



UNIVERSIDADE FEDERAL DO CEARÁ
CAMPUS DE CRATEÚS
CURSO DE ENGENHARIA CIVIL

GUILHERME OLIVEIRA RODRIGUES

**ANÁLISE DE PREENCHIMENTO DE FALHAS EM SÉRIES HISTÓRICAS DE
PRECIPITAÇÃO MENSAL NO CEARÁ**

CRATEÚS

2026

GUILHERME OLIVEIRA RODRIGUES

ANÁLISE DE PREENCHIMENTO DE FALHAS EM SÉRIES HISTÓRICAS DE
PRECIPITAÇÃO MENSAL NO CEARÁ

Trabalho de Conclusão de Curso I apresentado
ao Curso de Engenharia Civil da Universidade
Federal do Ceará, como requisito parcial à
obtenção do título de bacharel em Engenharia
Civil

Orientador: Prof. Dr. Alan Michell Barros
Alexandre.

CRATEÚS

2026

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

R613a Rodrigues, Guilherme Oliveira.

Análise de preenchimento de falhas em séries históricas de precipitação mensal no Ceará / Guilherme Oliveira Rodrigues. – 2026.

45 f. : il. color.

Trabalho de Conclusão de Curso (graduação) – Universidade Federal do Ceará, Campus de Crateús, Curso de Engenharia Civil, Crateús, 2026.

Orientação: Prof. Dr. Alan Michell Barros Alexandre.

1. Precipitação pluviométrica. 2. Preenchimento de falhas. 3. Séries históricas. 4. Métodos estatísticos. I. Título.

CDD 620

GUILHERME OLIVEIRA RODRIGUES

ANÁLISE DE PREENCHIMENTO DE FALHAS EM SÉRIES HISTÓRICAS DE
PRECIPITAÇÃO MENSAL NO CEARÁ

Trabalho de Conclusão de Curso I apresentado
ao Curso de Engenharia Civil da Universidade
Federal do Ceará, como requisito parcial à
obtenção do título de bacharel em Engenharia
Civil

Orientador: Prof. Dr. Alan Michell Barros
Alexandre.

Aprovado em 22/01/2026.

BANCA EXAMINADORA

Prof. Dr. Alan Michell Barros Alexandre (Orientador)
Universidade Federal do Ceará (UFC)

Profa. Dra. Deborah Mithya Barros Alexandre
Companhia de Gestão de Recursos Hídricos

Prof. José Kerlly Soares de Araújo
Universidade Federal do Ceará – Departamento de Engenharia Hidráulica e Ambiental

AGRADECIMENTOS

Agradeço primeiramente a Deus, por me conceder saúde, força e perseverança ao longo de toda essa jornada.

À minha mãe, Sandra, ao meu pai, Junior, e à minha avó, Maria, deixo meu mais sincero agradecimento pelo amor, apoio incondicional, incentivo constante e por nunca medirem esforços para que eu pudesse alcançar meus objetivos.

Ao meu orientador, Professor Alan, agradeço imensamente pelos ensinamentos compartilhados, pela paciência, dedicação e pelas valiosas orientações que foram fundamentais para o desenvolvimento e conclusão deste trabalho.

Agradeço também aos colegas e amigos que fizeram parte dessa caminhada, contribuindo com apoio, companheirismo, trocas de conhecimento e momentos que tornaram essa trajetória mais leve e significativa.

A todos que, direta ou indiretamente, contribuíram para a realização deste trabalho, meu muito obrigado.

“Na vida, nada deve ser temido, apenas compreendido”

Marie Curie

RESUMO

A análise de séries históricas de precipitação é essencial para estudos climatológicos e hidrológicos, principalmente em áreas com alta variabilidade espacial e temporal das chuvas, como o Ceará. No entanto, a confiabilidade das análises pode ser afetada pela ocorrência frequente de falhas nos registros pluviométricos. Neste cenário, o objetivo deste estudo foi analisar a eficácia de diversos métodos tradicionais de preenchimento de falhas em séries históricas mensais de precipitação no Ceará, levando em conta estratégias com e sem a separação sazonal mensal. Analisaram-se séries de registros com cerca de 50 anos, obtidas da plataforma HidroWeb/ANA. Esses dados foram distribuídos em dez regiões climáticas diferentes do estado, com uma estação principal definida para cada região e estações secundárias situadas a um máximo de 75 km de distância. Os métodos analisados foram a Ponderação Regional (PR), a Regressão Linear Múltipla (RLM), a Razão Normal (MRN) e a Ponderação Regional com Base em Regressões Lineares (PRBRL). Esses métodos foram aplicados de maneira consistente às séries, e o desempenho foi avaliado por meio de validação cruzada e métricas estatísticas de erro. Os resultados mostraram que, na maioria das regiões analisadas, os métodos tiveram desempenhos semelhantes, com diferenças pequenas entre as métricas avaliadas. A regressão linear múltipla se destacou como a técnica com melhor desempenho médio, embora com ganhos modestos em comparação com as outras. A análise da sazonalidade mensal afetou o desempenho dos métodos, mas não trouxe ganhos consistentes; em 8 de 10 regiões o desempenho foi inferior ao cenário global (Δ MAE positivo) e as melhorias ocorreram em poucas regiões e com pequena magnitude. Isso sugere que a seleção da técnica de preenchimento de falhas deve levar em conta não só o desempenho estatístico, mas também a simplicidade, a disponibilidade de dados e o custo computacional.

Palavras-chave: precipitação pluviométrica; preenchimento de falhas; séries históricas; métodos estatísticos.

ABSTRACT

A historical precipitation time series analysis is essential for climatological and hydrological studies, especially in areas with high spatial and temporal variability of rainfall, such as Ceará. However, the reliability of the analyses can be affected by the frequent occurrence of gaps in pluviometric records. In this scenario, the objective of this study was to analyze the effectiveness of several traditional methods for filling gaps in monthly historical precipitation time series in Ceará, considering strategies with and without monthly seasonal separation. Series of records spanning about 50 years were analyzed, obtained from the HidroWeb/ANA platform. These data were distributed across ten different climatic regions of the state, with a main station defined for each region and secondary stations located at a maximum distance of 75 km. The methods analyzed were Regional Weighting (PR), Multiple Linear Regression (RLM), Normal Ratio (MRN), and Regional Weighting Based on Linear Regressions (PRBRL). These methods were applied consistently to the series, and performance was evaluated through cross-validation and statistical error metrics. The results showed that, in most of the regions analyzed, the methods had similar performance, with small differences among the evaluated metrics. Multiple linear regression stood out as the technique with the best average performance, although with modest gains compared to the others. The analysis of monthly seasonality affected the performance of the methods, but did not yield consistent gains; in 8 out of 10 regions performance was inferior to the global scenario (positive Δ MAE), and improvements occurred in only a few regions and with small magnitude. This suggests that the selection of the gap-filling technique should take into account not only statistical performance, but also simplicity, data availability, and computational cost.

Keywords: pluviometric precipitation; gap filling; historical time series; statistical methods.

LISTA DE SIGLAS

ANA	Agência Nacional de Águas e Saneamento Básico
ASAN	Anticiclone Subtropical do Atlântico Norte
ASAS	Anticiclone Subtropical do Atlântico Sul
Bias	Viés (diferença média entre valores estimados e observados)
ENOS	El Niño–Oscilação Sul
FMA	Fevereiro-Março-Abril (Trimestre)
FUNCEME	Fundação Cearense de Meteorologia e Recursos Hídricos
LOTO	Leave-One-Time-Out (validação cruzada)
MAE	Mean Absolute Error (Erro Médio Absoluto)
MAM	Março-Abril-Maio (Trimestre)
MRN	Método da Razão Normal
OLS	Ordinary Least Squares
OMM	Organização Meteorológica Mundial
PR	Ponderação Regional
PRBRL	Ponderação Regional com Base em Regressões Lineares
RLM	Regressão Linear Múltipla
RMSE	Root Mean Square Error (Raiz do Erro Quadrático Médio)
SIRGAS	Sistema de Referência Geocêntrico para as Américas
UTM	Universal Transverse Mercator (Sistema de Coordenadas)
WMO	World Meteorological Organization
ZCIT	Zona de Convergência Intertropical

LISTA DE FIGURAS

Figura 1 - Fluxograma da metodologia.....	24
Figura 2 - Mapa de localização das estações pluviométricas selecionadas.....	26
Figura 3 - Distribuição dos valores de MAE por método de preenchimento no cenário sem sazonalidade.....	35
Figura 4 - Ranking relativo dos métodos de preenchimento por região, baseado no MAE no cenário sem sazonalidade.....	36
Figura 5 - Valores mensais de MAE por método de preenchimento, considerando a sazonalidade.....	37
Figura 6 - Perfil sazonal do viés (média mensal) por método de preenchimento com consideração da sazonalidade.....	38
Figura 7 - Distribuição espacial do método de preenchimento com melhor desempenho por região.....	41

LISTA DE TABELAS

Tabela 1 - Caracterização da base de dados pluviométricos utilizados no estudo.....	28
Tabela 2 - Critérios adotados para seleção das estações pluviométricas.....	29
Tabela 3 - Cenários considerados na implementação dos métodos de preenchimento.....	31
Tabela 4 - Métricas utilizadas na avaliação do desempenho dos métodos.....	32
Tabela 5 - Métricas globais de desempenho dos métodos de preenchimento de falhas no cenário sem consideração da sazonalidade.....	33
Tabela 6 - Comparação do desempenho dos métodos de preenchimento por região nos cenários com e sem consideração da sazonalidade, com base no MAE.....	39

SUMÁRIO

1. INTRODUÇÃO.....	13
1.1 Contextualização.....	13
1.2 Justificativa.....	14
1.3. Objetivos.....	15
<i>1.3.1 Objetivo geral.....</i>	<i>15</i>
<i>1.3.2 Objetivos específicos.....</i>	<i>15</i>
1.4 Roteiro de trabalho.....	15
2. REFERENCIAL TEÓRICO.....	17
2.1 Séries históricas e o problema de falhas.....	17
<i>2.1.1 Importância de séries longas e contínuas.....</i>	<i>17</i>
<i>2.1.2 Causas e consequências das falhas.....</i>	<i>17</i>
<i>2.1.3 Necessidade de métodos de imputação.....</i>	<i>18</i>
2.2 Clima e precipitação no Ceará.....	19
2.3 Métodos de preenchimento de falhas.....	19
<i>2.3.1 Método da ponderação regional (PR).....</i>	<i>19</i>
<i>2.3.2 Método da regressão linear múltipla (RLM).....</i>	<i>20</i>
<i>2.3.3 Método da razão normal (MRN).....</i>	<i>21</i>
<i>2.3.4 Método da ponderação por correlação (PRBRL).....</i>	<i>22</i>
3. METODOLOGIA.....	24
3.1 Região de estudo.....	24
3.2 Base de dados.....	27
<i>3.2.1 Coleta de dados pluviométricos.....</i>	<i>27</i>
<i>3.2.2 Critérios de seleção e consistência dos dados.....</i>	<i>28</i>
3.3 Métodos de preenchimento de falhas avaliados.....	29
3.4 Implementação computacional e estratégia de validação.....	30
3.5 Métricas de avaliação de desempenho.....	31
4. RESULTADOS E DISCUSSÕES.....	33
4.1 Apresentação geral dos resultados.....	33
4.2 Avaliação global dos métodos (sem consideração da sazonalidade).....	33
4.3 Avaliação com consideração da sazonalidade mensal.....	36
<i>4.3.1 Apresentação geral da análise sazonal.....</i>	<i>36</i>
<i>4.3.2 Desempenho mensal dos métodos.....</i>	<i>37</i>
4.4 Comparação entre abordagens com e sem consideração da sazonalidade.....	38
4.5 Análise regional dos resultados.....	40
5. CONCLUSÃO.....	42
REFERÊNCIAS.....	43

1. INTRODUÇÃO

1.1 Contextualização

A precipitação atmosférica constitui um dos principais elementos do clima e representa o mecanismo fundamental de transferência de água para a superfície terrestre no ciclo hidrológico. De acordo com Chow, Maidment e Mays (1988), a água precipitada pode ser interceptada pela vegetação, escoar superficialmente ou infiltrar no solo, contribuindo para a recarga de aquíferos e para a alimentação de cursos d'água. Dessa forma, a precipitação exerce papel central na hidrologia de superfície e subterrânea, sendo um componente essencial para a compreensão dos processos hidrológicos e para o planejamento e a gestão dos recursos hídricos.

No âmbito da climatologia e hidrologia, a análise de dados de precipitação é uma atividade essencial para o planejamento em diversas áreas. Segundo Tucci (2009, p. 177), a precipitação “é entendida em hidrologia como toda água proveniente do meio atmosférico que atinge a superfície terrestre. Neblina, chuva, granizo, saraiva, orvalho, geada e neve são formas diferentes de precipitação”. O conhecimento sobre a distribuição deste fenômeno é, portanto, de suma importância para a agricultura, gestão de recursos hídricos e prevenção de desastres naturais.

Dentre as diferentes formas de precipitação, a precipitação pluvial destaca-se como a mais importante, tanto pela sua frequência, quanto pelo volume de água. Segundo Ayoade (2010), a chuva constitui o principal tipo de precipitação analisada nos estudos climatológicos, sendo fundamental para a compreensão da distribuição espacial e temporal da umidade, bem como para a análise das variações climáticas. Por essa razão, a maior parte das análises relacionadas à precipitação baseia-se nos registros pluviométricos, os quais permitem avaliar a intensidade, a duração e variabilidade das chuvas ao longo dos anos.

O principal objetivo de um posto pluviométrico é de se obter uma série histórica de precipitações ao longo dos anos. Porém, na prática podem ocorrer falhas ou ausência de dados em certos períodos, exigindo uma análise dos dados antes de sua utilização. Essas falhas podem ser causadas por erros sistemáticos, como preenchimento errado do valor na caderneta, soma errada do número de provetas, até problemas instrumentais, como danos à aparelhagem, problemas mecânicos no visor (TUCCI, 2009, p. 182).

Além dos erros sistemáticos e dos problemas instrumentais, a ocorrência de falhas em séries históricas de precipitação podem estar associadas a aspectos relacionados à operação e à gestão dos postos de medição. Tais fatores podem comprometer a continuidade e a

consistência dos registros pluviométricos, reforçando a necessidade de uma análise criteriosa dos dados antes de sua utilização em estudos hidrológicos.

A utilização de séries históricas contendo falhas pode acarretar em consequências para a qualidade e confiabilidade de qualquer estudo hidrológico. A ausência de dados pode levar a análises tendenciosas e resultados precipitados, comprometendo os resultados da pesquisa (VIEIRA et al., 2018). Segundo Wanderley, Amorim e Carvalho (2014), isso ocorre porque uma série incompleta pode exibir padrões de variabilidade temporal que não condizem com a realidade. Dessa forma, a etapa de tratamento e preenchimento de dados torna-se indispensável.

Nesse contexto, a literatura enfatiza que a avaliação da qualidade e da consistência dos dados hidrológicos constitui uma etapa indispensável para garantir a confiabilidade das análises estatísticas e hidrológicas. De acordo com as diretrizes da Organização Meteorológica Mundial, procedimentos de controle de qualidade, padronização e tratamento de dados são fundamentais para a identificação de inconsistências, a verificação da representatividade temporal dos registros e a mitigação dos efeitos de falhas nas séries históricas. A adoção dessas práticas contribui para o aumento da robustez dos estudos baseados em séries de precipitação (WMO, 2008).

Apesar dessas inconsistências serem comuns, existe um grande interesse em métodos que possam preencher essas falhas. Técnicas como a Ponderação Regional, onde é possível fazer o preenchimento dessas falhas utilizando dados de postos pluviométricos vizinhos, esse tipo de preenchimento não substitui os dados originais, e somente podem ser aplicados para dados em intervalos mensais ou anuais. Sendo possível também, a utilização de um método simplificado, utilizando regressões lineares simples ou múltiplas para gerar informações nos períodos com falha (COLLISCHONN; DORNELLES, 2015).

1.2 Justificativa

A necessidade de dados pluviométricos com boa qualidade é de grande importância para o estado do Ceará. Caracterizado por clima semiárido e uma forte dependência de seus reservatórios para a segurança hídrica, erros nas séries históricas podem ter impactos significativos na tomada de decisão por parte dos órgãos gestores e usuários de recursos hídricos.

Com o exposto, este trabalho se justifica pela necessidade de selecionar e aplicar, de forma criteriosa, métodos de preenchimento de falhas em séries históricas de precipitação no Ceará. Falhas e inconsistências em dados pluviométricos são frequentes e podem

comprometer estimativas hidrológicas e climatológicas utilizadas no planejamento e na operação de sistemas hídricos. Além disso, diferentes métodos podem apresentar desempenhos distintos conforme as características climáticas e a distribuição espacial das estações, o que torna inadequado adotar uma única técnica sem avaliação prévia. Assim, este estudo compara diferentes métodos de preenchimento e critérios de sazonalidade, utilizando métricas de desempenho e validação, com o objetivo de indicar abordagens mais adequadas para diferentes regiões do estado, contribuindo para aumentar a confiabilidade das séries e subsidiar decisões de gestão de recursos hídricos.

1.3. Objetivos

1.3.1 Objetivo geral

Avaliar a eficiência de diferentes métodos de preenchimento de falhas em séries históricas pluviométricas mensais, com o objetivo de determinar a técnica mais adequada para diferentes regiões do Ceará.

1.3.2 Objetivos específicos

Este estudo tem como objetivos específicos:

- Selecionar os postos pluviométricos de referência e auxiliares a serem estudados nas diferentes regiões de análise, a partir das séries históricas disponível;
- Aplicar os métodos clássicos de preenchimento de falhas em séries pluviométricas, avaliando sua utilização com e sem a consideração da sazonalidade mensal;
- Avaliar o desempenho dos métodos de preenchimento de falhas por meio de métricas estatísticas de erro;
- Comparar os resultados obtidos entre as regiões analisadas e diferentes abordagens temporais, de modo a determinar os métodos de preenchimento mais adequados.

1.4 Roteiro de trabalho

Este trabalho é estruturado em cinco seções principais. A primeira seção apresenta o contexto e a relevância da aplicação de métodos de preenchimento de falhas em séries de precipitação no Ceará, juntamente com a definição dos objetivos gerais e específicos da pesquisa. A segunda seção reúne o referencial teórico essencial do estudo, abordando métodos

estatísticos de imputação, regressões lineares, técnicas baseadas em razões normalizadas e ponderações por correlação.

A terceira seção descreve a metodologia adotada, incluindo a seleção e o pré-processamento dos dados, a implementação computacional dos métodos avaliados e o procedimento de validação. Na quarta seção, são apresentados e discutidos os resultados obtidos a partir das análises globais, sazonais, mensais e regionais. Por fim, a quinta e última seção sintetiza as conclusões do estudo e destaca suas principais contribuições e implicações.

2. REFERENCIAL TEÓRICO

2.1 Séries históricas e o problema de falhas

2.1.1 Importância de séries longas e contínuas

A análise climática de longo prazo exige a disponibilidade de séries históricas contínuas e consistentes, especialmente em regiões com grande variabilidade climática, como o Ceará. De acordo com Silva et al. (2025, p. 534), “a normal climatológica escolhida para representar o período mais recente (1981-2010) se deve ao fato da ampla ausência de dados existentes na última série temporal (1991-2020), o que inviabiliza a análise espacial da região Nordeste.” Isso reforça a importância de uma base de dados consistente e sem falhas. No mesmo, os autores identificaram uma redução de 8% na média de precipitação da região no período mais recente, além de um aumento de 1°C na temperatura (Silva et al., 2025).

Além de possibilitarem uma análise robusta da variabilidade climática, as séries históricas contínuas são indispensáveis para a identificação de tendências de longo prazo, como alteração dos regimes pluviométricos e variações de temperatura. De acordo com a Organização Mundial de Meteorologia (OMM), os estudos climáticos devem considerar séries mínimas de 30 anos, denominadas normais climatológicas, a fim de garantir representatividade estatística e confiabilidade.

2.1.2 Causas e consequências das falhas

As falhas representam um desafio significativo para análises consistentes e confiáveis. Segundo Diaz et al. (2018), essas falhas podem acontecer tanto em coletas convencionais, que são realizadas manualmente por um operador, quanto em estações automáticas, devido a erros de operação ou sobreposição de dados. A ocorrência dessas lacunas compromete de maneira direta a integralidade das séries temporais.

Conforme Vieira et al. (2018), a presença de dados faltantes pode ocasionar resultados tendenciosos ou equivocados, afetando negativamente as análises climáticas. Séries incompletas dificultam a identificação de padrões temporais e comprometem a confiabilidade de estudos técnicos e científicos que dependem da continuidade e da representatividade dos dados.

2.1.3 Necessidade de métodos de imputação

A existência de falhas em séries de dados pluviométricos representa um obstáculo para estudos hidrológicos, especialmente aqueles que incluem análises estatísticas, modelagem e avaliações de longo prazo. Silva et al. (2020) ressaltam que séries descontínuas dificultam o ajuste de distribuições estatísticas aos dados históricos, diminuem a eficácia dos modelos hidrológicos e podem até tornar certas aplicações inviáveis. Nesse contexto, os autores destacam a importância do preenchimento de falhas como etapa fundamental para assegurar a continuidade temporal das séries e a confiabilidade dos resultados obtidos em estudos climatológicos e hidrológicos.

A continuidade das séries históricas é um requisito metodológico essencial para a realização de análises hidrológicas e climatológicas confiáveis. Segundo Vieira et al. (2018), somente séries históricas longas e temporalmente contínuas possibilitam a identificação de mudanças no comportamento das variáveis hidrometeorológicas, pois a existência de falhas compromete análises estatísticas, a identificação de tendências e a comparação entre diferentes períodos. Os autores também destacam que é comum encontrar falhas nos dados das bacias hidrográficas brasileiras, o que torna a simples exclusão de dados uma solução inadequada. Nesse cenário, a utilização de métodos de imputação de dados torna-se uma abordagem metodológica fundamental para a preservar a integridade temporal das séries e assegurar a confiabilidade dos resultados.

Considerando essas limitações, os métodos de imputação de dados são amplamente aceitos como uma opção eficaz para o tratamento de falhas em séries pluviométricas. Santos et al. (2023), ao analisarem a imputação de dados diários de precipitação no Nordeste brasileiro, mostram que a reconstrução de valores faltantes usando métodos estatísticos adequados permitem manter a estrutura temporal e sazonal das séries, além de ajudar na criação de uma base de dados mais robusta. De forma semelhante, Oliveira et al (2010), em um estudo comparativo realizado no estado de Goiás, destacam que a estimativa de valores ausentes constitui uma etapa essencial para prevenir a perda de informação e garantir a continuidade e a confiabilidade das séries históricas de precipitação.

Portanto, a implementação de métodos de imputação é essencial para garantir a continuidade e a confiabilidade das séries de dados pluviométricos utilizados neste estudo. A literatura oferece uma variedade de métodos para o preenchimento de falhas, que vão desde abordagens empíricas de ponderação até técnicas estatísticas baseadas em regressões.

2.2 Clima e precipitação no Ceará

O clima no Nordeste do Brasil é significativamente influenciado por sistemas atmosféricos de grande escala, como os anticiclones subtropicais do Atlântico. De acordo com Kayano e Andreoli (2009), dois anticiclones localizados, um no hemisfério norte (Anticiclone Subtropical do Atlântico Norte - ASAN) e outro no hemisfério sul (Anticiclone Subtropical do Atlântico Sul - ASAS) atuam principalmente nessa dinâmica. Entre esses sistemas, forma-se o cavado equatorial, cuja sua posição e intensidade variam sazonalmente, controlando o clima na região.

No eixo cavado equatorial localiza-se a Zona de Convergência Intertropical (ZCIT) que, correlacionada com a Zona de Confluência dos Alísios, exerce grande influência para o regime de chuvas na porção norte do Nordeste Brasileiro.

Rodrigues et al. (2021) e Reboita e Santos (2015) realizaram uma análise abrangente dos efeitos de El Niño e La Niña no Ceará durante a estação chuvosa, constatando que eventos de El Niño normalmente correspondem à redução das precipitações, enquanto a La Niña tende a aumentar os volumes pluviométricos nos trimestres Fevereiro-Março-Abril (FMA) e Março-Abril-Maio (MAM). Os autores destacam que, embora o ENOS seja um importante determinante da variabilidade das chuvas, outros fatores, como as condições térmicas no Atlântico, também exercem papel relevante na climatologia do Estado.

2.3 Métodos de preenchimento de falhas

2.3.1 Método da ponderação regional (PR)

Uma das abordagens mais convencionais usadas para preencher falhas em séries históricas de precipitação é o método da Ponderação Regional, que é frequentemente utilizado em análises hidrológicas com dados mensais e anuais. O método baseia-se na suposição de que estações pluviométricas situadas em uma mesma região climática exibem padrões de precipitação semelhantes, o que permite a estimativa de dados ausentes a partir das estações vizinhas.

Segundo Oliveira et al. (2010), o método se fundamenta na relação proporcional entre os totais observados nas estações vizinhas e as respectivas médias mensais históricas, assumindo que essas médias representam de maneira adequada o padrão de precipitação local. Assim, o valor faltante em uma estação pode ser estimado combinando os dados disponíveis nas estações próximas, ponderadas pelas suas médias históricas.

A Equação 1 pode ser usada para expressar a formulação do método da Ponderação Regional.

$$D_x = \frac{1}{n} \sum_{i=1}^n \left(\frac{M_x}{M_i} \right) D_i \quad (1)$$

Em que D_x representa o valor mensal de precipitação a ser estimado na estação com falha, D_i corresponde ao valor observado na estação vizinha de ordem i , M_x é a média mensal histórica da estação com falha, M_i é a média mensal histórica da estação vizinha e n é o número de estações utilizadas no processo de estimativa.

De acordo com Ruezzene et al (2019), o método da Ponderação Regional destaca-se pela sua simplicidade conceitual e pela facilidade de uso, sendo comumente utilizado quando há limitações na disponibilidade de dados ou na qualidade das séries que impedem a aplicação de métodos mais complexos. No entanto, os autores destacam que o desempenho do método depende diretamente do nível de correlação espacial entre as estações e da homogeneidade climática da área em estudo.

Estudos comparativos sugerem que, apesar de o método da Ponderação Regional oferecer resultados satisfatórios em várias aplicações, ele tende a suavizar a variabilidade dos dados estimados, o que pode limitar sua capacidade de representar eventos de precipitação extremos. Bier e Ferraz (2017) notaram que, para séries mensais de precipitação o desempenho do método pode ser comparável ao de outras técnicas tradicionais, não sendo possível apontar uma metodologia universalmente superior, pois os resultados variam de acordo com as características regionais e a qualidade dos dados disponíveis.

Com base na discussão apresentada por Bier e Ferraz (2017), o método da Ponderação Regional permanece como uma opção para o preenchimento de falhas em séries pluviométricas, sobretudo em aplicações que demandam simplicidade metodológica e consistência estatística, desde que suas limitações sejam consideradas na interpretação dos resultados.

2.3.2 Método da regressão linear múltipla (RLM)

Baseada no método dos Mínimos Quadrados Ordinários (Ordinary Least Squares - OLS), a Regressão Linear Múltipla pode ser utilizada para preencher falhas em séries históricas de precipitação, pois permite o uso simultâneo de dados de várias estações

pluviométricas. De acordo com Ruezzeno et al. (2019), essa metodologia é comumente utilizada em pesquisas hidrológicas, pois permite uma representação mais precisa da variabilidade espacial da precipitação, principalmente quando há dados simultâneos disponíveis em estações próximas.

O método da Regressão Linear Múltipla pode ser descrito por meio da Equação 2.

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + e \quad (2)$$

Em que Y representa o valor de precipitação a ser estimado na estação com falha, X_1 , X_2 , X_3 correspondem aos valores observados nas estações vizinhas utilizadas como variáveis explicativas, β_0 é o termo independente do modelo, β_1 , β_2 , β_3 são os coeficientes associados a cada variável explicativa e e representa o termo de erro do modelo.

Segundo Oliveira et al. (2010), a regressão linear múltipla geralmente oferece um desempenho melhor do que métodos mais simples de preenchimento de falhas, como a ponderação regional, especialmente em áreas onde a rede pluviométrica possui uma boa densidade espacial e uma alta correlação entre as estações.

No entanto, Bier e Ferraz (2017) destacam que o uso da regressão linear múltipla requer pressupostos estatísticos mais exigentes, sendo seu desempenho diretamente ligado à qualidade das séries, à estabilidade das relações lineares e à presença de períodos comuns de observação entre as estações. Além disso, o aumento do número de variáveis explicativas nem sempre resulta em uma melhoria significativa no desempenho do modelo, podendo introduzir redundâncias quando as estações têm comportamento muito correlacionado.

2.3.3 Método da razão normal (MRN)

Conforme descrito por Ruezzeno et al. (2019), o Método da Razão Normal é uma técnica de preenchimento de falhas em séries pluviométricas baseada na ponderação direta dos valores observados em estações vizinhas. O objetivo é ajustar as discrepâncias sistemáticas entre os padrões de precipitação das séries envolvidas. Ao contrário dos métodos de regressão, o MRN não ajusta um modelo estatístico de forma explícita, sendo considerado um método empírico de ponderação aprimorada.

Segundo Young (1992 apud Ruezzeno et al., 2019), a estimativa do dado faltante é obtida por meio de uma média ponderada, expressa pela Equação 3.

$$D_x = \frac{\sum_{i=1}^n D_i W_i}{\sum_{i=1}^n W_i} \quad (3)$$

Em que D_x representa o dado mensal faltante a ser estimado na estação com falha, D_i corresponde ao valor observado na estação vizinha de ordem i no mês de ocorrência da falha e W_i é o peso ponderativo atribuído a cada estação vizinha, descrito pela Equação 4.

$$W_i = r_i^2 \left(\frac{n_i - 2}{1 - r_i^2} \right) \quad (4)$$

Em que r_i corresponde ao coeficiente de correlação entre a estação teste e a estação vizinha de ordem i , e n_i representa o número de meses com sobreposição de dados entre as duas séries.

Bier e Ferraz (2017) destacam que o MRN demonstra um desempenho satisfatório em aplicações que envolvem séries mensais de precipitação, especialmente quando comparado à ponderação regional simples. Entretanto, os autores ressaltam que o método ainda é sensível à escolha das estações vizinhas e pode ter limitações na representação de eventos extremos, pois não inclui mecanismos explícitos de ajuste estatístico.

Assim, o Método da Razão Normal se estabelece como uma técnica intermediária entre a ponderação regional e os métodos de regressão, sendo indicado para estudos que demandam simplicidade de aplicação, ao mesmo tempo em que buscam aprimorar a qualidade das estimativas por meio da ponderação baseada na similaridade entre séries.

2.3.4 Método da ponderação por correlação (PRBRL)

Segundo Oliveira et al. (2010), a Ponderação Regional com Base em Regressões Lineares (PRBRL) é um método híbrido que combina os fundamentos da ponderação regional com as relações estatísticas derivadas de regressões lineares entre a estação com falha e suas estações vizinhas. A proposta do método é integrar dados de correlação estatística ao processo de ponderação, buscando aprimorar a estimativa dos valores ausentes em séries pluviométricas.

O método se fundamenta no ajuste de regressões lineares simples entre a estação com

falha e cada uma das estações vizinhas, usando os coeficientes de correlação como pesos no processo de estimativa. Assim, o PRBRL se distingue da ponderação regional tradicional ao levar em conta de forma explícita a intensidade da relação estatística entre as séries, sem utilizar um modelo multivariado único, como acontece na regressão linear múltipla.

A estimativa do valor mensal faltante pelo método PRBRL pode ser expressa pela Equação (5):

$$D_x = \frac{\sum_{i=1}^n r_{D_x D_i} D_i}{\sum_{i=1}^n} \quad (5)$$

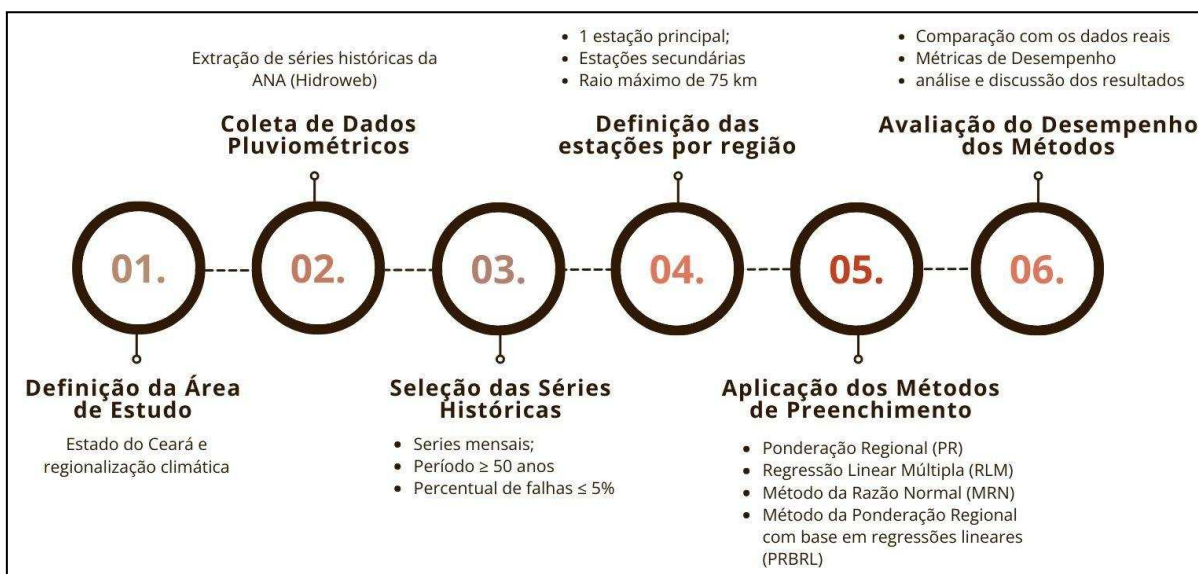
Oliveira et al. (2010), em estudos comparativos, indicam que o desempenho do PRBRL pode ser aceitável em algumas aplicações, mas não necessariamente melhor do que métodos mais simples, como a regressão linear múltipla ou a ponderação regional. Isso depende das características da rede pluviométrica e da disponibilidade de dados simultâneos. Além disso, é importante destacar que o método é mais sensível à qualidade das correlações estimadas, o que pode restringir sua utilização em áreas com séries temporais curtas ou com baixa correlação.

Dessa forma, a Ponderação Regional com Base em Regressões Lineares se apresenta como uma opção intermediária entre os métodos de ponderação e as abordagens estatísticas mais elaboradas, sendo apropriada para pesquisas exploratórias ou comparativas, desde que suas restrições sejam levadas em conta ao interpretar os resultados.

3. METODOLOGIA

A metodologia que será adotada neste estudo tem como objetivo avaliar o desempenho de diferentes métodos de preenchimentos de falhas em séries pluviométricas no estado do Ceará, sendo exemplificado por meio da figura 1, a seguir:

Figura 1 - Fluxograma da metodologia



Fonte: Autoria Própria (2025).

3.1 Região de estudo

Este estudo foi desenvolvido a partir de séries históricas de precipitação provenientes de estações pluviométricas localizadas no estado do Ceará, na região Nordeste do Brasil. O estado do Ceará apresenta elevada variabilidade espacial e temporal da precipitação, associada principalmente à atuação de diferentes sistemas atmosféricos como a Zona de Convergência Intertropical, o fenômeno El Niño-Oscilação Sul e as anomalias da temperatura da superfície do mar do Atlântico Tropical, refletindo padrões climáticos característicos do semiárido nordestino (SILVA et al., 2023).

Para levar em consideração a diversidade espacial da precipitação no estado, as estações pluviométricas estudadas foram organizadas em dez áreas com diferentes características climáticas e geográficas. Essa opção possibilita uma representação mais precisa dos variados regimes de precipitação observados no Ceará, minimizando possíveis distorções resultantes da análise combinada de áreas com comportamentos climáticos distintos.

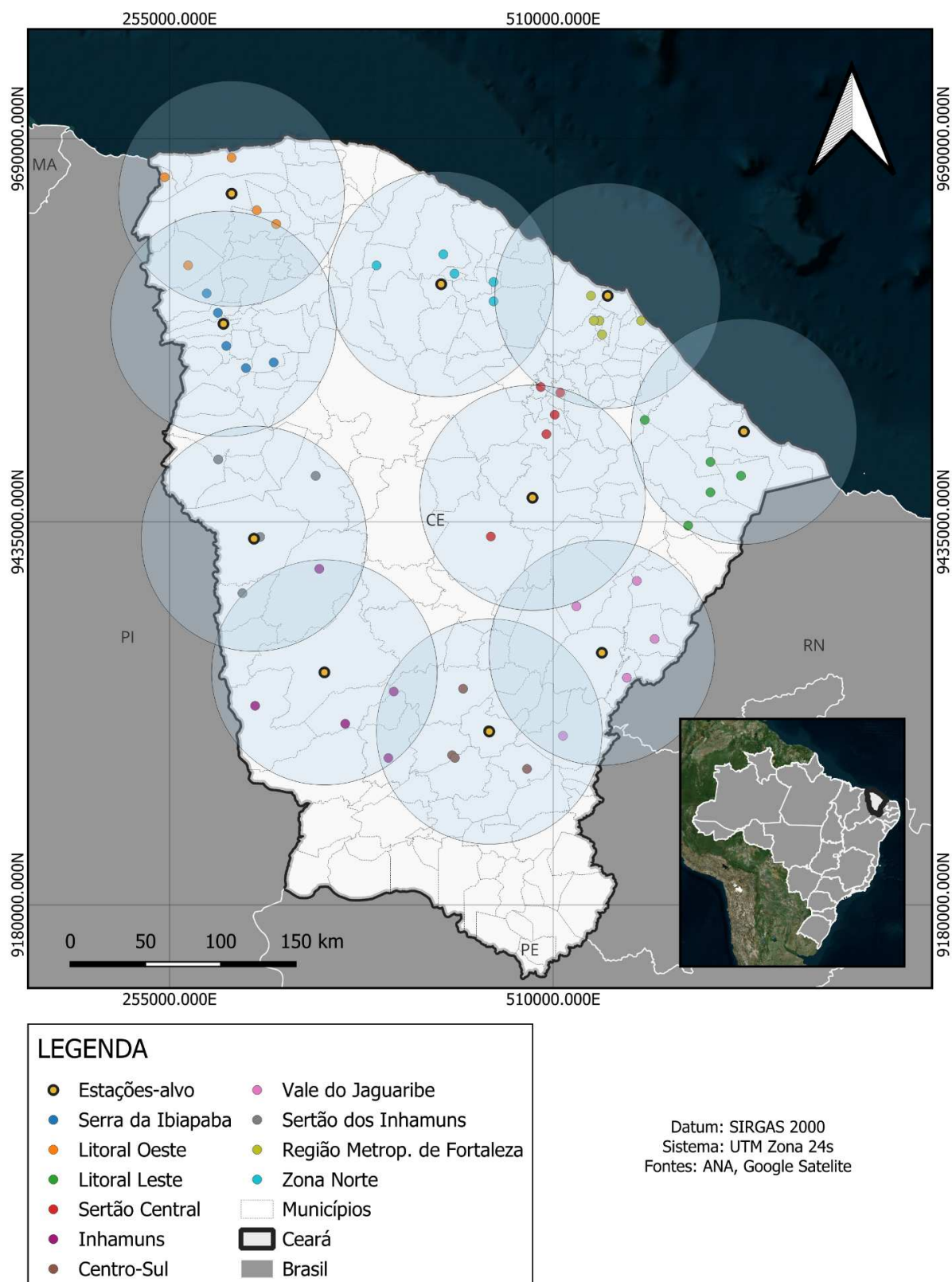
Para a análise espacial, as estações foram organizadas em 10 regiões estabelecidas especificamente para este estudo. Embora existam regionalizações de referência para o estado

do Ceará, como as regiões pluviometricamente homogêneas propostas por Xavier (2007) e adotadas pela FUNCEME, neste caso utilizou-se um critério operacional fundamentado na distribuição das estações e em um raio de 75km, com o objetivo de assegurar a cobertura do estado e possibilitar a comparação entre distintos padrões de precipitação.

As regiões analisadas foram: Litoral Leste, Litoral Oeste, Serra da Ibiapaba, Zona Norte, Sertão Central, Inhamuns, Centro-sul, Vale do Jaguaribe, Sertão dos Inhamuns, Região Metropolitana de Fortaleza. Essa divisão foi estabelecida para analisar o desempenho dos métodos de preenchimento de falhas em relação a diferentes padrões espaciais de precipitação, permitindo uma comparação mais sólida e confiável.

A Figura 2 apresenta o mapa de localização das estações pluviométricas escolhidas, bem como a definição das regiões levadas em consideração no estudo. A distribuição geográfica das estações mostra a extensão geográfica da base de dados usada, permitindo a análise dos métodos de imputação em diferentes contextos climáticos no estado do Ceará.

Figura 2 - Mapa de localização das estações pluviométricas selecionadas.



Fonte: Autoria Própria (2025).

3.2 Base de dados

3.2.1 Coleta de dados pluviométricos

Os dados utilizados neste estudo foram obtidos a partir da plataforma HidroWeb, mantida pela Agência Nacional de Águas e Saneamento Básico (ANA). Essa plataforma oferece séries históricas de precipitação observadas em postos pluviométricos espalhados por todo o Brasil, sendo umas das principais fontes oficiais de dados hidrometeorológicos do país.

Foram selecionadas séries mensais com aproximadamente 50 anos de registros, de modo a garantir uma base temporal ampla para avaliar os métodos de preenchimento de falhas. Para cada uma das dez regiões definidas no estudo, foram consideradas uma estação pluviométrica principal, e estações pluviométricas secundárias, utilizadas como suporte nos procedimentos de preenchimentos dos dados ausentes.

Como critério extra de seleção, as estações secundárias foram escolhidas considerando um raio máximo de 75 km em relação à estação principal de cada região. Esse critério espacial foi adotado com o objetivo de garantir maior similaridade climática entre as estações utilizadas, dado que a precipitação exibe uma alta variabilidade espacial. No total, foram examinadas 60 estações pluviométricas, com seis estações vinculadas a cada região.

A Tabela 1 exibe os principais aspectos da base de dados empregada, incluindo o período de análise, código da estação principal e identificação das estações secundárias em cada área.

Tabela 1 - Caracterização da base de dados pluviométricos utilizados no estudo

Região	Estação principal (código ANA)	Estações secundárias (código ANA)	Período analisado
Litoral Leste	437000	437006, 437009, 437010, 538009, 438022	1961-2021
Litoral Oeste	340017	240002, 341010, 340008, 341018, 340039	1961-2021
Serra da Ibiapaba	340018	440000, 440030, 340031, 340030, 440023	1961-2021
Zona Norte	339037	339078, 339084, 339004, 339034, 339023	1961-2021
Sertão Central	439001	438032, 539012, 438053, 438010, 439008	1961-2021
Inhamuns	640002	640015, 640003, 639013, 540114, 539005	1961-2021
Centro-sul	639084	639044, 639030, 639026, 639002, 639021	1961-2021
Vale do Jaguaribe	538029	638093, 638011, 538030, 538066, 538008	1961-2021
Sertão dos Inhamuns	540020	540114, 540003, 540019, 440018, 440007	1961-2021
Região Metropolitana de Fortaleza	338001	338009, 338007, 338005, 338016, 338002	1961-2021

Fonte: Autoria Própria (2025)

3.2.2 Critérios de seleção e consistência dos dados

Para cada uma das regiões definidas no estudo, foram escolhidos postos pluviométricos que demonstraram maior consistência e integridade dos dados ao longo do período analisado. O objetivo dessa etapa foi assegurar a consistência das séries históricas empregadas, reduzindo ao máximo o impacto de falhas prolongadas ou descontinuidades que pudessem afetar a implementação e análise dos métodos de preenchimento de falhas.

Como critério de qualidade, consideraram-se apenas estações com percentual de falhas igual ou inferior a 5% do total da série histórica, garantindo que os dados observados representassem adequadamente o comportamento pluviométrico de cada região. As séries foram organizadas na escala temporal mensal, compatível com os métodos adotados e com o

objetivo do estudo.

A organização da base de dados envolveu a padronização das séries temporais, a identificação dos períodos com dados ausentes e a separação das estações em principais e secundárias, conforme descrito na seção anterior. Essa estrutura permitiu a aplicação sistemática dos métodos de imputação.

A Tabela 2 exhibe os critérios utilizados para a seleção e organização das estações pluviométricas, bem como as características mais relevantes levadas em conta na formação da base de dados.

Tabela 2 - Critérios adotados para seleção das estações pluviométricas

Critério	Descrição
Período mínimo	50 anos
Percentual máximo de falhas	$\leq 5\%$
Escala temporal	Mensal
Fonte	HidroWeb/ANA

Fonte: Autoria Própria (2025)

3.3 Métodos de preenchimento de falhas avaliados

Foram avaliados métodos clássicos de preenchimento de falhas em séries temporais de precipitação, amplamente utilizados na literatura hidrológica, todos baseados na utilização de informações vindas de estações pluviométricas vizinhas. A escolha desses métodos teve como objetivo comparar diferentes formas de exploração das relações estatísticas entre as séries, mantendo-se constante a base de dados e os critérios de seleção das estações auxiliares.

Em todos os métodos que utilizam dados de estações vizinhas, o preenchimento das falhas foi realizado levando em conta apenas as estações que forneceram informações no mês correspondente ao valor ausente da estação-alvo. Além disso, as estimativas foram realizadas apenas quando havia pelo menos três estações vizinhas com registros disponíveis neste mesmo mês, adotando-se esse critério como condição mínima para a execução, de modo a garantir a robustez estatística das estimativas. Nos casos em que esse critério mínimo não foi atendido, os valores ausentes permaneceram sem estimativa.

No método da Ponderação Regional (PR), as estimativas foram obtidas a partir das médias ponderadas das precipitações observadas nas estações secundárias. Os fatores de ponderação foram definidos com base na relação entre a média histórica da estação-alvo e as

médias históricas das estações vizinhas. Esse procedimento foi adotado com o objetivo de minimizar a influência das diferentes sistemáticas entre as estações.

No método da Regressão Linear Múltipla (RLM), os modelos foram ajustados utilizando exclusivamente períodos em que havia dados disponíveis simultaneamente na estação-alvo quanto nas estações secundárias consideradas em cada estimativa. O modelo foi recalculado a cada etapa do processo de validação cruzada, de forma que o valor a ser estimado não fosse incluído no conjunto de treinamento, garantindo a independência entre os dados utilizados para ajuste e aqueles usados para avaliar o desempenho do método.

No método da Razão Normal (MRN), as estimativas foram calculadas com base nas médias ponderadas das precipitações registradas nas estações secundárias. Os pesos foram determinados a partir dos coeficientes de correlação linear entre cada estação vizinha e a estação-alvo, além do tamanho da sobreposição amostral entre as séries no conjunto de treinamento. Dessa forma, foram atribuídos maiores pesos às estações que apresentaram maior similaridade estatística e maior número de observações simultâneas com a série a ser preenchida, aumentando a confiabilidade das estimativas.

De maneira semelhante, no método da Ponderação Regional com Base em Regressões Lineares (PRBRL), a estimativa dos valores ausentes foi feita com base em médias ponderadas das precipitações nas estações vizinhas, considerando pesos diretamente proporcionais ao coeficiente de correlação linear entre as séries da estação-alvo e das estações secundárias. Foram consideradas apenas correlações positivas e somente estações que apresentaram, no mínimo, três ocorrências simultâneas com a estação-alvo no conjunto de treinamento.

Os procedimentos de preenchimento de falhas e de validação cruzada foram implementados por meio de rotinas computacionais desenvolvidas em linguagem Python, utilizando as bibliotecas *pandas* (McKinney, 2010) e *NumPy* (Harris et al., 2020).

3.4 Implementação computacional e estratégia de validação

Os métodos de preenchimento de falhas descritos na Seção 3.3 foram implementados em linguagem Python, por meio de rotinas computacionais criadas para este estudo. A implementação foi feita de forma modular, permitindo a aplicação consistente dos métodos a diferentes regiões, além de favorecer e reprodutibilidade dos resultados.

A validação cruzada do tipo Leave-One-Time-Out (LOTO) foi utilizada para avaliar o desempenho dos métodos. Nesse processo, cada observação mensal da estação-alvo é removida individualmente da série e estimada com base nos dados restantes.

A implementação abrangeu dois cenários de análise diferentes, determinados pela consideração ou não da sazonalidade das chuvas. Esses cenários diferem quanto à composição do conjunto de treinamento utilizado em cada estimativa e estão resumidos na Tabela 3.

Tabela 3 - Cenários considerados na implementação dos métodos de preenchimento

Cenário	Descrição	Conjunto de treinamento
Sem sazonalidade	Série completa	Todos os meses disponíveis
Com sazonalidade	Análise mensal	Apenas meses equivalentes

Fonte: Autoria Própria (2025)

No caso em que não há sazonalidade, todas as observações da série histórica foram usadas simultaneamente no processo de treinamento. No caso de sazonalidade, as estimativas foram feitas de maneira estratificada por mês, de forma que cada observação fosse calculada usando apenas os dados históricos referentes ao mesmo mês do ano. Essa metodologia possibilitou a análise do impacto da sazonalidade no desempenho dos métodos de preenchimento.

3.5 Métricas de avaliação de desempenho

O desempenho dos métodos de preenchimento de falhas foi avaliado usando métricas estatísticas frequentemente aplicadas na literatura para a quantificação de erros em séries hidrológicas. Essas métricas possibilitam a avaliação tanto da magnitude média dos erros quanto da presença de padrões sistemáticos nas estimativas.

Foram utilizadas as métricas MAE (Mean Absolute Error), RMSE (Root Mean Square Error) e Bias (viés). o MAE foi usado para mensurar o erro médio absoluto entre os valores observados e estimados, enquanto o RMSE foi adotado por sua maior sensibilidade a erros de grande magnitude. O Bias foi levado em conta para identificar possíveis tendências de superestimação ou subestimação associadas aos métodos avaliados. A Tabela 4 exibe a interpretação geral dessas métricas.

Tabela 4 - Métricas utilizadas na avaliação do desempenho dos métodos

Métrica	Descrição	Interpretação
MAE	Erro médio absoluto entre valores observados e estimados	Valores menores indicam melhor desempenho
RMSE	Raiz do erro quadrático médio	Valores menores indicam melhor desempenho, sendo mais sensível a erros de maior magnitude
Bias	Diferença média entre valores estimados e observados	Valores próximos de zero indicam ausência de viés

Fonte: Autoria Própria (2025)

As métricas foram determinadas com base nos resultados da validação cruzada Leave-One-Time-Out (LOTO), levando em conta apenas os períodos em que foi possível estimar os valores ausentes. Isso garantiu uma avaliação consistente e comparável do desempenho dos métodos em todos os cenários analisados.

Para o cálculo das métricas, considere n pares de valores, em que O_i representa o valor observado e E_i o valor estimado para o instante i . Assim, o Erro Médio Absoluto (MAE) foi calculado conforme a Equação (6):

$$MAE = \frac{1}{n} \sum_{i=1}^n |E_i - O_i| \quad (6)$$

A Raiz do Erro Quadrático Médio (RMSE) foi determinada conforme a Equação (7):

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (E_i - O_i)^2} \quad (7)$$

Por fim, o viés (Bias) foi calculado conforme a Equação (8):

$$Bias = \frac{1}{n} \sum_{i=1}^n (E_i - O_i) \quad (8)$$

4. RESULTADOS E DISCUSSÕES

4.1 Apresentação geral dos resultados

Esta seção apresenta os resultados obtidos a partir da aplicação dos métodos de preenchimento de falhas em séries históricas de precipitação no Ceará. As análises foram feitas considerando diferentes cenários, incluindo avaliações globais, mensais e regionais, bem como abordagens como e sem consideração da sazonalidade pluviométrica.

Inicialmente, os resultados são apresentados de forma agregada, com o objetivo de fornecer uma visão geral do desempenho dos métodos avaliados. Em seguida, são discutidas análises mais detalhadas, contemplando a variabilidade mensal, a distinção entre períodos úmidos e secos e as diferenças observadas entre as regiões analisadas.

4.2 Avaliação global dos métodos (sem consideração da sazonalidade)

Os resultados obtidos a partir da aplicação dos métodos de preenchimento de falhas sem consideração da sazonalidade pluviométrica são apresentados a seguir. Nesse cenário, todos os dados disponíveis das séries históricas foram utilizados conjuntamente no processo de treinamento, independente do mês do ano, conforme descrito na metodologia.

O desempenho dos métodos foi avaliado com base nas métricas MAE, RMSE e Bias, obtidas por meio da validação cruzada do tipo Leave-One-Time-Out (LOTO), permitindo uma comparação agregada do comportamento médio das abordagens ao longo da série histórica.

A Tabela 5 apresenta os valores médios globais das métricas de desempenho para cada método avaliado.

Tabela 5 - Métricas globais de desempenho dos métodos de preenchimento de falhas no cenário sem consideração da sazonalidade

Método	MAE médio	RMSE médio	Bias médio
RLM	23,40	41,86	0,40
PR	27,22	48,26	0,04
MRN	26,78	47,52	-3,63
PRBRL	29,85	52,73	-2,36

Fonte: Autoria Própria (2025)

Nota-se que o método de regressão linear múltipla (RLM) exibiu os menores valores médios de MAE e RMSE, sugerindo uma maior precisão geral na previsão dos valores de precipitação em comparação com os outros métodos.

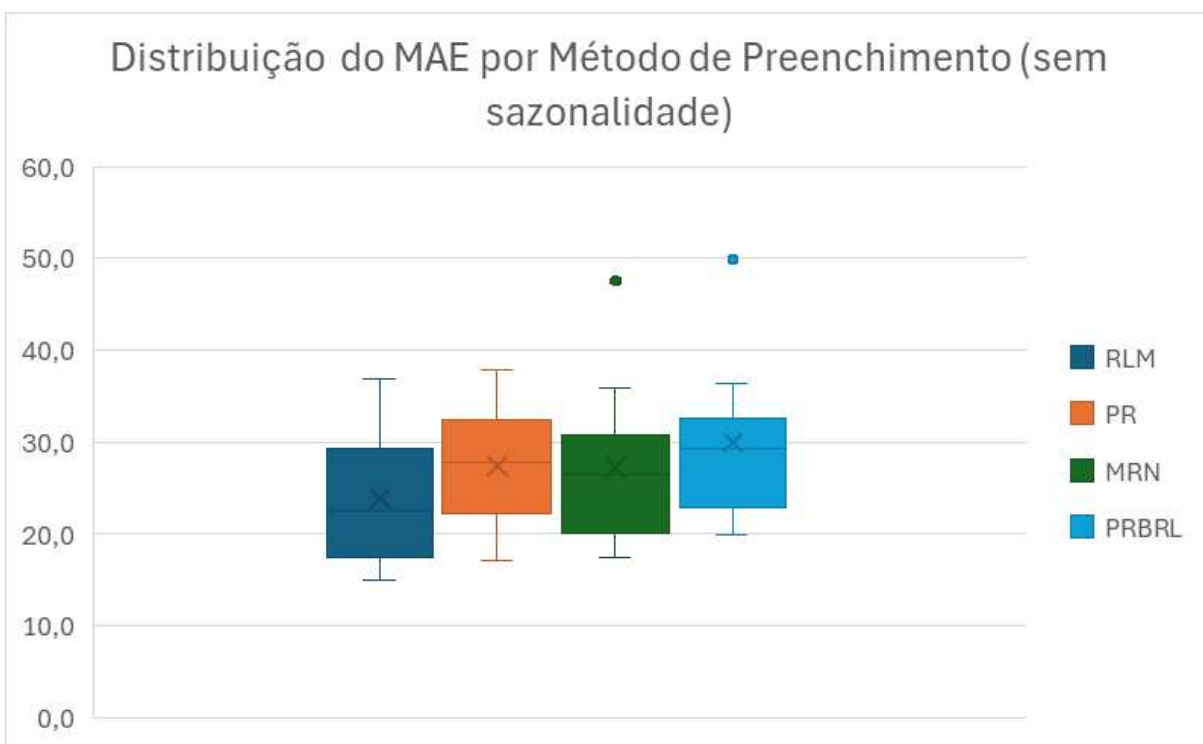
Os métodos da razão normal (MRN) e da ponderação regional (PR) exibiram desempenhos medianos, com valores médios de erro bastante semelhantes. O método PRBRL exibiu os maiores valores médios de MAE e RMSE, indicando um desempenho geral inferior quando utilizado sem estratificação sazonal.

No que diz respeito ao Bias médio, os valores obtidos para os métodos RLM e PR estiveram próximos de zero, o que indica que não houve uma tendência sistemática significativa de superestimação ou subestimação. Em contrapartida, os métodos MRN e PRBRL exibiram valores negativos de viés, indicando uma tendência de subestimar as precipitações estimadas.

A Figura 3, que mostra a distribuição dos valores de MAE por método e por meio de boxplots, evidencia a variabilidade do desempenho dos métodos em diferentes regiões. Nota-se que o método RLM exibe uma mediana e uma dispersão menores dos valores, o que demonstra uma maior consistência no desempenho ao longo das regiões estudadas.

Os métodos PR e MRN exibem um comportamento intermediário, com distribuições comparáveis. Por outro lado, o método PRBRL mostra maior dispersão e valores mais altos de MAE, sugerindo uma sensibilidade maior às variações regionais das séries pluviométricas.

Figura 3 - Distribuição dos valores de MAE por método de preenchimento no cenário sem sazonalidade



Fonte: Autoria Própria (2025).

Como complemento à avaliação global, a Figura 4 apresenta uma matriz de ranking com a classificação relativa dos métodos de preenchimento de falhas por região, levando em conta o cenário sem sazonalidade. Essa representação resume a estabilidade do desempenho dos métodos no espaço, destacando padrões regionais que serão abordados com mais profundidade na análise regional apresentada a seguir.

Figura 4 - Ranking relativo dos métodos de preenchimento por região, baseado no MAE no cenário sem sazonalidade

Região	RLM	PR	MRN	PRBRL
Litoral Leste	2	1	4	3
Litoral Oeste	1	4	2	3
Serra da Ibiapaba	1	2	3	4
Zona Norte	1	3	2	4
Sertão Central	1	3	2	4
Inhamus	2	1	4	3
Centro Sul	1	4	2	3
Vale do Jaguaribe	1	3	2	4
Sertão dos inhamus	1	4	2	3
Região Metrop. De Fort.	1	2	3	4

Fonte: Autoria Própria (2025).

Os resultados globais apresentados fornecem um ponto de partida para a avaliação levando em conta a sazonalidade mensal. Isso possibilita investigar, nas seções seguintes, se a estratificação temporal ajuda a melhorar o desempenho dos métodos de preenchimento de falhas.

4.3 Avaliação com consideração da sazonalidade mensal

4.3.1 Apresentação geral da análise sazonal

A análise levando em conta a sazonalidade mensal foi realizada para examinar como a variabilidade sazonal da precipitação afeta o desempenho dos métodos de preenchimento de falhas. Ao contrário da abordagem global descrita anteriormente, as estimativas foram feitas de maneira estratificada por mês, de forma que cada valor ausente fosse calculado usando apenas os dados históricos referentes ao mesmo mês do ano.

Essa abordagem possibilita a avaliação de como levar em conta a sazonalidade pode aprimorar as estimativas, particularmente em áreas com alta variabilidade de precipitação ao longo do ano. Os resultados são primeiramente expostos de maneira minuciosa por mês e, em seguida, resumidos em relação aos períodos secos e úmidos, a fim de tornar mais fácil a interpretação e a comparação com o cenário sem sazonalidade.

4.3.2 Desempenho mensal dos métodos

A abordagem com sazonalidade mensal é representada no mapa de calor dos valores de MAE para cada método de preenchimento ao longo dos meses do ano, conforme ilustrado na Figura 5. Nota-se que a performance dos métodos oscila consideravelmente ao longo do ano, evidenciando a influência direta do padrão de precipitação sobre a precisão das estimativas.

Figura 5 - Valores mensais de MAE por método de preenchimento, considerando a sazonalidade.

Mês	MRN	RLM	PR	PRBRL	MÉDIA GERAL
JANEIRO	37,65	35,65	39,06	41,89	38,56
FEVEREIRO	46,11	43,18	47,82	50,46	46,89
MARÇO	63,05	58,37	67,38	69,87	64,67
ABRIL	59,10	52,21	62,00	65,81	59,78
MAIO	38,54	35,17	39,42	42,45	38,89
JUNHO	22,75	20,36	21,84	24,53	22,37
JULHO	14,35	13,44	14,12	15,24	14,29
AGOSTO	4,89	5,22	5,40	5,18	5,17
SETEMBRO	4,24	4,85	5,42	4,38	4,72
OUTUBRO	3,78	4,15	5,19	4,13	4,32
NOVEMBRO	5,79	6,08	6,32	5,84	6,01
DEZEMBRO	16,29	16,75	16,06	16,90	16,50

Fonte: Autoria própria (2025)

Em geral, os meses relacionados ao período de chuvas, que se caracteriza por maior variabilidade e ocorrência de eventos de precipitação mais intensos, concentram os maiores valores de erro. Durante esses meses, as variações de desempenho entre os métodos se tornam mais claras, revelando uma maior sensibilidade das abordagens às condições de precipitação mais complexas. Em particular, observa-se um crescimento nos erros médios de fevereiro a abril, período em que ocorrem os maiores volumes de chuva no estado.

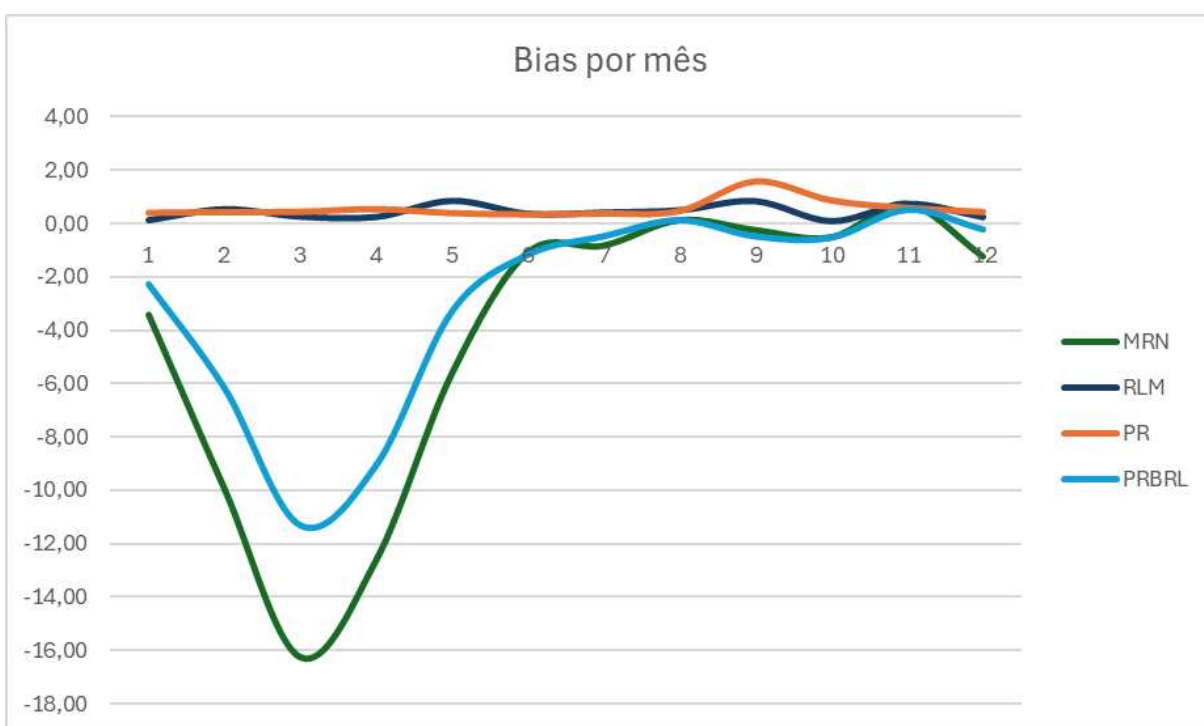
Por outro lado, nos meses correspondentes ao período seco, os valores de MAE são significativamente menores e apresentam comportamento mais semelhante entre os métodos avaliados. Esse padrão sugere maior estabilidade das estimativas em condições de menor variabilidade pluviométrica, independentemente do método utilizado.

A análise do comportamento sazonal é complementada pela Figura 6, que apresenta a variação mensal do Bias médio por método de preenchimento. Observa-se que os métodos RLM e PR mantêm valores de Bias próximo de zero tanto nos meses associados ao período

chuvoso quanto nos meses do período seco, indicando ausência de tendência de superestimação ou subestimação ao longo do ano.

Enquanto isso, os métodos MRN e PRBRL apresentam viés negativo durante o período chuvoso, indicando uma tendência de subestimação das precipitações em condições de maior complexidade. Nos meses que correspondem ao período seco, os valores de Bias aproximam-se de zero, evidenciando redução da tendência de subestimação e maior estabilidade.

Figura 6 - Perfil sazonal do viés (média mensal) por método de preenchimento com consideração da sazonalidade



Fonte: Autoria própria (2025)

4.4 Comparação entre abordagens com e sem consideração da sazonalidade

A análise dos cenários com e sem consideração da sazonalidade mensal foi conduzida por meio da identificação do método mais adequado para cada região, além dos valores correspondentes de MAE, RMSE e Bias. Essa análise possibilita a avaliação da eficácia da estratificação mensal na melhoria do desempenho dos métodos de preenchimento de falhas.

A Tabela 6 apresenta, para cada região analisada, o método mais eficaz identificado em cada cenário, os respectivos valores de MAE e a diferença entre eles, expressa por meio do indicador Δ MAE, definido com a diferença entre o MAE obtido com sazonalidade mensal

e o MAE obtido no cenário sem sazonalidade. Valores positivos de Δ MAE indicam piora no desempenho com a inclusão da sazonalidade, enquanto valores negativos indicam melhoria.

Tabela 6 - Comparação do desempenho dos métodos de preenchimento por região nos cenários com e sem consideração da sazonalidade, com base no MAE

Região	Melhor método (sem sazonalidade)	MAE (sem sazonalidade) (mm)	Melhor método (com sazonalidade)	MAE (com sazonalidade) (mm)	ΔMAE (mm)
Litoral Leste	PR	27,44	PR	27,61	0,17
Litoral Oeste	RLM	25,22	RLM	24,71	-0,51
Serra da Ibiapaba	RLM	36,92	RLM	38,01	1,09
Zona Norte	RLM	23,01	RLM	22,83	-0,18
Sertão Central	RLM	18,52	RLM	19,54	1,02
Inhamuns	PR	17,17	PR	17,43	0,26
Centro-Sul	RLM	15	MRN	15,71	0,71
Vale do Jaguaribe	RLM	17,07	RLM	18,36	1,29
Sertão dos Inhamuns	RLM	19,95	MRN	20,11	0,16
Região Metropolitana de Fortaleza	RLM	32,63	PR	34,63	2

Fonte: Autoria própria (2025)

Os resultados mostram que, na maioria das regiões analisadas, a consideração da sazonalidade mensal não resultou em redução do erro médio absoluto. Em oito das dez regiões avaliadas, os valores de Δ MAE foram positivos, indicando que a abordagem mensal apresentou desempenho inferior ao cenário global. As melhorias percebidas com a consideração da sazonalidade foram limitadas, afetando apenas duas regiões, e mostraram-se de magnitude reduzida.

Em relação ao Bias, os valores mantiveram-se, de modo geral, próximos de zero em ambos os casos, sugerindo que a incorporação da sazonalidade mensal não teve um impacto significativo na tendência de superestimação ou subestimação das estimativas. Dessa forma,

as diferenças observadas entre as abordagens estão associadas predominantemente à magnitude dos erros, e não à introdução de viés sistemático.

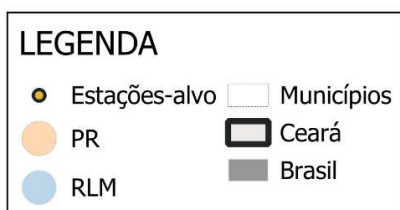
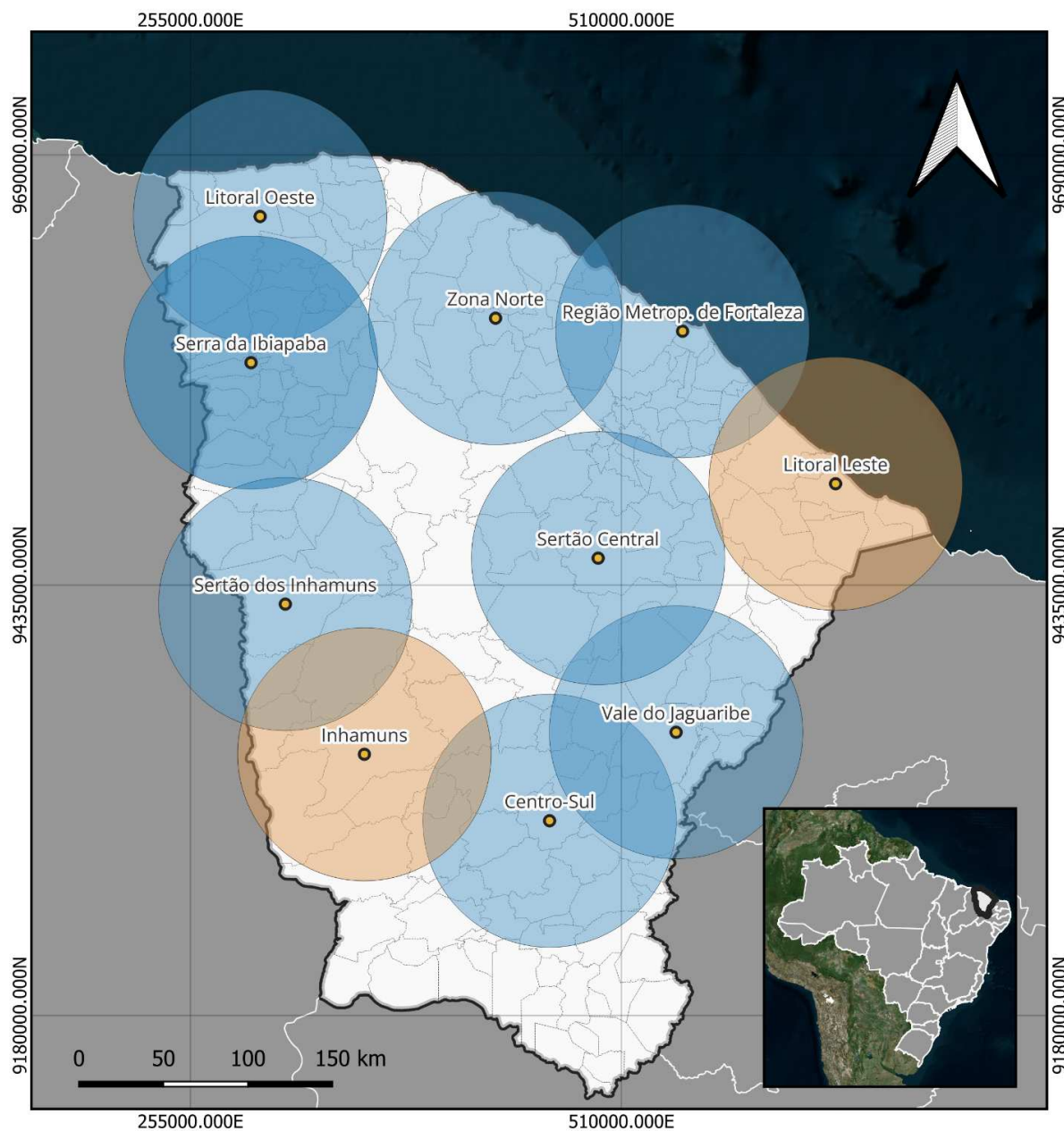
4.5 Análise regional dos resultados

A análise regional dos resultados possibilita a avaliação da variabilidade no desempenho dos métodos de preenchimento de falhas, considerando as particularidades climáticas e espaciais das regiões analisadas no Ceará. As diferenças observadas entre as regiões refletem a influência do regime pluviométrico, da variabilidade espacial da precipitação e da disponibilidade de estações vizinhas com séries históricas consistentes.

A Figura 7 apresenta o desempenho relativo dos métodos de preenchimento por região, com base no MAE, evidenciando que o método de regressão linear múltipla (RLM) apresentou desempenho superior na maioria das regiões analisadas. Esse padrão sugere que, em áreas onde existe maior correlação espacial entre as séries pluviométricas, métodos baseados em relações lineares costumam gerar estimativas mais precisas.

Os resultados sugerem que, na maioria das regiões analisadas, os métodos avaliados tiveram desempenhos semelhantes, com diferenças entre as métricas de erro sendo relativamente pequenas. Apesar de o método de regressão linear múltipla (RLM) ter mostrado um desempenho superior na maioria das regiões, os ganhos em comparação com os outros métodos foram, em geral, modestos. Esse comportamento indica que a implementação de métodos mais complexos nem sempre leva a melhorias significativas. Em muitos casos, é mais vantajoso usar abordagens mais simples e eficientes em termos computacionais, especialmente quando elas oferecem desempenho semelhante.

Figura 7 - Distribuição espacial do método de preenchimento com melhor desempenho por região.



Datum: SIRGAS 2000
 Sistema: UTM Zona 24s
 Fontes: ANA, Google Satellite

5. CONCLUSÃO

O objetivo deste estudo foi comparar métodos tradicionais de preenchimento de falhas em séries históricas de precipitação no Ceará, considerando a aplicação direta das séries e a separação sazonal mensal. Os resultados indicaram que a sazonalidade mensal altera o desempenho dos métodos, porém não gerou ganhos consistentes, o que sugere que sua adoção deve ser avaliada caso a caso.

Entre as técnicas analisadas, os métodos baseados em regressão linear apresentaram, em geral, os melhores desempenhos, com destaque para a Regressão Linear Múltipla (RLM). Ainda assim, em diversas situações, as diferenças em relação aos demais métodos foram pequenas, indicando que abordagens mais complexas nem sempre resultam em melhorias relevantes na estimativa de valores ausentes.

Conclui-se que a escolha do método deve considerar não apenas o desempenho estatístico, mas também aspectos práticos, como disponibilidade de dados, facilidade de implementação e custo computacional. Como trabalhos futuros, recomenda-se avaliar outras escalas temporais e comparar técnicas geoestatísticas e métodos baseados em aprendizado de máquina.

REFERÊNCIAS

AGÊNCIA NACIONAL DE ÁGUAS E SANEAMENTO BÁSICO (ANA). HidroWeb: Sistema de Informações Hidrológicas. Brasília: ANA, 2025. Disponível em: <https://www.snirh.gov.br/hidroweb>. Acesso em: 16 jul. 2025.

AYOADE, J. O. **Introdução à climatologia para os trópicos**. 12. ed. Rio de Janeiro: Bertrand Brasil, 2010.

BIER, A. A.; FERRAZ, S. E. T. **Comparação de metodologias de preenchimento de falhas em dados meteorológicos para estações no Sul do Brasil**. Revista Brasileira de Meteorologia, v. 32, n. 2, p. 215–226, 2017. DOI: 10.1590/0102-77863220008.

CHOW, V. T.; MAIDMENT, D. R.; MAYS, L. W. **Applied hydrology**. New York: McGraw-Hill, 1988.

COLLISCHONN, W.; DORNELLES, F. **Hidrologia para engenharia e ciências ambientais**. 2. ed. rev. e ampl. Porto Alegre: ABRH, 2015.

DIAZ, C. C. F.; PEREIRA, J. A. S.; NÓBREGA, R. S. **Comparação de dados estimados por dois diferentes métodos para o preenchimento de falhas de precipitação pluvial na bacia do rio Pajeú, Pernambuco, Brasil**. Revista Brasileira de Climatologia, v. 22, p. 324–339, 2018.

HARRIS, Charles R. et al. Array programming with NumPy. *Nature*, v. 585, p. 357-362, 2020.

KAYANO, M. T.; ANDREOLI, R. V. Clima da Região Nordeste do Brasil. In: CAVALCANTI, I. F. A. et al. (org.). **Tempo e clima no Brasil**. São Paulo: Oficina de Textos, 2009. p. 179–193.

MCKINNEY, Wes. Data structures for statistical computing in Python. In: *PROCEEDINGS OF THE 9TH PYTHON IN SCIENCE CONFERENCE (SCIPY)*. 2010. p. 56-61.

OLIVEIRA, L. F. C. de; FIOREZE, A. P.; MEDEIROS, A. M. M.; SILVA, M. A. S. **Comparação de metodologias de preenchimento de falhas de séries históricas de precipitação pluvial anual**. Revista Brasileira de Engenharia Agrícola e Ambiental, v. 14, n. 11, p. 1186–1192, 2010.

ORGANIZAÇÃO METEOROLÓGICA MUNDIAL. **Guide to hydrological practices: hydrology – from measurement to hydrological information**. 6. ed. Geneva: World Meteorological Organization, 2008. (WMO-No. 168).

REBOITA, M. S.; SANTOS, I. **Influência de alguns padrões de teleconexão na precipitação no norte e nordeste do Brasil**. Revista Brasileira de Climatologia, v. 15, p. 106–122, 2015.

RODRIGUES, B. D. et al. **Uma análise sobre as chuvas no Ceará baseada nos eventos de El Niño, La Niña e no Dipolo do Servain durante a estação chuvosa**. Revista Brasileira de Climatologia, v. 28, p. 507–519, 2021. Disponível em: <https://ojs.ufgd.edu.br/rbclima/article/view/14723>. Acesso em: 16 jul. 2025.

RUEZZENE, C. B.; MIRANDA, R. B. de; MAUAD, F. F. **Revisão bibliográfica dos principais métodos de preenchimento de falhas em dados de precipitação**. In: ENCONTRO DE SUSTENTABILIDADE EM PROJETO (ENSUS), 7., 2019, Florianópolis. *Anais [...]* Florianópolis: UFSC, 2019. p. 153–164.

SANTOS, L. F. et al. **Imputation of precipitation data in Northeast Brazil**. Anais da Academia Brasileira de Ciências, 2023.

SILVA, A. D. G. et al. **Mudanças climáticas e variabilidade termopluiométrica na região Nordeste do Brasil**. Revista Brasileira de Climatologia, v. 36, n. 21, p. 533–559, 2025. DOI: 10.55761/abclima.v36i21.18695. Disponível em: <https://ojs.ufgd.edu.br/rbclima/article/view/18695>. Acesso em: 16 jul. 2025.

SILVA, A. F. et al. **Preenchimento de falhas e espacialização de dados pluviométricos: desafios e perspectivas**. Revista Brasileira de Meteorologia, v. 35, n. 4, 2020.

SILVA, M. D. P. B. da; ALCÓCER, J. C. A.; SILVA, M. L. da; SANDI, N. P. M.; ALFARO, E. J.; HIDALGO, H. G. **Variabilidade climática da precipitação no estado do Ceará, Nordeste do Brasil**. Revista Brasileira de Climatologia, v. 33, p. 492–512, jul./dez. 2023.

TUCCI, C. E. M. **Hidrologia: ciência e aplicação**. 4. ed. Porto Alegre: Editora da Universidade, UFRGS; ABRH, 2009.

VIEIRA, S. et al. **Metodologia de imputação de dados hidrometeorológicos para análise de séries históricas – Bacia do Rio dos Sinos, RS, Brasil.** Revista Brasileira de Climatologia, v. 23, p. 189–204, 2018.

WANDERLEY, H. S.; AMORIM, R. F. C. de; CARVALHO, F. O. de. **Interpolação espacial de dados médios mensais pluviométricos com redes neurais artificiais.** Revista Brasileira de Meteorologia, v. 29, n. 3, p. 389–396, 2014.

XAVIER, T. de M. B. S; XAVIER, A. F. S; ALVES, J. M. B. **Quantis e eventos extremos: aplicações em ciências da terra e ambientais.** Fortaleza: RDS, 2007. 278 p.