CLUSTERIZAÇÃO DE NASCENTES DA BACIA DO MAMANGUAPE: MODELAGEM COMPUTACIONAL PARA CONSERVAÇÃO HÍDRICA

Isis Gomes de Medeiros [1]
Jonas Otaviano Praça de Souza [2]
Daniel Baptista Vio [3]

RESUMO

O presente trabalho teve como objetivo desenvolver um modelo de clusterização para agrupar nascentes da Bacia do Mamanguape, na Paraíba, conforme características ambientais. A pesquisa fundamenta-se na importância das nascentes para a dinâmica hidrológica e para a manutenção dos recursos hídricos, considerando a influência de fatores geológicos, geomorfológicos, climáticos e antrópicos sobre esses ambientes. Para isso, foram utilizadas sete variáveis ambientais: altitude, declividade, curvatura vertical, TWI (Topographic Wetness Index), precipitação, vegetação e geologia. Os dados foram obtidos por meio de imagens SRTM, informações pluviométricas da AESA, dados do MapBiomas e bases geológicas da SUDEMA, sendo processados em ambiente SIG utilizando os softwares ArcMap e QGIS. Posteriormente, aplicou-se o algoritmo K-Means em linguagem Python, após a normalização das variáveis ambientais, permitindo a identificação de quatro agrupamentos distintos de nascentes. Os resultados evidenciaram padrões espaciais e ambientais associados principalmente ao relevo, à cobertura vegetal e à compartimentação geológica da Bacia do Mamanguape. Os grupos localizados sobre o Planalto da Borborema apresentaram maiores altitudes e relevo mais dissecado, enquanto o agrupamento associado ao Tabuleiro Costeiro destacou-se pelos maiores índices de umidade do terreno. Além disso, observou-se que os baixos percentuais de vegetação em parte dos grupos indicam maior vulnerabilidade ambiental e influência antrópica sobre as APPs de nascente. Dessa forma, a clusterização demonstrou potencial como ferramenta de apoio à gestão ambiental e à identificação de áreas prioritárias para conservação hídrica.

Palavras-chave: Nascentes, K-Means, Conservação hídrica, Bacia do Mamanguape.

[1] Graduanda do Curso de Geografia da Universidade Federal da Paraíba – UFPB, Isisgomesdm@gmail.com. Pesquisa vinculada ao projeto de Iniciação Científica "Clusterização de Nascentes do Semiárido: Modelagem Computacional para Conservação Hídrica", financiado pelo CNPq;
[2] Doutor em Geografia pela Universidade Federal de Pernambuco - UFPE, Professor do Departamento de Geociências da Universidade Federal da Paraíba – UFPB, jonas.souza@academico.ufpb.br;
[3] Mestre em Ciência da Computação da Universidade Federal da Paraíba - UFPB, dbvio82@gmail.com.

INTRODUÇÃO

As nascentes desempenham papel fundamental na dinâmica hidrológica e na manutenção dos recursos hídricos, constituindo áreas de grande relevância ambiental, social e econômica. Esses ambientes correspondem aos locais de exfiltração natural da água subterrânea, podendo originar cursos d'água e influenciando diretamente processos relacionados ao relevo, infiltração, escoamento superficial e disponibilidade hídrica. Além disso, suas características estão intrinsecamente associadas aos aspectos geológicos, geomorfológicos e climáticos da região em que se inserem, tornando o estudo dessas áreas essencial para compreensão da dinâmica ambiental e para o planejamento de ações de conservação.

No Brasil, a preservação de nascentes é prevista pela Lei Federal nº 12.651/12, o Código Florestal Brasileiro, que estabelece essas áreas como Áreas de Preservação Permanente (APPs), definindo um raio mínimo de 50 metros ao redor das nascentes como área protegida. Entretanto, diversos fatores antrópicos, como expansão urbana, atividades agropecuárias e supressão da cobertura vegetal, comprometem a integridade ambiental dessas áreas, afetando diretamente a dinâmica hídrica e a capacidade de manutenção das nascentes.

Nesse contexto, o uso de técnicas computacionais associadas ao geoprocessamento apresenta grande potencial para análises ambientais integradas. A disponibilidade de dados geoespaciais e o avanço das técnicas de Inteligência Artificial permitem a aplicação de métodos de machine learning capazes de identificar padrões em grandes conjuntos de dados ambientais. Dentre essas abordagens, destaca-se a clusterização, que possibilita agrupar elementos com características semelhantes, contribuindo para a compreensão espacial e ambiental das nascentes.

Assim, a presente pesquisa tem como objetivo desenvolver um modelo de clusterização para agrupar nascentes da Bacia do Mamanguape conforme características ambientais e graus de degradação. Para isso, foram utilizadas variáveis relacionadas à altitude, declividade, curvatura vertical, TWI (Topographic Wetness Index), precipitação, vegetação e geologia, obtidas por meio de técnicas de geoprocessamento em ambiente SIG.

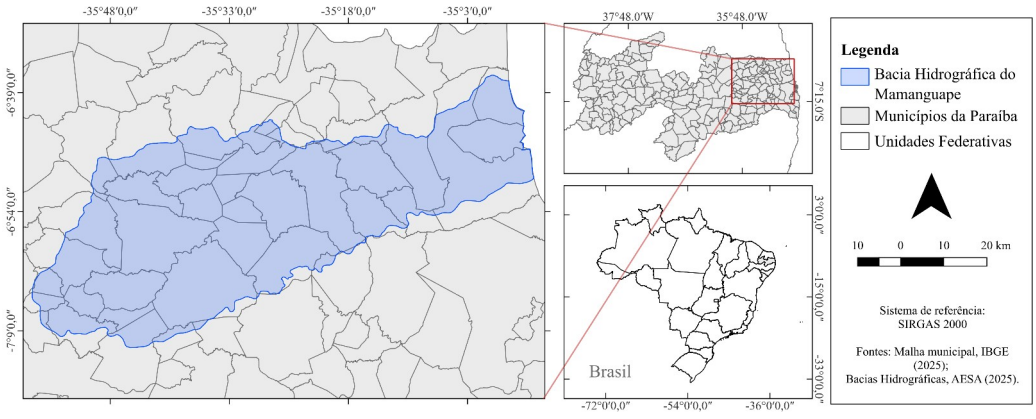
Posteriormente, aplicou-se o algoritmo K-Means para identificação de agrupamentos ambientais entre as nascentes. Os resultados evidenciaram distintos padrões espaciais e ambientais associados principalmente ao relevo, à cobertura vegetal e à compartimentação geológica da Bacia do Mamanguape, permitindo identificar áreas com diferentes níveis de conservação e vulnerabilidade ambiental.

METODOLOGIA

Área de Estudo

A área de estudo corresponde às nascentes da Bacia do Rio Mamanguape, localizada no estado da Paraíba, abrangendo cerca de 46 municípios das Regiões Imediatas de Campina Grande, Guarabira, João Pessoa e Mamanguape–Rio Tinto (Figura 1). A escolha da área deve-se à elevada concentração de nascentes em relação às demais bacias hidrográficas do estado, totalizando 687 nascentes identificadas por meio de dados do Cadastro Ambiental Rural (CAR).

Figura 1 – Mapa de localização da Bacia do Mamanguape



Fonte: Elaboração própria (2026).

Aquisição e Interpolação dos dados

Foram definidas sete variáveis ambientais consideradas relevantes para a caracterização das nascentes, sendo elas: altitude, declividade, curvatura vertical, TWI (Topographic Wetness Index), precipitação, vegetação e geologia (Quadro 1). As variáveis altitude, declividade, curvatura vertical e TWI foram obtidas a partir de imagens SRTM

(Shuttle Radar Topography Mission) com resolução espacial de 30 metros, sendo realizados procedimentos de mosaico, recorte e processamento utilizando o software ArcMap 10.5.

Quadro 1 – Variáveis ambientais utilizadas na caracterização das nascentes

Variável	Importância para análise das nascentes
Altitude	Influencia as condições climáticas locais, a dinâmica de escoamento e os padrões de drenagem, podendo condicionar a ocorrência e manutenção de nascentes.
Declividade	Controla o escoamento superficial e a infiltração da água no solo, influenciando diretamente a recarga hídrica das nascentes.
Curvatura vertical	Indica áreas de convergência ou dispersão do fluxo superficial, auxiliando na identificação de zonas favoráveis à concentração de água.
TWI (*Topographic Wetness Index*)	Representa o potencial de acumulação de umidade no terreno, sendo importante para identificar áreas com maior tendência à saturação hídrica.
Precipitação	Principal fonte de recarga hídrica, condicionando a disponibilidade de água e a permanência do fluxo nas nascentes.
Vegetação	Atua na proteção do solo, favorecendo a infiltração da água, reduzindo processos erosivos e contribuindo para a conservação ambiental.
Geologia	Controla a permeabilidade e a capacidade de armazenamento de água no subsolo, influenciando a formação e a vazão das nascentes.

Fonte: Elaboração própria (2026).

Para a variável precipitação foram coletados dados pluviométricos das estações ativas da AESA a partir da plataforma HidroWeb, utilizando aquelas estações que se encontravam dentro dos municípios ao qual a bacia do Mamanguape faz parte, com o recorte temporal entre 01/01/1996 e 31/12/2025, sendo feita a média de chuva diária de cada estação. Posteriormente, as estações foram espacializadas e transformadas em superfície contínua por interpolação espacial, utilizando o método Kriging, com resolução espacial de 30 m.

A variável vegetação foi extraída da coleção 10.1 do MapBiomas, utilizando o ano de 2024, sendo reclassificada em vegetação natural e usos antrópicos. Já a variável geologia foi obtida a partir de dados da SUDEMA, sendo utilizada como suporte analítico espacial dos agrupamentos devido ao seu caráter categórico. Posteriormente, foram gerados buffers de 50 metros para cada nascente identificada na Bacia do Mamanguape. Em seguida, foram extraídos os valores das variáveis ambientais para cada buffer utilizando o software QGIS 3.42, consolidando um banco de dados ambiental das nascentes.

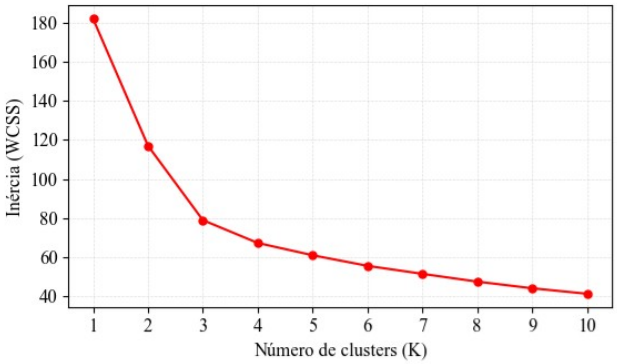
Clusterização

Para definição do número ideal de agrupamentos foi utilizado o Método do Cotovelo, por meio da análise da métrica Within-Cluster Sum of Squares (WCSS), sendo adotado K=4

por apresentar melhor adequação aos objetivos da pesquisa, conforme apresentado no Gráfico 1.

As variáveis ambientais foram normalizadas por meio da padronização Min-Max, estabelecendo intervalo entre 0 e 1 para os dados utilizados. Posteriormente, aplicou-se o algoritmo K-Means em linguagem Python utilizando as bibliotecas pandas, numpy, matplotlib e scikit-learn, permitindo a identificação de agrupamentos ambientais entre as nascentes da Bacia do Mamanguape.

Gráfico 1 – Gráfico de Inércia

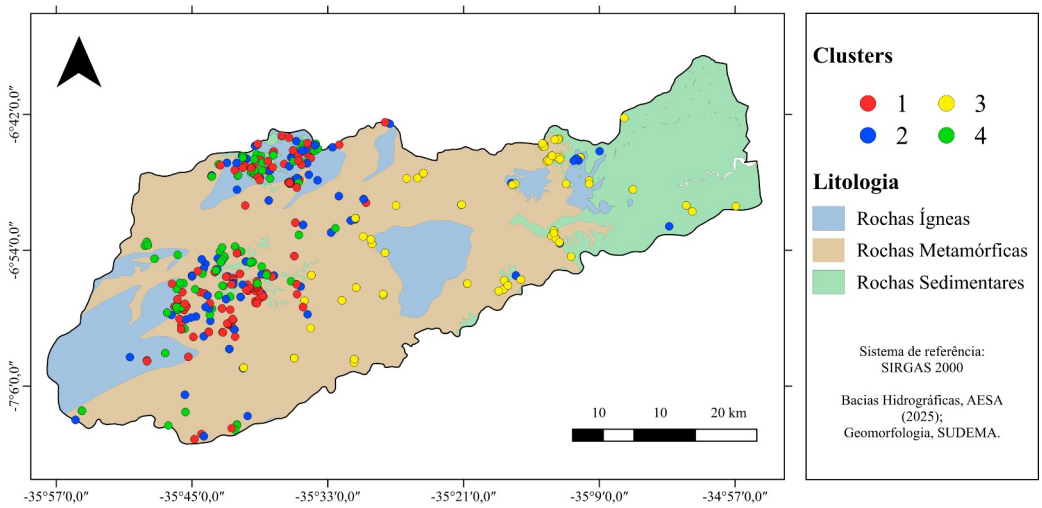


Fonte: Elaboração própria (2026).

RESULTADOS E DISCUSSÃO

A partir do Mapa de Espacialização dos Clusters é possível observar certas tendências espaciais de cada grupo, como a maior concentração a oeste da Bacia pelo grupo 1, representado por pontos vermelhos, e o grupo 4, representado por pontos verdes, enquanto o grupo com maior dispersão sobre a Bacia é o grupo 3, pontos amarelos (Figura 2). Sendo, também, o grupo com menor número de nascentes, com 124 representantes de um total de 687, apresentando os grupos 1 e 4 os maiores números de nascentes, sendo 248 e 188, respectivamente, o que pode indicar relativa homogeneidade das características ambientais nas nascentes que compõem a Bacia do Mamanguape.

Figura 2 – Espacialização dos Clusters e Litologia na Bacia do Mamanguape



Fonte: Elaboração própria (2026).

Sendo assim, o grupo 1, com o maior número de nascentes entre os agrupamentos, apresenta valores médios intermediários nas seis variáveis aplicadas à clusterização quando comparado aos demais grupos (Tabela 1), podendo ser caracterizado como um grupo de transição ambiental, associado a altitudes intermediárias e declividade moderada. Todas as suas nascentes encontram-se sobre o Planalto da Borborema, em áreas de formações geológicas ígneas ou metamórficas, inseridas em um relevo mais movimentado, marcado por maior dissecação e presença de fraturas, o que evidencia um forte condicionamento dessas nascentes à dinâmica cristalina.

Tabela 1 - Média das Variáveis Ambientais por Cluster

Cluster	Altitude (m)	Declividade (%)	Curvatura	TWI	Precipitação diária (mm)	Vegetação (%)
1	472,647	20,092	0,155	3,356	3,029	4,663
2	403,862	20,871	0,165	3,787	3,014	80,895
3	105,075	9,831	0,059	4,546	2,921	3,004
4	497,988	11,091	0,205	5,075	2,903	3,033

Fonte: Elaboração própria (2026).

Além disso, o valor médio de 4.6% de vegetação indica uma cobertura vegetal relativamente esparsa, o que pode refletir uma maior influência antrópica sobre essas APPs. Essa baixa cobertura vegetal tende a comprometer processos importantes para a manutenção

hídrica das nascentes, como a infiltração da água no solo, a proteção contra processos erosivos e a estabilidade ambiental. Ainda assim, por apresentar características intermediárias em relação às variáveis analisadas, o grupo 1 pode representar áreas de transição entre setores mais preservados e áreas mais degradadas da Bacia do Mamanguape.

O grupo 4, por sua vez, segundo grupo mais representativo, compreende as nascentes com os maiores valores médios de altitude, com cerca de 497 m, e de curvatura vertical, com média de 0,205. Assim como o grupo 1, todas as suas nascentes estão localizadas sobre o Planalto da Borborema. Seus elevados valores de altitude e curvatura indicam nascentes associadas a áreas de maior dissecação do relevo, responsáveis pela origem dos cursos fluviais de primeira ordem.

Entretanto, apesar de sua elevada importância para a recarga hídrica e manutenção da drenagem da Bacia, o grupo 4 apresenta média de apenas 3,0% de vegetação, evidenciando forte supressão da cobertura vegetal em suas APPs. Grande parte dessas nascentes encontra-se muito próxima a centros urbanos, especialmente nos municípios de Solânea e Bananeiras, ao noroeste da Bacia, o que pode indicar maior pressão antrópica sobre esses ambientes. Dessa forma, apesar de representar áreas estratégicas para a dinâmica hidrológica da Bacia do Mamanguape, o grupo 4 também se configura como um dos mais suscetíveis à degradação ambiental.

Já o grupo 2 é definido principalmente pelo seu elevado percentual médio de vegetação, com cerca de 80%, valor que se destaca significativamente em relação aos demais grupos. Além disso, apresenta o maior valor médio de declividade (20,8%), estando associado às áreas mais dissecadas do Planalto da Borborema, assim como os grupos 1 e 4. Apesar de também apresentar grande concentração de nascentes nos municípios de Solânea, Bananeiras e Areia, observa-se que suas nascentes se distribuem em setores relativamente mais afastados dos principais centros urbanos, o que pode ter contribuído para a permanência da cobertura vegetal nessas áreas.

Além disso, o relevo mais acidentado e com maiores declividades pode dificultar processos de ocupação humana e expansão agrícola intensiva, favorecendo a conservação da vegetação natural. Essa característica auxilia na infiltração da água, proteção do solo e

estabilidade hidrológica desses ambientes. Dessa forma, o grupo 2 representa o agrupamento com melhores condições de conservação ambiental entre os clusters identificados, podendo servir como referência para estratégias de preservação e manejo das nascentes da Bacia do Mamanguape.

Por fim, o grupo 3 apresenta as menores altitudes médias, com cerca de 105 m, destacando-se também pelos altos valores de TWI (Topographic Wetness Index). Suas nascentes encontram-se distribuídas principalmente na transição entre o Planalto da Borborema e a Depressão Sertaneja, com maior concentração sobre o Tabuleiro Costeiro. Nessas áreas, a presença de um relevo mais planificado contribui para maiores valores do índice de umidade, indicando maior tendência à saturação hídrica do terreno e favorecendo a permanência da umidade superficial.

Esse grupo também corresponde ao agrupamento com maior dispersão espacial das nascentes sobre a Bacia do Mamanguape, diferentemente dos grupos mais concentrados sobre o Planalto da Borborema. Em relação à cobertura vegetal, o grupo apresenta um dos menores percentuais médios entre os clusters, com cerca de 3,0%, indicando uma vegetação pouco expressiva nas áreas de entorno das nascentes. Essa baixa cobertura vegetal pode estar relacionada à intensa ocupação antrópica observada em setores do Tabuleiro Costeiro, associados principalmente à expansão agrícola e urbana. Dessa forma, mesmo apresentando condições favoráveis à saturação hídrica e manutenção da umidade superficial, a baixa presença de vegetação pode comprometer a estabilidade ambiental dessas APPs, aumentando sua suscetibilidade à degradação e aos impactos sobre a dinâmica hídrica local.

Embora a precipitação tenha apresentado diferenças relativamente discretas entre os clusters, observou-se uma tendência de maiores valores nos grupos associados ao Planalto da Borborema, refletindo a influência altimétrica sobre a distribuição das chuvas. Entretanto, o grupo 3 demonstrou que elevados índices de umidade do terreno podem ocorrer mesmo em setores com menor precipitação média, evidenciando a forte influência do relevo e da dinâmica de saturação hídrica

CONSIDERAÇÕES FINAIS

O presente trabalho demonstrou que as nascentes da Bacia do Mamanguape apresentam distintos padrões espaciais e ambientais, relacionados a características geológicas e ambientais da região. Com os grupos associados ao Planalto da Borborema, sendo maioria e representados pelos grupos 1, 2 e 4, apresentando maiores valores de altitude e relevos dissecados associados a formações ígneas e metamórficas, indicando nascentes condicionadas à estrutura cristalina. Em contrapartida, o grupo 3 associado ao Tabuleiro Costeiro apresentou um alto valor de umidade do terreno, associado ao relevo mais suavizado e com tendência à saturação hídrica.

Além disso, o grupo 2 destacou-se pelo elevado percentual de vegetação, indicando melhores condições ambientais para manutenção das APPs de nascente, enquanto os demais grupos apresentaram baixos percentuais de cobertura vegetal, sugerindo maior vulnerabilidade ambiental, destacando áreas preservadas e áreas sob maior pressão antrópica. Em especial, o grupo 4 demonstrou elevada importância para a dinâmica hidrológica da Bacia, por concentrar nascentes de cabeceira, porém apresentando forte influência da expansão urbana em municípios como Solânea e Bananeiras.

Dessa forma, o uso do algoritmo K-Means mostrou-se eficiente na identificação de padrões ambientais e espaciais entre as nascentes da Bacia do Mamanguape, contribuindo para a interpretação diante da complexidade dos dados ambientais analisados. A utilização da clusterização demonstrou potencial como ferramenta de apoio à gestão ambiental, possibilitando identificar áreas prioritárias para a conservação e recuperação de APPs, assim como o monitoramento ambiental das nascentes. Sendo de suma importância para a manutenção da dinâmica hidrológica e da segurança hídrica regional. Recomenda-se, para pesquisas futuras, a incorporação de novas variáveis ambientais e análises de campo, visando aprofundar a compreensão da dinâmica das nascentes e aperfeiçoar estratégias de conservação na Bacia do Mamanguape.

REFERÊNCIAS

AMORIM, Rodrigo de Freitas; SOUZA, Jonas Otaviano Praça de; SOUSA, Silvio Braz de. Condição ambiental de nascentes perenes e intermitentes no Nordeste brasileiro – bacias no RN e na PB. **Revista Contexto Geográfico**, Maceió, v. 8, n. 17, p. 36-50, dez. 2023.

BRASIL. Lei nº 12.651, de 25 de maio de 2012. Dispõe sobre a proteção da vegetação nativa; altera as Leis nº 6.938, de 31 de agosto de 1981, nº 9.393, de 19 de dezembro de 1996, e nº 11.428, de 22 de dezembro de 2006; revoga as Leis nº 4.771, de 15 de setembro de 1965, e nº 7.754, de 14 de abril de 1989, e a Medida Provisória nº 2.166-67, de 24 de agosto de 2001; e dá outras providências. **Diário Oficial da União**: Brasília, DF, 28 maio 2012. Disponível em: http://www.planalto.gov.br/ccivil_03/_ato2011-2014/2012/lei/l12651.htm. Acesso em: 17 abril 2026.

BRASIL. Lei nº 12.727, de 17 de outubro de 2012. Altera a Lei nº 12.651, de 25 de maio de 2012, que dispõe sobre a proteção da vegetação nativa. **Diário Oficial da União**: Brasília, DF, 18 out. 2012. Disponível em: http://www.planalto.gov.br/ccivil_03/_ato2011-2014/2012/lei/l12727.htm. Acesso em: 17 abril 2026.

FELIPPE, Miguel Fernandes; GOMES, Cecília Siman; MAGALHÃES JÚNIOR, Antônio Pereira. Nascentes e áreas úmidas. In: MAGALHÃES JÚNIOR, Antônio Pereira; LOPES, Frederico Wagner de Azevedo. **Recursos hídricos: as águas na interface sociedade-natureza**. Belo Horizonte: Oficina de Texto, 2022. p. 49–57. Disponível em: https://integrada.minhabiblioteca.com.br/reader/books/978-65-86235-88-3. Acesso em: 17 abril 2026.

PEREIRA, Pedro Henrique Vaz; PEREIRA, Sueli Yoshinaga; YOSHINAGA, Alberto; PEREIRA, Paulo Ricardo Brum. Nascente: análise e discussão dos conceitos existentes. **Fórum Ambiental da Alta Paulista**, v. 7, n. 2, 2011. ISSN 1980-0827.

RUSSELL, Stuart; NORVIG, Peter. **Inteligência artificial**. Tradução da 3. ed. Rio de Janeiro: Elsevier, 2013. ISBN 978-85-352-3701-6.

SANTOS, Álvaro Rodrigues dos. As APPs associadas a nascentes: o que é uma nascente? Como identificá-la? **EcoDebate**, 21 jul. 2009. Disponível em: https://www.ecodebate.com.br/2009/07/21/as-apps-associadas-a-nascentes-o-que-e-umanascente-como-identifica-la-artigo-de-alvaro-rodrigues-dos-santos/. Acesso em: 17 abril 2026.

SOUZA NETO, Wilson Mendonça de. **Avaliação da distribuição espacial de zona de armazenamento de água em nascente perene de microbacia instável, Barra de Guaratiba, RJ**. Seropédica, RJ, 2010. Monografia (Graduação em Engenharia Florestal) – Universidade Federal Rural do Rio de Janeiro, Instituto de Florestas.