



I CONGRESSO PERNAMBUCANO DE RECURSOS HÍDRICOS

Água para o Desenvolvimento
Recife, 24, 25 e 26 de Março de 2026

IDENTIFICAÇÃO DE ZONAS HOMOGÊNEAS DE PRECIPITAÇÃO MENSAL EM PERNAMBUCO USANDO CLUSTERIZAÇÃO E ANÁLISE DE COMPONENTES PRINCIPAIS

Juliana Farias Santos de Moraes¹ ; Camila Oliveira de Britto Salgueiro² ; João Maria de Andrade³ ; Alfredo Ribeiro Neto⁴ ; Sylvana Melo dos Santos⁵ & Leidjane Maria Maciel de Oliveira⁶.

Palavras-chave: Machine Learning; Zona homogênea; Precipitação.

INTRODUÇÃO

Entender as condições pluviométricas de determinado local é um requisito essencial para o planejamento territorial, a gestão de recursos hídricos e a mitigação de riscos associados a eventos climáticos extremos. O estado de Pernambuco apresenta uma acentuada e heterogênea distribuição de chuvas, sendo que as mesorregiões da Região Metropolitana do Recife e Mata Pernambucana se destacam por apresentarem as maiores taxas de precipitação anual. A região litorânea, onde essas mesorregiões se localizam, é marcada pelo clima tropical e pelas menores altitudes do estado, registrando alturas de chuva acumuladas anuais na ordem de 1.000 a 2.200 mm, conforme Costa Neto *et al.* (2025).

Para a avaliação das condições pluviométricas do estado, pode-se considerar o zoneamento, que consiste no agrupamento de áreas geográficas com padrões de precipitação estatisticamente semelhantes. No entanto, em relação à discretização dessas áreas, a simples aplicação de métodos tradicionais, baseados em fatores puramente geográficos ou domínios climáticos definidos, apresenta limitações na indicação da variabilidade espacial das séries históricas de chuva (Miranda, Negri e Pampuch, 2022).

Este trabalho visa o desenvolvimento de um zoneamento pluviométrico para o estado de Pernambuco, aplicando técnicas de clusterização, uma técnica de aprendizado não supervisionado, que permite o agrupamento baseado na similaridade entre os elementos, revelando sub-regiões homogêneas e padrões de precipitação relevantes.

METODOLOGIA

Área de Estudo

O estado de Pernambuco, de acordo com o Plano Estadual de Recursos Hídricos (PERH/PE - PERNAMBUCO, 2022), é marcado pela irregularidade espaço-temporal na distribuição de chuvas, condição intrinsecamente ligada à variação climática que distingue a região litorânea, dominada pelo clima tropical, do interior, onde prevalece o

¹) Doutoranda do Programa de Pós-Graduação em Engenharia Civil - PPGEC, Universidade Federal de Pernambuco - UFPE. Rua Acadêmico Hélio Ramos, s/n, Cidade Universitária. 50.740-530. Recife/PE – Brasil. Fone: +55 81 2126.8223. E-mail: juliana.fmoraes@ufpe.br.

²) Doutoranda - PPGEC, UFPE, Recife-PE. E-mail: camila.salgueiro@ufpe.br.

³) Pós-Doutorando - PPGEC, UFPE, Recife-PE. E-mail: joao.jma@ufpe.br.

⁴) Professor Titular do Departamento de Engenharia Civil e Ambiental - DECIV, docente permanente - PPGEC, UFPE, Recife-PE. E-mail: alfredo.ribeiro@ufpe.br.

⁵) Professora Titular - DECIV, docente permanente - PPGEC, UFPE, Recife-PE. E-mail: sylvana.santos@ufpe.br.

⁶) Professora Adjunta - DECIV, docente permanente - PPGEC, UFPE, Recife-PE. E-mail: leidjane.oliveira@ufpe.br.

Clima Semiárido. A escolha do estado se justifica pela sua heterogeneidade pluviométrica, caracterizada pela transição abrupta do clima Tropical úmido (Zona da Mata e Região Metropolitana) para o Semiárido (Sertão), passando pelo Agreste de transição, situação que resulta em desafios ao planejamento territorial e à gestão dos recursos hídricos. Nesse contexto, a aplicação de técnicas de clusterização serve como uma ferramenta de identificação das regiões hidrologicamente homogêneas, fornecendo parâmetros técnicos para a administração eficiente da região frente às suas diferentes condições pluviométricas.

Levantamento dos pluviômetros de referência

Para a obtenção dos dados de precipitação na extensão do estado de Pernambuco, foram consideradas séries históricas do Portal *HidroWeb* da Agência Nacional de Águas e Saneamento Básico (ANA, 2024), integrante do Sistema Nacional de Informações sobre Recursos Hídricos (SNIRH), além das estações operadas pela Agência Pernambucana de Águas e Clima (APAC). Após a compilação inicial, iniciou-se o processo de filtragem dos dados, no qual foram considerados aspectos ligados à situação operacional, extensão temporal e qualidade de dados, sendo, assim, utilizados postos que apresentam percentual máximo de falhas diárias de 2,5%. Diante disso, foram analisadas 46 estações pluviométricas abrangendo toda a extensão do estado. Ressalta-se que, embora seja possível identificar áreas com menor densidade de monitoramento, a distribuição final das 46 estações representa a melhor cobertura espacial alcançada a partir dos critérios de qualidade e extensão temporal aplicados.

Aplicação do Machine Learning

O zoneamento pluviométrico foi desenvolvido com o auxílio de técnicas de Aprendizado de Máquina Não Supervisionado, utilizando a Análise de Componentes Principais (PCA), para a identificação dos atributos físico-climáticos que mais influenciam na variabilidade da precipitação no estado, e a Clusterização com algoritmo K-Means, para agrupar as estações pluviométricas em regiões hidrologicamente homogêneas com base na similaridade de seus valores, processados no *software* RStudio, versão 2025.09.1.

Foram utilizados os registros pluviométricos das 46 estações selecionadas, inicialmente obtidos para dados diários, depois agrupados em séries mensais e, em seguida, foram derivados: precipitação média mensal, precipitação acumulada total, precipitação máxima, desvio padrão de precipitação, quantidade de meses com chuva e sem chuva e altitude, a fim de caracterizar o comportamento espacial da chuva no estado.

Análise de Componentes Principais (PCA)

Visando a redução da dimensionalidade do conjunto de dados, foi inicialmente empregada a Análise de Componentes Principais (PCA), procedimento que visa identificar a relação entre os atributos calculados e determinar os fatores que explicam a maior parte da variância da precipitação. A utilização desse método é justificada porque a precipitação total, isoladamente, não seria suficiente para representar a variabilidade climática de Pernambuco. A matriz de dados de entrada para a PCA foi enriquecida com múltiplos atributos derivados das séries históricas das 46 estações pluviométricas, contribuindo para que a PCA gere componentes que captem a complexa variabilidade climática do estado, como o gradiente de umidade e a sazonalidade.

Clusterização com K-Means

Foi realizada a aplicação do método de clusterização *K-Means*, adotado como principal algoritmo de agrupamento, com o objetivo de dividir as 46 estações

pluviométricas em K *clusters* pré-definidos. Diante disso, cada grupo é formado de modo que as estações apresentem maior similaridade entre si do que em relação às estações pertencentes aos demais *clusters*.

Para a definição do número de *clusters* (K) utilizou-se o Método do Cotovelo (*Elbow Method*), procedimento que calcula a Soma dos Quadrados *Intra-Cluster* (WSS) para diferentes valores de K , métrica que quantifica a compactação interna dos agrupamentos.

RESULTADOS

Análise de Componentes Principais (PCA)

A Análise de Componentes Principais (PCA) indicou que o componente PC1 representa aproximadamente 63% da variância total dos dados utilizados (Tabela 1).

Tabela 1 – Importância das variáveis.

Atributos	PC1	PC2	PC3	PC4	PC5	PC6	PC7
Desvio Padrão	2.097	1.2130	0.8825	0.58120	0.12637	2.33e-16	9.706e-17
Variância	0.628	0.2102	0.1113	0.04826	0.00228	0.00e+00	1.00e+00
Variância acumulada	0.628	0.8382	0.9495	0.99772	1.000	0.00e+00	1.00e+00

Fonte: Autores.

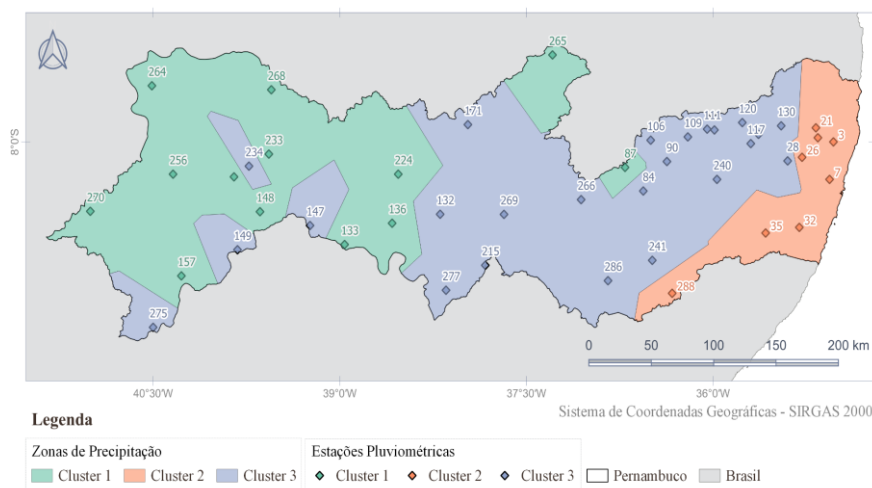
De acordo com a Matriz de Cargas, a componente PC1 tem como atributos de maior influência a precipitação total, a média mensal e precipitação máxima, além dos meses sem chuva, sugerindo ser uma componente de gradiente de umidade, caracterizada pela variação seca-chuva. Sendo assim, a PCA indicou que o gradiente de umidade explica aproximadamente 63% da variância dos dados de precipitação no estado de Pernambuco, separando as estações secas e úmidas. A componente PC2, que explicou 21% da variância dos dados analisados, teve como atributos de maior influência os meses com chuva, precipitação média, soma da precipitação e máxima precipitação, fazendo uma oposição entre a duração da estação chuvosa e a intensidade da chuva. Já a componente PC3 teve como predominância o atributo altitude, representando um eixo de topografia que explica 11% da variância dos dados.

A partir dos estudos, foi perceptível que a altitude é o atributo com menor importância na formação dos eixos, além de apresentar-se como atributo com baixa correlação com o volume total de chuva nas estações estudadas. Essa configuração estatística sugere que a variabilidade pluviométrica em Pernambuco é governada primordialmente pela distância do oceano e pela dinâmica de sistemas atmosféricos de grande escala, e não apenas por efeitos topográficos.

Clusterização com K-Means

A clusterização feita com o *K-Means* para $K = 3$, obtida nos estudos, resultou em três zonas homogêneas de precipitação, com espacialização coerente (Figura 1) ao ser comparada com os dados de precipitação anual média. Assim, foi possível identificar claramente a zona mais úmida, localizada próxima ao litoral (*Cluster 2*), com precipitação anual média variando de 950 a 2150 mm; a zona de transição entre áreas úmidas e secas (*Cluster 3*), com precipitação anual média variando de 290 a 927 mm; e a zona mais seca do estado (*Cluster 1*), com precipitação anual média variando de 210 a 835 mm.

Figura 1 – Clusterização das estações pluviométricas adotadas.



Fonte: Autores.

CONCLUSÕES

A aplicação combinada dos métodos PCA e K-Means permitiu identificar três zonas pluviometricamente homogêneas no estado de Pernambuco, apresentando resultados coerentes com a variabilidade climática da região. Essas zonas mostraram-se bem definidas, correspondendo a uma zona mais úmida, próxima ao litoral, uma zona de transição com precipitação moderada e uma zona mais seca, com baixos índices pluviométricos. Sendo assim, este estudo demonstrou o potencial de aplicação de técnicas de *Machine Learning* como ferramentas de suporte à gestão dos Recursos Hídricos.

AGRADECIMENTOS - Os autores agradecem à Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) pelo fomento à pesquisa e apoio financeiro por meio da concessão de bolsas de doutorado da primeira autora (Processo nº 88887.151009/2025-00) e da segunda autora (Processo nº 88887.134583/2025-00).

REFERÊNCIAS

- ANA - Agência Nacional de Águas e Saneamento Básico. **Séries históricas das estações**. Disponível em: <http://www.snirh.gov.br/hidroweb/serieshistoricas>. Acesso em: 18/12/2025.
- COSTA NETO, F. A. P. W.; SALGUEIRO, C. O. B.; MENEZES, R. B. G. G.; SANTOS, S. M.; OLIVEIRA, L. M. M.. Análise de concordância dos dados de precipitação estimados por sensoriamento remoto em Mesorregiões do estado de Pernambuco – Brasil. **Revista Brasileira de Climatologia**, 36 (21), 452–477, 2025. <https://doi.org/10.55761/abclima.v36i21.19499>
- MIRANDA, B. G.; NEGRI, R. G.; P., L. A.. Using clustering algorithms and GPM data to identify spatial precipitation patterns over southeastern Brazil. **Atmosfera**, 37, 365–381, 2023. <https://doi.org/10.20937/ATM.53155>
- PERNAMBUCO. **Secretaria de Infraestrutura e Recursos Hídricos**. *Plano Estadual de Recursos Hídricos de Pernambuco: tomo I: diagnósticos*. Recife: SEINFRA, 2022.