabela 3](#)). Desta forma, quanto mais próximos do valor 0 estiverem as medidas EM , EAM , EQM e $REQM$, indicam que a técnica de interpolação conseguiu boas estimativas. Por outro lado, as medidas NSE , D e R^2 que avalia a eficiência do modelo, quanto mais próximo do valor 1, exibem, em termos percentuais, o quanto o modelo se ajustou aos dados.

Tabela 3 – Valores obtidos através das medidas de avaliação de desempenho para os sete métodos de interpolação e suas variações.

Métodos	Medidas de avaliação						
	<i>EM</i>	<i>EAM</i>	<i>EQM</i>	<i>REQM</i>	<i>NSE</i>	<i>D</i>	<i>R</i> ²
<i>Inverse Distance Weighting</i>							
<i>p</i> = 0, 5	-0,65	32,90	3775,84	61,45	0,72	0,91	0,73
<i>p</i> = 1, 0	-0,45	31,49	3585,85	59,88	0,74	0,92	0,74
<i>p</i> = 1, 5	-0,24	30,52	3492,83	59,10	0,75	0,92	0,75
<i>p</i> = 2, 0	-0,10	30,18	3505,87	59,21	0,74	0,92	0,74
<i>p</i> = 2, 5	-0,03	30,29	3586,60	59,89	0,74	0,92	0,74
<i>p</i> = 3, 0	-0,01	30,65	3703,14	60,85	0,73	0,92	0,73
<i>p</i> = 3, 5	-0,01	31,12	3833,42	61,91	0,72	0,92	0,72
<i>p</i> = 4, 0	-0,02	31,59	3964,45	62,96	0,71	0,92	0,72
<i>p</i> = 4, 5	-0,03	32,03	4089,91	63,95	0,70	0,91	0,71
<i>p</i> = 5, 0	-0,04	32,44	4207,27	64,86	0,69	0,91	0,70
<i>Trend Surface Analysis</i>							
<i>n</i> = 1	-0,12	47,17	5533,83	74,39	0,60	0,86	0,60
<i>n</i> = 2	-0,03	40,98	4566,39	67,58	0,67	0,89	0,67
<i>n</i> = 3	0,02	37,80	4069,17	63,79	0,70	0,91	0,70
<i>n</i> = 4	0,03	36,35	3869,76	62,21	0,72	0,91	0,72
<i>n</i> = 5	0,02	34,92	3661,05	60,51	0,73	0,92	0,73
<i>Radial Basis Function</i>							
<i>Inverse Multiquadric</i>	0,05	39,94	19998,74	141,42	-0,46	0,69	0,31
<i>Multiquadric</i>	-1,83	59,05	88012,94	296,67	-5,42	0,36	0,09
<i>Multilog</i>	-0,29	47,86	41954,57	204,83	-2,06	0,53	0,18
<i>Natural Cubic Spline</i>	-0,92	66,88	32505,30	180,29	-1,37	0,61	0,22
<i>Thin Plate Spline</i>	-2,26	67,61	120372,89	346,95	-7,79	0,30	0,07
<i>Kriging</i>							
<i>Cubic</i>	-0,10	52,87	54382,78	233,20	-2,97	0,47	0,14
<i>Linear</i>	-0,13	31,63	3679,48	60,66	0,73	0,92	0,74
<i>Power</i>	-0,17	34,66	4821,76	69,44	0,65	0,90	0,67
<i>Exponential</i>	-0,16	31,58	3654,03	60,45	0,73	0,92	0,74
<i>Gaussian</i>	1,75	71,93	579561,48	761,29	-41,30	0,09	0,02
<i>Quadratic</i>	-0,25	31,62	3589,91	59,92	0,74	0,92	0,74
<i>Rational Quadratic</i>	-1,05	88,60	300869,22	548,52	-20,96	0,16	0,03
<i>Wave</i>	-0,99	43,34	85051,52	291,64	-5,21	0,33	0,08
<i>Logarithmic</i>	-104,68	879,83	216256339,10	14705,70	-15782,89	0,00	0,00
<i>Spherical</i>	-0,21	31,56	3584,47	59,87	0,74	0,93	0,74
<i>Pentasppherical</i>	-0,24	31,64	3596,31	59,97	0,74	0,92	0,74
<i>Modified Shepard</i>							
<i>Nearest Neighbours</i>	0,12	37,05	5574,62	74,66	0,59	0,89	0,64
<i>Natural Neighbours</i>	0,32	30,54	3770,61	61,41	0,73	0,92	0,73

Na Tabela 3, temos os resultados das medidas de avaliação para os valores estimados obtidos através da utilização dos sete métodos de interpolação e suas variações para o *IDW*, *Kriging*, *RBF* e *Trend Surface Analysis*, que foram aplicados aos 816 meses para o

período em estudo. O método que apresentou o melhor conjunto de medidas foi o *Modified Shepard* (Tabela 3). Entretanto, segundo SILVA et al. (2019), comparando métodos de interpolação para o estado de Pernambuco–Brasil, o *Modified Shepard* não apresentou um bom desempenho, com destaque para os valores de $NSE = -0,63$ e $R^2 = 0,29$, em um rank de 26 técnicas de interpolação com suas variações, o método ocupou a posição 16. Segundo Fim et al. (2017), se os dados não estiverem igualmente distribuídos, pode resultar em valores diferentes.

O segundo melhor método de interpolação foi o *IDW* para $p = 1,5$ dentre os 10 valores de p verificados. No entanto, os valores de avaliação para p igual 2,0, 2,5 e 3,0, foram bem semelhantes para as medidas NSE , D e R^2 . Entretanto, na literatura o parâmetro de ponderação mais indicado é $p = 2,0$ (FIM et al., 2017; SILVA et al., 2019), contudo segundo Pereira e Macedo (2017) o *IDW* apresentou o pior desempenho para o mapeamento da precipitação na bacia de Nova Ponte–MG.

Para o Kriging, dentre os 11 variogramas testados na interpolação, o que apresentou o melhor conjunto de medidas de avaliação foi o variograma *Spherical*. Os que apresentaram resultados similares ao variograma *Spherical* foram os *Liner*, *Exponential*, e o *Pentaspherical*. O *Power* e o *Quadratic* apresentaram também bons resultado, enquanto que o *Logarithmic* foi o que apresentou pior desempenho. Segundo Farias, Francisco e Senna (2017), avaliando diferentes método de interpolação para a região do litoral sul do estado do Rio de Janeiro, foi avaliado vários tipos de variogramas (*Exponential*, *linear*, *circular*, *esférico* e *gaussiano*) no *Kriging*, sendo o *Exponential* o que apresentou melhor resultado. Todavia, dentro do conjuntos de técnicas avaliadas no presente trabalho, o *Exponential* não apresentou bons resultados.

No método *Trend Surface Analysis*, foram utilizados diferentes graus de superfícies, tendo destaque a de quinto grau, $n = 5$. Por outro lado, o desempenho obtido em comparação aos demais método, o *Trend Surface Analysis* obteve o quarto melhor desempenho. Este resultado também é semelhante ao apresentado pelo *IDW* com $p = 3$, com destaque para as seguintes medidas *REQM*, NSE , D , R^2 . No entanto, para o $n = 4$, foram encontrados valores equivalentes em SILVA et al. (2019).

A técnica *Natural Neighbours* apresentou o quinto melhor desempenho. Pode-se observar na Tabela 3 que os valores das medidas NSE , D e R^2 são semelhantes para os métodos *Kriging* (*linear*, *exponential*, *spherial*, *pentaspherical* e *quadratic*), *Trend Surface Analysis* com $n = 5$ e 4 e para o método *IDW* com $p = 3$. Para MIRANDA et al. (2017), avaliando diferentes método de interpolação sobre dados meteorológicos para a cidade de Petrolina–PE, para a variável precipitação, o *Natural Neighbours* foi o que apresentou melhores resultados.

Dentre os 7 métodos avaliados, o que apresentou o pior desempenho foi o *Kriging Logarithmic*, com destaque para os valores de NSE , D , R^2 (Tabela 3). Dentre as funções base utilizadas a que apresentou o melhor desempenho foi a *Inverce Multiquadric*, porém em comparação aos resultados apresentados pelos demais métodos de interpolação, não é recomendado utilizar esta técnica. O Nearest Neighbours apresentou o segundo pior desempenho.

Após avaliar os 7 métodos de interpolação e suas variações para o *IDW*, *Trend Surface Analysis*, *Radial Basis Function* e *Kriging*, totalizando 34 variações, foi identificado que o método *Modified Shepard* reuniu o melhor conjunto de medidas de avaliação. Em seguida foi realizada a interpolação pelo método *Modified Shepard* e avaliado o desempenho temporal da técnica, utilizando as medidas de avaliação da Tabela 2 em cada mês do período entre Janeiro de 1950 a Dezembro de 2017.

Na Figura 8 pode-se observar a variação do *EM* em cada mês para o período em estudo. O menor valor foi de $-6,66$ para o mês de Maio do ano de 1963 e Fevereiro de 1965, enquanto que o seu maior valor foi $7,48$ para o mês de Novembro do ano de 1954 e Março de 1960. Para os demais meses e anos, os valores foram bem distribuídos entre os valores de $-1,96$ a $1,80$. Para os valores do *EMA* Figura 9, observa-se que não estão bem distriuídos, ressaltando o período entre os anos de 1980 a 2007, em que os valores variaram entre $3,48$ a $25,6$. Porém, no período de 1951 a 1973, os erros variaram entre os valores $25,68$ a $55,28$. Para o *EQM*, Figura 10, percebe-se que no mês de Agosto de 1985 e Outubro de 2002, apresentaram valores extremos acima de 40000 , enquanto que na maior parte do período, os valores variaram entre $42,59$ a $2923,59$, o que está próximo do valor global para o *EQM*. Por outro lado, os valores para a *REQM*, Figura 11, apresentam similaridade em sua distriuição com relação ao *EMA* (Figura 9).

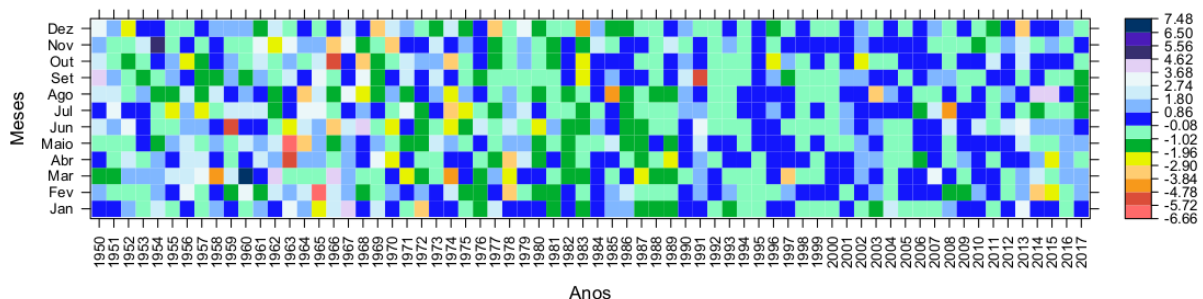


Figura 8 – Erro Médio mensal de 1950 à 2017 para o *Modified Shepard*.

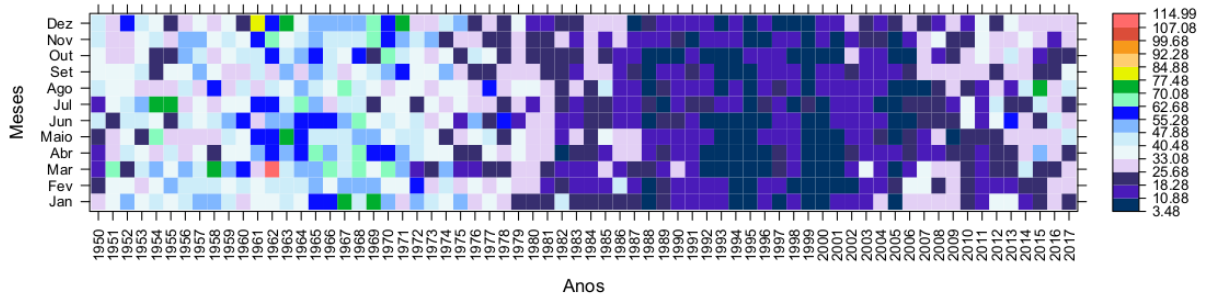


Figura 9 – Erro Médio Absoluto mensal de 1950 à 2017 para o *Modified Shepard*.

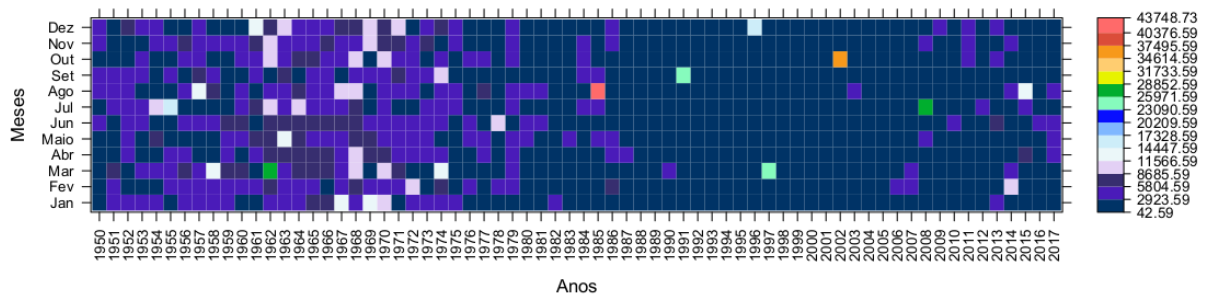


Figura 10 – Erro Quadrático Médio mensal de 1950 à 2017 para o *Modified Shepard*.

Analisando a eficiência do método utilizando a medida NSE , Figura 12, os valores estão bem distribuídos entre 0,68 e 0,82, próximos do valor geral para essa medida. No entanto, para os períodos de 1982 a 1983 e período de 1987 a 1989, para todos os meses temos os melhores resultados. Semelhante ao NSE , em distribuição temporal, temos o R^2 , que pode ser verificado na Figura 13. Para a medida D , D de Willmott, os valores estão distribuídos em sua grande maioria entre 0,81 a 0,97, que é semelhante ao valor global de D .

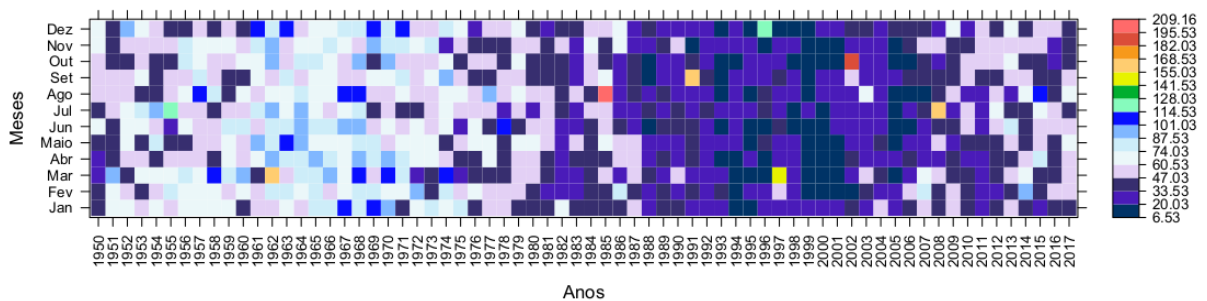


Figura 11 – Raiz do Erro Quadrático Médio mensal de 1950 à 2017 para o *Modified Shepard*.

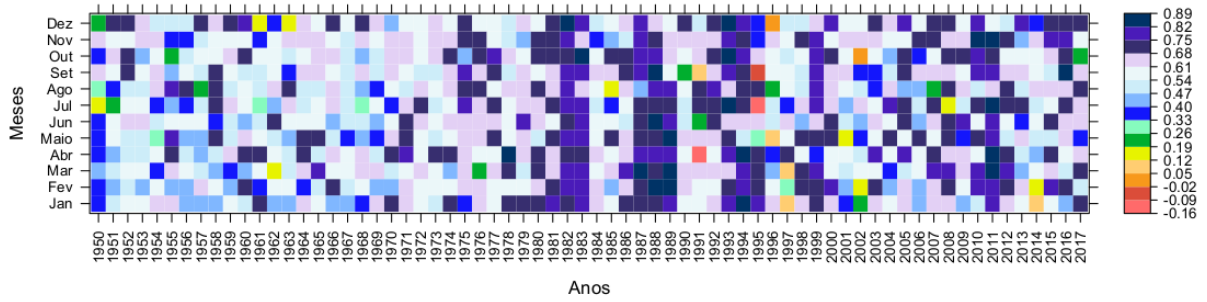


Figura 12 – Coeficiente de Nash-Sutcliffe mensal de 1950 à 2017 para o *Modified Shepard*.

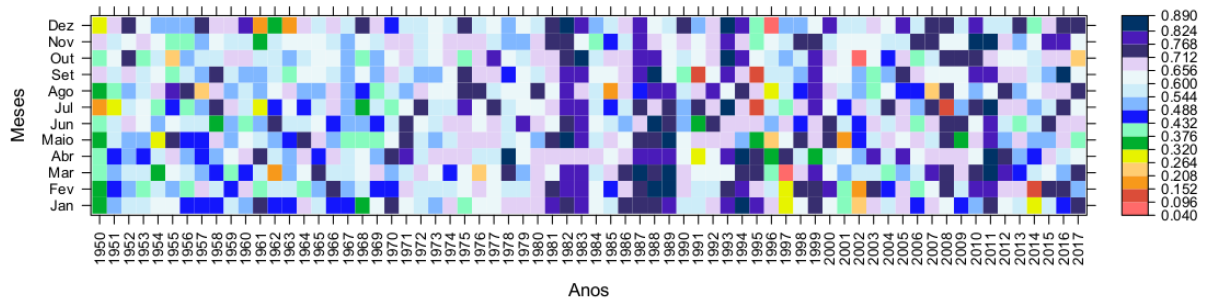


Figura 13 – Coeficiente de determinação R^2 mensal de 1950 à 2017 para o *Modified Shepard*.

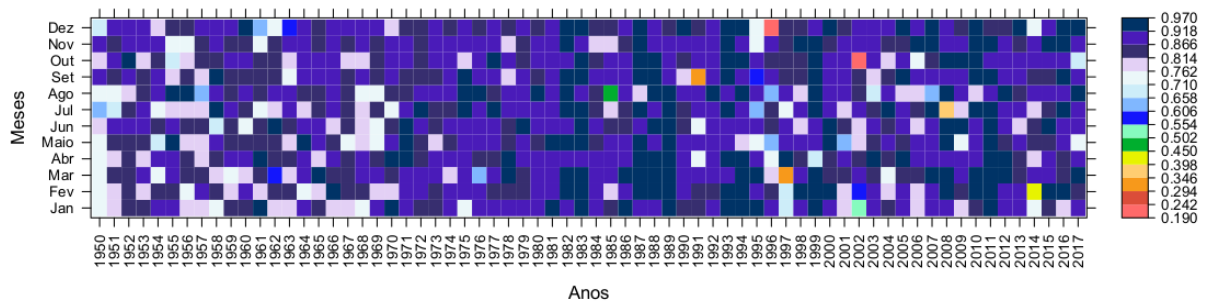


Figura 14 – D de Willmott's mensal de 1950 à 2017 para o *Modified Shepard*.

Nas Figuras 15 a 18 estão ilustrados os boxplots que correspondem às medidas de erro (EM , EMA , EQM e $REQM$) e de eficiência (NSE , D e R^2) para avaliação do desempenho do método de interpolação *Modified Shepard*. Por inspeção, para a medida EM na Figura 15a verifica-se que somente no mês de Janeiro a mediana apresentou valor mais distante do valor zero, indicando que o método não apresentou um resultado satisfatório durante este período. Além disso, observa-se maior variação nos dados nos meses de Fevereiro, Março, Junho e Agosto. Para a medida EMA (Figura 15b), o mês com menor valor da mediana foi Junho e o de maior valor foi Agosto. Isto indica que para

o mês de Junho o desempenho foi superior ao observado no mês de Agosto. Por outro lado, os meses com maior variabilidade dos valores estimados foram Março e Junho.

Ainda podemos ressaltar que foram observados menos *outliers* na Figura 15b em comparação com a Figura 15a. Segundo Gois et al. (2019), as variações de sistemas meteorológicos podem contribuir para o surgimento de *outliers*. Isso ocorre devido a complexidade de séries climáticas, uma vez que precipitação trata-se de uma variável não linear. Analisando as Figura 16a e Figura 16b, os meses de Março, Junho, Agosto e Dezembro apresentaram maior variabilidade nas medidas de avaliação quando comparado com os demais meses. Consequentemente, os meses com menor dispersão foram Julho e Setembro.

Avaliando a eficiência do método *Modified Shepard*, o mês de Abril expôs o valor mais próximo do ótimo ($NSE = 1$), revelando uma melhor estimacão entre o valor observado e o estimado (ver Figura 17a). Em contrapartida, além do mês de Abril, os meses de Julho, Setembro, Outubro, Novembro e Dezembro exibiram os maiores valores para a medida D (ver Figura 17b). Os menores valores observados tanto para a medida NSE como para R^2 ocorreram no mês de Fevereiro. Enquanto D apresentou menor valor no mês de Maio.

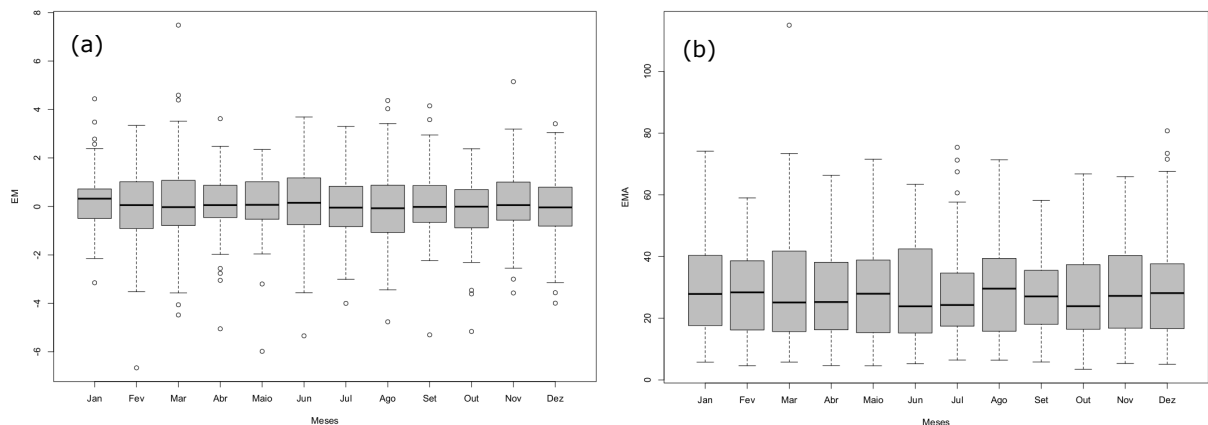


Figura 15 – Boxplot das medidas de avaliação (a) EM e (b) EMA para o período de Jan-1950 a Dez-2017.

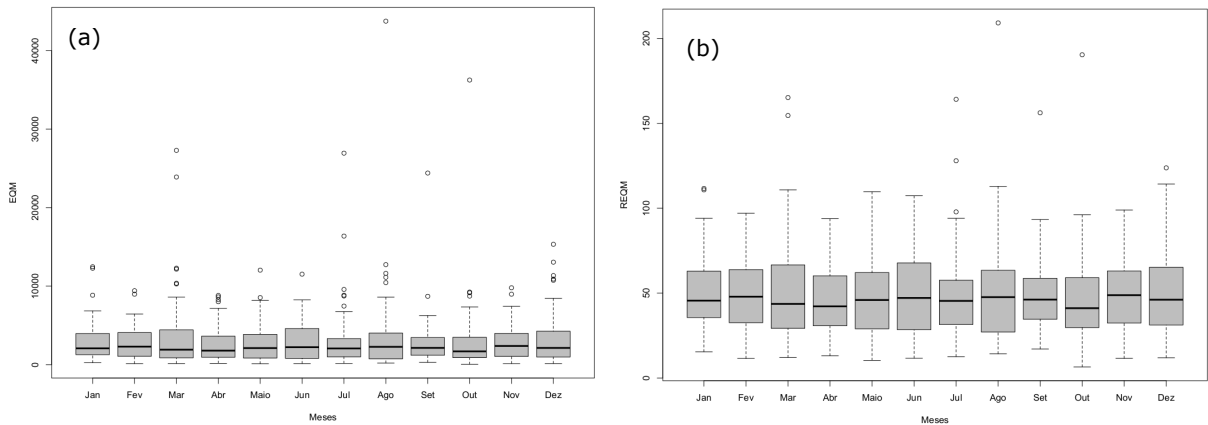


Figura 16 – Boxplot das medidas de avaliação (a) EQM e (b) $REQM$ para o período de Jan-1950 a Dez-2017.

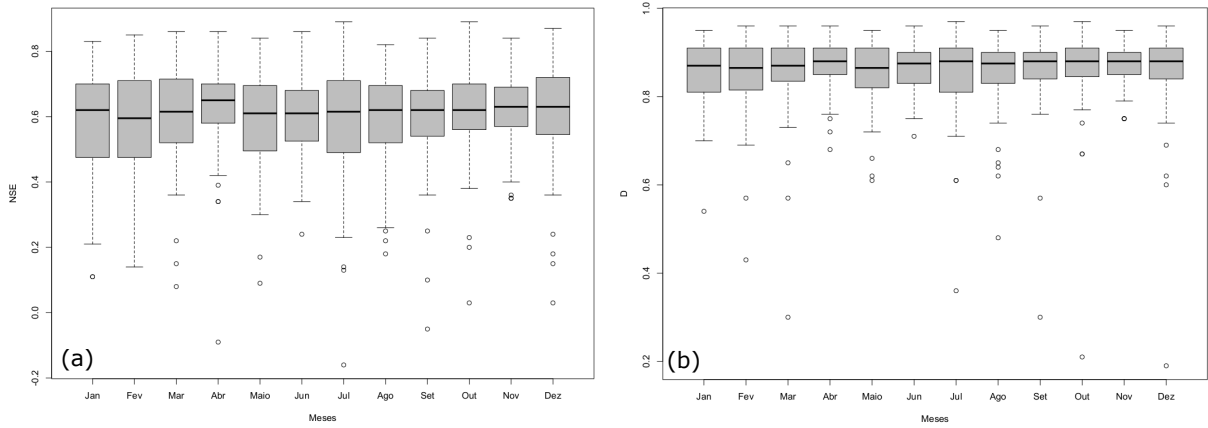


Figura 17 – Boxplot das medidas de avaliação (a) NSE e (b) D de Willmott's para o período de Jan-1950 a Dez-2017.

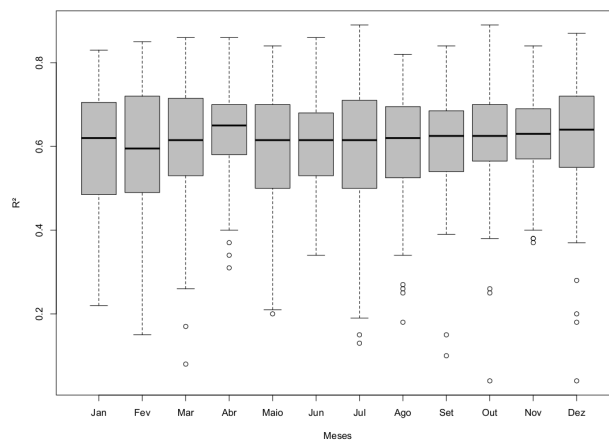


Figura 18 – Boxplot da medida de avaliação R^2 para o período de Jan-1950 a Dez-2017. para o *Modified Shepard*

6 Conclusão

Na comparação global dos métodos de interpolação, o *Modified Shepard* apresentou o melhor conjunto de medidas ($EM = -0,03$, $EMA = 30,64$ e $REQM = 57,69$, $NSE = 0,76$, $D = 0,93$ e $R^2 = 0,76$) sugerindo ser o melhor interpolador dentre os métodos avaliados para o conjunto de dados de precipitação mensal para o Nordeste brasileiro. O *IDW* com ponderação $p = 1,5$ se destacou como a segunda técnica bem avaliada, podendo ser uma outra opção importante para interpolar os dados faltantes. Seguindo um *ranking* de desempenho, apresentaram bons resultados em relação aos demais: o *Kriging* com variograma *Spherical*, o *Trend Surface Analysis* ($p = 1,5$), o *Natural Neighbours*, o *Nearest Neighbours* e o *Radial Basis Function*. Por outro lado, o pior desempenho foi alcançado pelo *Radial Basis Function* utilizando a função base *Thin Plate Spline*, com $NSE = -7,79$, $D = 0,30$ e $R^2 = 0,07$.

Os resultados mostram que é possível obter excelentes estimativas para precipitação mensal no Nordeste brasileiro a partir de método de interpolação em regiões onde possuam uma rede bem distribuída de estações meteorológicas. Por outro lado, limitações na disponibilidade dos dados podem dificultar futuros estudos. Objetivando solucionar esta dificuldade de limitações de dados, ferramenta de acesso e visualização da precipitação para o Nordeste brasileiro, pode ser desenvolvida e disponibilizada em site de instituições de pesquisas e ensino afim de servir de suporte para trabalhos futuros bem como instrumento de auxílio para políticas públicas.

Referências Bibliográficas

- ALBUQUERQUE, L. S. de et al. A relação entre o relevo e o clima como proposição de caracterização da fisiologia da paisagem em pernambuco, brasil. **Geo UERJ**, n. 34, p. 40942, 2019.
- AURENHAMMER, F. Voronoi diagrams—a survey of a fundamental geometric data structure. **ACM Computing Surveys (CSUR)**, ACM New York, NY, USA, v. 23, n. 3, p. 345–405, 1991.
- BRITO, S. et al. Frequency, duration and severity of drought in the semiarid northeast brazil region. **International Journal of Climatology**, Wiley Online Library, v. 38, n. 2, p. 517–529, 2018.
- CAVALCANTI, I. F. **Tempo e clima no Brasil**. São Paulo: Oficina de textos, 2016.
- CHENG, M. et al. Performance assessment of spatial interpolation of precipitation for hydrological process simulation in the three gorges basin. **Water**, Multidisciplinary Digital Publishing Institute, v. 9, n. 11, p. 838, 2017.
- COLLISCHONN, W.; DORNELLES, F. Hidrologia para engenharia e ciências ambientais. **Porto Alegre: Editora ABRH**, 2013.
- DELAUNAY, B. et al. Sur la sphere vide. **Izv. Akad. Nauk SSSR, Otdelenie Matematicheskii i Estestvennyka Nauk**, v. 7, n. 793-800, p. 1–2, 1934.
- FARIAS, O. G.; FRANCISCO, C. N.; SENNA, M. C. A. Avaliação de métodos de interpolação espacial aplicados à pluviosidade em região montanhoso no litoral sul do estado do Rio de Janeiro. **Revista Brasileira de Climatologia**, v. 21, 2017.
- FERREIRA, A. G.; MELLO, N. G. da S. Principais sistemas atmosféricos atuantes sobre a região nordeste do brasil e a influência dos oceanos pacífico e atlântico no clima da região. **Revista brasileira de climatologia**, v. 1, n. 1, 2005.
- FILHO, D. F. F. et al. Aplicação de técnicas de interpolação para espacialização de chuvas da rede hidrográfica: Estudo de caso calha norte–pa. **Revista Brasileira de Climatologia**, v. 24, 2019.
- FIM, B. M. et al. Interpolação espacial da precipitação mensal média no distrito federal, Brasil. In: XXII SIMPOSIO BRASILEIRO DE RECURSOS HIDRICOS. Florianópolis–SC, 2017.
- FORNBERG, B.; LARSSON, E.; WRIGHT, G. A new class of oscillatory radial basis functions. **Computers & Mathematics with Applications**, v. 51, n. 8, p. 1209–1222, 2006.

- FRANKE, R.; NIELSON, G. Smooth interpolation of large sets of scattered data. **International journal for numerical methods in engineering**, Wiley Online Library, v. 15, n. 11, p. 1691–1704, 1980.
- GASPARIN, E. et al. **Otimização econômica com análise dos riscos na produção agrícola de uma propriedade rural do oeste paranaense**. Dissertação (Mestrado) — Universidade Tecnológica Federal do Paraná, 2018.
- GOIS, G. de et al. Variabilidade anual e mensal da chuva e da temperatura do ar no município de resende, rio de janeiro. **Revista Brasileira de Climatologia**, v. 24, 2019.
- GOLDEN SOFTWARE, LLC. **Surfer13. R: Gridding e Superfícies 3D**. 2015. Disponível em: <<https://www.goldensoftware.com/>>.
- HALLAK, R.; FILHO, A. J. P. Metodologia para análise de desempenho de simulações de sistemas convectivos na região metropolitana de são paulo com o modelo arps: sensibilidade a variações com os esquemas de advecção e assimilação de dados. **Revista Brasileira de Meteorologia**, v. 26, n. 4, p. 591–608, 2011.
- HARZHEIM, L.; WARNECKE, U. Robustness optimization of the position of an anti-roll bar link to avoid the toggling of a rear axle stabilizer. **Structural and Multidisciplinary Optimization**, Springer, v. 42, n. 2, p. 315–323, 2010.
- JAKOB, A. A. E.; YOUNG, A. F. O uso de métodos de interpolação espacial de dados nas análises sociodemográficas. **Anais**, p. 1–22, 2016.
- JAVARI, M. Comparison of interpolation methods for modeling spatial variations of precipitation in iran. **International Journal of Environmental and Science Education**, v. 12, n. 5, p. 1037–1054, 2017.
- KRIGE, D. G. A statistical approach to some basic mine valuation problems on the witwatersrand. **Journal of the Southern African Institute of Mining and Metallurgy**, Southern African Institute of Mining and Metallurgy, v. 52, n. 6, p. 119–139, 1951.
- LI, J.; HEAP, A. D. A review of comparative studies of spatial interpolation methods in environmental sciences: Performance and impact factors. **Ecological Informatics**, Elsevier, v. 6, n. 3-4, p. 228–241, 2011.
- LI, X. et al. Influencing factors of the spatial–temporal variation of layered soils and sediments moistures and infiltration characteristics under irrigation in a desert oasis by deterministic spatial interpolation methods. **Water**, Multidisciplinary Digital Publishing Institute, v. 11, n. 7, p. 1483, 2019.
- LIMA, F. et al. Avaliação das estimativas de irradiação solar do brams e desenvolvimento de uma técnica estatística de pós-processamento para o norte do brasil. **Revista Brasileira de Geografia Física**, v. 10, n. 01, p. 079–093, 2017.
- LUNDGREN, W. J. C.; SILVA, J. A. A. d.; FERREIRA, R. L. C. A precisão da estimativa do erro da krigagem pela validação cruzada. **Floresta e Ambiente**, Instituto de Florestas/UFRRJ, v. 24, p. 0–0, 2018.

- MARUYAMA, T.; KAWACHI, T.; SINGH, V. P. Entropy-based assessment and clustering of potential water resources availability. **Journal of Hydrology**, Elsevier, v. 309, n. 1-4, p. 104–113, 2005.
- MEDEIROS, M. H. F. de et al. Análise do índice de chuva dirigida em cidades do paran e sua importncia no projeto de fachadas de edifcios. **Revista Tcnico-Cientfica**, v. 1, 2017.
- MENEGHETTI, G.; FERREIRA, N. Variabilidade sazonal e interanual da precipitao no nordeste brasileiro. **Simpsio Brasileiro de Sensoriamento Remoto**, v. 14, p. 1685–1689, 2009.
- MIRANDA, R. A. C. de; OLIVEIRA, M. V. S. de; SILVA, D. F. da. Ciclo hidrogrfico planetrio: abordagens e conceitos. **Geo Uerj**, v. 1, n. 21, p. 109–119, 2010.
- MIRANDA, R. d. Q. et al. Avaliao de dez mtodos diferentes de interpolao sobre dados meteorolgicos em petrolina, pernambuco. In: CONGRESSO BRASILEIRO DE AGROMETEOROLOGIA, 20; SIMPSIO DE MUDANAS . . . **Embrapa Semirido-Artigo em anais de congresso (ALICE)**. Juazeiro–BA, 2017.
- MOLION, L. C. B.; BERNARDO, S. d. O. Uma reviso da dinmica das chuvas no nordeste brasileiro. **Revista Brasileira de Meteorologia**, v. 17, n. 1, p. 1–10, 2002.
- MONTAO, M.; SOUZA, M. P. d. Integrao entre planejamento do uso do solo e de recursos hdricos: a disponibilidade hdrica como critrio para a localizao de empreendimentos. **Engenharia Sanitaria e Ambiental**, SciELO Brasil, v. 21, n. 3, p. 489–495, 2016.
- NASH, J. E.; SUTCLIFFE, J. V. River flow forecasting through conceptual models part i-a discussion of principles. **Journal of hydrology**, Elsevier, v. 10, n. 3, p. 282–290, 1970.
- NOBRE, C. et al. Aspectos da climatologia dinmica do brasil. **Climanlise, Nmero especial**, 1986.
- NOGUEIRA, A. K. M.; SANTANA, A. C. de. Influncia das chuvas na oferta de castanha-do-brasil e o impacto no benefcio socioeconmico e ambiental, no oeste do estado do par. **Desenvolvimento e Meio Ambiente**, v. 45, 2018.
- NOLLI, A. C. Estudo comparativo dos mtodos de dimensionamento de reservatriotimo em aproveitamento de guas pluviais. Universidade Federal de Minas Gerais, 2017.
- OLIVEIRA, L. B. de. Uso e manejo da gua na regio semirida do nordeste do brasil. **Anais da Academia Pernambucana de Cincia Agronmica**, v. 11/12, p. 50–64, 2015.
- OLIVEIRA, L. F. G. S. Dimensionamento e anlise de desempenho hidrulico de estacionamentos com drenagem convencional e pavimento permevel, apoiado por modelagem computacional. **Rio de Janeiro**, 2018.
- PENG, J. et al. A review of spatial downscaling of satellite remotely sensed soil moisture. **Reviews of Geophysics**, Wiley Online Library, v. 55, n. 2, p. 341–366, 2017.

- PEREIRA, A. L. S.; MACEDO, D. R. Análise de métodos de interpolação para o mapeamento da precipitação na bacia do reservatório de Nova Ponte/MG. **Cadernos do LESTE**, v. 17, n. 17, 2017.
- QUININO, R. C.; REIS, E. A.; BESSEGATO, L. F. O coeficiente de determinação r^2 como instrumento didático para avaliar a utilidade de um modelo de regressão linear múltipla. **Belo Horizonte: UFMG**, 1991.
- SANTOS, B. C. D. et al. Análise do efeito orográfico na distribuição espacial das chuvas no município de Itirapina-sp. XIII Simpósio Brasileiro de Climatologia Geográfica/UFRRJ, 2018.
- SANTOS, G. T.; SOUZA, M. C. d.; FORTES, M. Use of radial basis functions for meshless numerical solutions applied to financial engineering barrier options. **Pesquisa Operacional**, SciELO Brasil, v. 29, n. 2, p. 419–437, 2009.
- SANTOS, W. M. dos et al. Variabilidade espacial da sazonalidade da chuva no semiárido brasileiro. **Journal of Environmental Analysis and Progress**, v. 2, n. 4, p. 368–376, 2017.
- SHEPARD, D. A two-dimensional interpolation function for irregularly-spaced data. In: ACM. **Proceedings of the 1968 23rd ACM national conference**. [S.l.], 1968. p. 517–524.
- SIBSON, R. A brief description of natural neighbour interpolation. **Interpreting multivariate data**, John Wiley & Sons, 1981.
- SILVA, A. S. A. et al. Comparison of interpolation methods for spatial distribution of monthly precipitation. **JOURNAL OF HYDROLOGIC ENGINEERING**, Elsevier, v. 24, p. 04018068–1–04018068–11, 2019.
- SILVA, C. B.; OLIVEIRA, L. F. C. de. Relação intensidade-duração-frequência de chuvas extremas na região nordeste do Brasil. **Revista Brasileira de Climatologia**, v. 20, 2017.
- SILVA, C. M. dos S.; OLIVEIRA, J. V. da C.; OLIVEIRA, L. de. Estudo climatológico de precipitação para as capitais do nordeste brasileiro. 2017.
- SILVA, J. C. da et al. Análise de distribuição de chuva para Santa Maria, RS. **R. Bras. Eng. Agríc. Ambiental**, SciELO Brasil, v. 11, n. 1, p. 67–72, 2007.
- SILVA, S. do N. et al. Efeitos da escassez hídrica na economia do perímetro irrigado de São Gonçalo, Paraíba. **Revista Verde de Agroecologia e Desenvolvimento Sustentável**, Grupo Verde de Agroecologia e Abelhas, v. 12, n. 1, p. 132–137, 2017.
- STEINKE, E. T. **Climatologia fácil**. [S.l.]: Oficina de Textos, 2016.
- TAYOB, L. J. D. et al. Determinação da relação entre a intensidade da chuva e refletividade radar no litoral do estado de Alagoas usando disdrômetro a laser. Universidade Federal de Alagoas, 2014.

- THIESSEN, A. H. Precipitation averages for large areas. **Monthly weather review**, v. 39, n. 7, p. 1082–1089, 1911.
- VIEIRA, H. P. M. et al. Análise dos índices pluviométricos como subsídio a gestão dos processos de degradação ambiental: Caso da sub-bacia santo antônio do maratuã-silva jardim/rj. 2014.
- VILLELA, S. M.; MATTOS, A. **Hidrologia Aplicada**. [S.l.]: McGrawHill do Brasil, 1975.
- VORONOI, G. Nouvelles applications des paramètres continus à la théorie des formes quadratiques. deuxième mémoire. recherches sur les paralléloèdres primitifs. **Journal für die reine und angewandte Mathematik**, v. 134, p. 198–287, 1908.
- WANG, S. et al. Comparison of interpolation methods for estimating spatial distribution of precipitation in ontario, canada. **International Journal of Climatology**, Wiley Online Library, v. 34, n. 14, p. 3745–3751, 2014.
- WEBSTER, R.; OLIVER, M. A. et al. **Statistical methods in soil and land resource survey**. [S.l.]: Oxford University Press (OUP), 1990.
- WILLMOTT, C. J. On the validation of models. **Physical geography**, Taylor & Francis, v. 2, n. 2, p. 184–194, 1981.
- WONG, T.-T. Performance evaluation of classification algorithms by k-fold and leave-one-out cross validation. **Pattern Recognition**, Elsevier, v. 48, n. 9, p. 2839–2846, 2015.
- XAVIER, A. C.; KING, C. W.; SCANLON, B. R. Daily gridded meteorological variables in brazil (1980–2013). **International Journal of Climatology**, Wiley Online Library, v. 36, n. 6, p. 2644–2659, 2016.
- XIAO, Y. et al. Geostatistical interpolation model selection based on arcgis and spatio-temporal variability analysis of groundwater level in piedmont plains, northwest china. **SpringerPlus**, Springer, v. 5, n. 1, p. 425, 2016.
- YNOUE, R. Y. et al. **Meteorologia: noções básicas**. São Paulo: Oficina de Textos, 2017.