



UNIVERSIDADE FEDERAL DE CAMPINA GRANDE
CENTRO DE TECNOLOGIA E RECURSOS NATURAIS
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA CIVIL E AMBIENTAL
ÁREA: RECURSOS HÍDRICOS E SANEAMENTO AMBIENTAL

PATRÍCIA PINHEIRO DE BARROS

**MODELAGEM E ANÁLISE DA QUALIDADE DA ÁGUA EM RESERVATÓRIOS
UTILIZANDO MAPAS AUTO-ORGANIZÁVEIS**

Campina Grande – PB
2026

PATRÍCIA PINHEIRO DE BARROS

**MODELAGEM E ANÁLISE DA QUALIDADE DA ÁGUA EM RESERVATÓRIOS
UTILIZANDO MAPAS AUTO-ORGANIZÁVEIS**

Dissertação de mestrado submetida ao Programa de Pós-graduação em Engenharia Civil e Ambiental da Universidade Federal de Campina Grande para a obtenção do título de mestre em Engenharia Civil e Ambiental.

Orientador:

Prof. Dr. Camilo Allyson Simões de Farias

Campina Grande – PB

2026

Universidade Federal de Campina Grande - UFCG - Sistema de
Bibliotecas - SISTEMOTECA
Catalogação de Publicação na Fonte. UFCG - Biblioteca Central

B277m

Barros, Patrícia Pinheiro de.

Modelagem e análise da qualidade da água em reservatórios utilizando mapas auto-organizáveis / Patrícia Pinheiro de Barros. – 2026.

117 f. : il. color.

Dissertação (mestrado em Engenharia Civil e Ambiental) – Universidade Federal de Campina Grande, Centro de Tecnologia e Recursos Naturais, 2026.

“Orientação: Prof. Dr. Camilo Allyson Simões de Farias”.

Referências.

1. Qualidade da Água. 2. Redes Neurais Artificiais. 3. Self-Organizing Map. 4. Reservatórios. 5. Semiárido. I. Farias, Camilo Allyson Simões de. II. Título.

UFCG/BC

CDU 628.16(043.3)

PATRÍCIA PNHEIRO DE BARROS

**MODELAGEM E ANÁLISE DA QUALIDADE DA ÁGUA EM RESERVATÓRIOS
UTILIZANDO MAPAS AUTO-ORGANIZÁVEIS**

Dissertação de mestrado submetida ao Programa de Pós-Graduação em Engenharia Civil e Ambiental da Universidade Federal de Campina Grande para a obtenção do título de Mestre em Engenharia Civil e Ambiental em 16 de março de 2026 e avaliada pela seguinte comissão examinadora:

COMISSÃO EXAMINADORA

Prof. Dr. Camilo Allyson Simões de Farias

Universidade Federal de Campina Grande (UFCG)

Orientador

Prof. Dr. Carlos de Oliveira Galvão

Universidade Federal de Campina Grande (UFCG)

Examinador Interno

Prof. Dr. Cícero Fellipe Diniz de Santana

Universidade Federal de Campina Grande (UFCG)

Examinador Externo

Campina Grande – PB

2026



MINISTÉRIO DA EDUCAÇÃO
UNIVERSIDADE FEDERAL DE CAMPINA GRANDE
POS-GRADUACAO ENGENHARIA CIVIL AMBIENTAL
Rua Aprigio Veloso, 882, - Bairro Universitario, Campina Grande/PB, CEP 58429-900

REGISTRO DE PRESENÇA E ASSINATURAS

1. **ATA DA DEFESA PARA CONCESSÃO DO GRAU DE MESTRE EM ENGENHARIA CIVIL E AMBIENTAL**
2. **ALUNO(A): PATRÍCIA PINHEIRO DE BARROS / COMISSÃO EXAMINADORA: DR. CAMILO ALLYSON SIMÕES DE FARIAS - PPGECA/UFCG (PRESIDENTE) - ORIENTADOR, DR. CARLOS DE OLIVEIRA GALVÃO – PPGECA/UFCG – EXAMINADOR INTERNO, DR. CÍCERO FELLIPE DINIZ DE SANTANA – CCTA/UFCG – EXAMINADOR EXTERNO (PORTARIA 14/2026). / TÍTULO DA DISSERTAÇÃO: “MODELAGEM E ANÁLISE DA QUALIDADE DA ÁGUA EM RESERVATÓRIOS UTILIZANDO MAPAS AUTO-ORGANIZÁVEIS” / ÁREA DE CONCENTRAÇÃO: RECURSOS HÍDRICOS E SANEAMENTO AMBIENTAL / HORA DE INÍCIO: 17:00 HORAS / DE FORMA PRESENCIAL, NO LABORATÓRIO DE HIDRÁULICA II, BLOCO BU, CAMPUS SEDE DA UFCG, EM CAMPINA GRANDE.**
3. **EM SESSÃO REALIZADA DE FORMA PRESENCIAL, APÓS EXPOSIÇÃO DE CERCA DE 40 MINUTOS, O(A) CANDIDATO(A) FOI ARGUIDO(A) ORALMENTE PELOS MEMBROS DA COMISSÃO EXAMINADORA, TENDO DEMONSTRADO SUFICIÊNCIA DE CONHECIMENTO E CAPACIDADE DE SISTEMATIZAÇÃO NO TEMA DE SUA DISSERTAÇÃO, SENDO-LHE ATRIBUÍDA O CONCEITO “EM EXIGÊNCIA”, SENDO QUE A POSSIBILIDADE DE APROVAÇÃO ESTÁ CONDICIONADA À AVALIAÇÃO DA NOVA VERSÃO DO TRABALHO FINAL, SEGUINDO PROCEDIMENTOS PREVISTOS NA RESOLUÇÃO DO PROGRAMA. O PRESIDENTE DA COMISSÃO EXAMINADORA, OUVIDOS OS DEMAIS MEMBROS, DEVERÁ FICAR RESPONSÁVEL POR ATESTAR QUE AS CORREÇÕES SOLICITADAS NA LISTA DE EXIGÊNCIAS FORAM ATENDIDAS NA VERSÃO FINAL DO TRABALHO. A COMISSÃO EXAMINADORA CUMPRINDO OS PRAZOS REGIMENTAIS, ESTABELECE UM PRAZO MÁXIMO DE 30 DIAS PARA QUE SEJAM FEITAS AS ALTERAÇÕES EXIGIDAS. APÓS O DEPÓSITO FINAL DO DOCUMENTO DE DISSERTAÇÃO, DEVIDAMENTE REVISADO E MEDIANTE ATESTADO DO ORIENTADOR, O CONCEITO “EM EXIGÊNCIA” PASSARÁ IMEDIATAMENTE PARA O DE “APROVADO”. NA FORMA REGULAMENTAR, FOI LAVRADA A PRESENTE ATA, QUE É ASSINADA POR MIM, FLÁVIO PEREIRA DA CUNHA, SECRETÁRIO(A), ALUNO(A) E OS MEMBROS DA COMISSÃO EXAMINADORA PRESENTES.**
4. **CAMPINA GRANDE, 16 DE MARÇO DE 2026.**



Documento assinado eletronicamente por **PATRÍCIA PINHEIRO DE BARROS, Usuário Externo**, em 17/03/2026, às 10:15, conforme horário oficial de Brasília, com fundamento no art. 8º, caput, da [Portaria SEI nº 002, de 25 de outubro de 2018](#).



Documento assinado eletronicamente por **CAMILO ALLYSON SIMOES DE FARIAS, PROFESSOR(A) DO MAGISTERIO SUPERIOR**, em 17/03/2026, às 14:59, conforme horário oficial de Brasília, com fundamento no art. 8º, caput, da [Portaria SEI nº 002, de 25 de outubro de 2018](#).



Documento assinado eletronicamente por **FLAVIO PEREIRA DA CUNHA, SECRETÁRIO (A)**, em 17/03/2026, às 15:41, conforme horário oficial de Brasília, com fundamento no art. 8º, caput, da [Portaria SEI nº 002, de 25 de outubro de 2018](#).



Documento assinado eletronicamente por **CARLOS DE OLIVEIRA GALVAO, PROFESSOR(A) DO MAGISTERIO SUPERIOR**, em 17/03/2026, às 16:36, conforme horário oficial de Brasília, com fundamento no art. 8º, caput, da [Portaria SEI nº 002, de 25 de outubro de 2018](#).



Documento assinado eletronicamente por **Cícero Fellipe Diniz de Santana, Usuário Externo**, em 18/03/2026, às 08:55, conforme horário oficial de Brasília, com fundamento no art. 8º, caput, da [Portaria SEI nº 002, de 25 de outubro de 2018](#).



A autenticidade deste documento pode ser conferida no site <https://sei.ufcg.edu.br/autenticidade>, informando o código verificador **6334020** e o código CRC **7A4BC2F4**.

DEDICATÓRIA

A Deus, por tudo.

À minha mãe e meu irmão, por todo amor e incentivo.

AGRADECIMENTOS

Agradeço, primeiramente, a Deus, pelo dom da vida, pelas oportunidades que me concedeu e por ser meu refúgio e fortaleza, socorro sempre presente nos momentos em que mais precisei. Foi a minha fé n'Ele que me trouxe até aqui, e reconheço, em cada etapa, o Seu cuidado em minha vida.

Expresso minha profunda gratidão à minha família, em especial à minha mãe, Mercês Pinheiro, por todo o amor, cuidado e dedicação. Sempre esteve ao meu lado, fazendo o possível e o impossível por mim. Este título e tudo o que sou carregam um pouco da sua essência. Ao meu irmão, Felipe Pinheiro, por nunca medir esforços para me ver feliz e realizada, me apoiando, incentivando e cuidando de mim. Ao meu tio, Edilton Rodrigues, por ser um dos maiores incentivadores dos meus estudos.

Às amigas que construí durante a graduação e que permanecem comigo até hoje, tornando meus dias mais leves e felizes: Petra Marinho, Lucas Alves, Hebert Jonathan, Ayrton Natthan e Daniel Bezerra.

Sou igualmente grata aos amigos que tive o privilégio de conhecer durante o mestrado, verdadeiros presentes em minha vida, que levarei para sempre: Luana Alves, Isabel Martins, Karla Azevedo, Ingrid Moreira, Pedro Augusto, Roneide Martins, Sonaly Felix, Renan Filipe, Gabriele Souza, Saulo Victor, Ana Luiza, Fernanda Souto, Laisa Daiana, Silvia Raphaele e Sabrina Vivian.

Agradeço ainda às amigas fora do ambiente acadêmico, que sempre estiveram ao meu lado com palavras de apoio, carinho e orações: Wellington Velez, Yasmim Lopes, Lucas Macena e Lorrainy Souza.

Ao meu orientador, Prof. Dr. Camilo Allyson Simões de Farias, pela acolhida, orientação e dedicação ao longo desta trajetória.

Aos professores Prof. Dr. Carlos Galvão e Profa. Dra. Márcia Ribeiro, por me receberem com tanto carinho na UFCG, pelas conversas, pelos ensinamentos e pelo encorajamento ao longo do caminho.

À CAPES, pelo apoio financeiro, e a todos que contribuíram, direta ou indiretamente, para a realização deste trabalho, o meu mais sincero agradecimento.

RESUMO

A qualidade da água em reservatórios do semiárido brasileiro é fortemente influenciada pela variabilidade climática, pelos múltiplos usos da água e pelas pressões antrópicas, o que torna sua análise um desafio devido à complexidade e à multidimensionalidade dos dados ambientais. Nesse contexto, o presente estudo teve como objetivo aplicar redes neurais artificiais do tipo Self-Organizing Map (SOM) para analisar padrões de qualidade da água em sete reservatórios localizados na bacia hidrográfica do Rio Paraíba, no estado da Paraíba. Os dados utilizados foram provenientes do projeto PELD-RIPA, que contempla variáveis físicas, químicas e biológicas associadas à qualidade da água. Inicialmente, foram aplicadas técnicas estatísticas exploratórias e multivariadas, incluindo estatística descritiva, análise de correlação e Análise de Componentes Principais (PCA), com o intuito de compreender as relações entre as variáveis e selecionar indicadores relevantes para a modelagem. Em seguida, foram desenvolvidos modelos SOM em duas abordagens complementares: uma modelagem integrada, considerando conjuntamente os dados dos sete reservatórios, e modelagens individuais para cada sistema. Os resultados evidenciaram a capacidade do SOM de identificar padrões complexos e agrupamentos ambientais associados às condições de qualidade da água. A modelagem integrada permitiu uma visão comparativa entre os reservatórios, enquanto os modelos individuais permitiram identificar particularidades limnológicas de cada sistema. De modo geral, os resultados demonstram que os mapas auto-organizáveis constituem uma ferramenta eficiente para análise exploratória de dados ambientais multivariados, contribuindo para o monitoramento e para o suporte à gestão dos recursos hídricos em regiões semiáridas.

Palavras-chave: qualidade da água; redes neurais artificiais; Self-Organizing Map; reservatórios; semiárido.

ABSTRACT

Water quality in reservoirs located in the Brazilian semiarid region is strongly influenced by climatic variability, multiple water uses, and anthropogenic pressures, which makes its analysis challenging due to the complexity and multidimensionality of environmental data. In this context, the present study aimed to apply artificial neural networks of the Self-Organizing Map (SOM) type to analyze water quality patterns in seven reservoirs located in the Paraíba River basin, in the state of Paraíba, Brazil. The data used were obtained from the PELD-RIPA project and include physical, chemical, and biological variables associated with water quality. Initially, exploratory and multivariate statistical techniques were applied, including descriptive statistics, correlation analysis, and Principal Component Analysis (PCA), in order to understand the relationships among variables and to select relevant indicators for modeling. Subsequently, SOM models were developed using two complementary approaches: an integrated model considering the combined data from the seven reservoirs and individual models for each system. The results demonstrated the ability of SOM to identify complex patterns and environmental clusters associated with water quality conditions. The integrated modeling provided a comparative view among the reservoirs, while the individual models allowed the identification of specific limnological characteristics of each system. Overall, the results indicate that Self-Organizing Maps constitute an efficient tool for the exploratory analysis of multivariate environmental data, contributing to monitoring efforts and supporting water resources management in semiarid regions.

Keywords: water quality; artificial neural networks; Self-Organizing Map; reservoirs; semi-arid region.

LISTA DE FIGURAS

Figura 1 – Seleção do portfólio bibliográfico pelo método ProKnow-C e análise bibliométrica.	27
Figura 2 – Evolução temporal da produção científica sobre aplicação de SOM na análise da qualidade da água	28
Figura 3 – Rede de coocorrência das palavras-chave associadas aos estudos sobre aplicação de SOM na análise da qualidade da água	29
Figura 4 – Mapa temático da literatura científica sobre aplicação de SOM em estudos de qualidade da água.	30
Figura 5 – Reservatórios selecionados para o estudo	32
Figura 6 – Fluxograma das etapas metodológicas do estudo	39
Figura 7 – Histogramas das variáveis analisadas	50
Figura 8 – Boxplot das variáveis limnológicas utilizadas na análise estatística descritiva.....	51
Figura 9 – Mapa de calor da matriz de correlação de Pearson entre as variáveis de qualidade da água analisadas	52
Figura 10 – Círculo de correlação (PC1 × PC2).....	54
Figura 11 – U-Matrix do SOM (12 × 12)	56
Figura 12 – Mapas de componentes: variáveis físico-químicas (abordagem integrada).....	57
Figura 13 – Mapas de componentes: variáveis químicas e biológicas (abordagem integrada)	59
Figura 14 – Método do Cotovelo (Elbow Method)	61
Figura 15 – Índice de Silhueta.....	61
Figura 16 – Distribuição dos clusters no mapa SOM.....	62
Figura 17 – U-Matrix Acauã	66
Figura 18 – U-Matrix Camalaú	67
Figura 19 – U-Matrix Cordeiro	68
Figura 20 – U-Matrix Epitácio Pessoa	68
Figura 21 – U-Matrix Sumé	69
Figura 22 – U-Matrix Poções	69
Figura 23 – U-Matrix Taperoá	70
Figura 24 – Mapas de Componentes: variáveis físico-químicas de Acauã.....	71
Figura 25 – Mapas de Componentes: variáveis físico-químicas de Camalaú.....	71
Figura 26 – Mapas de Componentes: variáveis químicas e biológicas de Acauã.....	73

Figura 27 – Mapas de Componentes: variáveis químicas e biológicas de Camalaú	74
Figura 28 – Mapas de Componentes: variáveis físico-químicas de Cordeiro	76
Figura 29 – Mapas de Componentes: variáveis físico-químicas de Epitácio Pessoa	76
Figura 30 – Mapas de Componentes: variáveis químicas e biológicas de Cordeiro	78
Figura 31 – Mapas de Componentes: variáveis químicas e biológicas de Epitácio Pessoa	79
Figura 32 – Mapas de Componentes: variáveis físico-químicas de Poções	81
Figura 33 – Mapas de Componentes: variáveis físico-químicas de Sumé	81
Figura 34 – Mapas de Componentes: variáveis químicas e biológicas de Poções	83
Figura 35 – Mapas de Componentes: variáveis químicas e biológicas de Sumé	83
Figura 36 – Mapas de Componentes: variáveis físico-químicas de Taperoá	85
Figura 37 – Mapas de Componentes: variáveis químicas e biológicas de Taperoá	86
Figura 38 – Mapa de clusters do SOM de Acauã	89
Figura 39 – Mapa de clusters do SOM de Camalaú	90
Figura 40 – Mapa de clusters do SOM de Cordeiro	92
Figura 41 – Mapa de clusters do SOM de Epitácio Pessoa	93
Figura 42 – Mapa de clusters do SOM de Poções	95
Figura 43 – Mapa de clusters do SOM de Sumé	96
Figura 44 – Mapa de clusters do SOM de Taperoá	98

LISTA DE TABELAS

Tabela 1 – Estatística descritiva das variáveis de qualidade da água utilizadas na modelagem integrada	49
Tabela 2 – Variância explicada pelos componentes principais	54
Tabela 3 – Média das variáveis analisadas em cada cluster identificado no SOM	64
Tabela 4 – Erro de Quantização (QE) dos modelos SOM individuais.....	66
Tabela 5 – Distribuição das amostras por cluster identificado nos mapas individuais dos reservatórios analisados.....	88
Tabela 6 – Valores médios das variáveis ambientais para os clusters identificados no SOM individual do reservatório de Acauã.....	90
Tabela 7 – Valores médios das variáveis ambientais para os clusters identificados no SOM individual do reservatório de Camalaú.....	91
Tabela 8 – Valores médios das variáveis ambientais para os clusters identificados no SOM individual do reservatório de Cordeiro.....	92
Tabela 9 – Valores médios das variáveis ambientais para os clusters identificados no SOM individual do reservatório de Epitácio Pessoa.....	94
Tabela 10 – Valores médios das variáveis ambientais para os clusters identificados no SOM individual do reservatório de Poções.....	96
Tabela 11 – Valores médios das variáveis ambientais para os clusters identificados no SOM individual do reservatório de Sumé.....	97
Tabela 12 – Valores médios das variáveis ambientais para os clusters identificados no SOM individual do reservatório de Taperoá.....	99

LISTA DE EQUAÇÕES

Equação 1	41
Equação 2	41
Equação 3	42
Equação 4	42
Equação 5	43
Equação 6	44
Equação 7	44
Equação 8	44
Equação 9	44
Equação 10	45
Equação 11	45
Equação 12	46

LISTA DE ABREVIATURAS E SIGLAS

AESA	Agência Executiva de Gestão das Águas
ANA	Agência Nacional de Águas e Saneamento Básico
BMU	Best Matching Unit
CAGEPA	Companhia de Água e Esgotos da Paraíba
CNPq	Conselho Nacional de Desenvolvimento Científico e Tecnológico
DNOCS	Departamento Nacional de Obras Contra as Secas
IA	Inteligência Artificial
ML	Machine Learning
PCA	Principal Component Analysis
PELD	Programa de Pesquisa Ecológica de Longa Duração
PELD-RIPA	Projeto Ecológico de Longa Duração do Rio Paraíba Integrado
PIS	Perímetro Irrigado de Sumé
PISF	Projeto de Integração do Rio São Francisco
RNA	Rede Neural Artificial
QE	Quantization Error
SOM	Self-Organizing Map
UFCG	Universidade Federal de Campina Grande
U-Matrix	Unified Distance Matrix

SUMÁRIO

1	INTRODUÇÃO.....	15
1.1	CONTEXTO GERAL	15
1.2	JUSTIFICATIVA	18
1.3	OBJETIVOS.....	19
1.3.1	Objetivo Geral	19
1.3.2	Objetivos Específicos	19
1.4	ESTRUTURA DA DISSERTAÇÃO	19
2	REVISÃO DE LITERATURA	22
2.1	RECURSOS HÍDRICOS E QUALIDADE DA ÁGUA	22
2.2	APRENDIZADO DE MÁQUINA APLICADO À MODELAGEM DE QUALIDADE DA ÁGUA	22
2.3	MAPAS AUTO-ORGANIZÁVEIS	24
2.4	APLICAÇÃO DO SOM EM ESTUDOS DE QUALIDADE DA ÁGUA.....	25
2.4.1	Procedimento de busca e construção do portfólio bibliográfico	26
2.4.2	Panorama da produção científica sobre SOM e qualidade da água.....	27
3	METODOLOGIA.....	32
3.1	ÁREA DE ESTUDO	32
3.1.1	Caracterização da Bacia Hidrográfica	32
3.1.2	Caracterização dos Reservatórios.....	33
3.1.2.1	Sumé	33
3.1.2.2	Taperoá	34
3.1.2.3	Epitácio Pessoa	34
3.1.2.4	Camalaú	35
3.1.2.5	Poções.....	36
3.1.2.6	Cordeiro	36
3.1.2.7	Acauã	37
3.2	AQUISIÇÃO DE DADOS	37
3.2.1	Base de Dados	37
3.2.2	Variáveis Analisadas	38
3.2.3	Tratamento Preliminar	39
3.3	ANÁLISE ESTATÍSTICA.....	41
3.3.1	Estatística Descritiva	41
3.3.2	Análise Multivariada Convencional	42
3.4	MODELAGEM COM SELF-ORGANIZING MAPS (SOM).....	42
3.4.1	Preparação dos dados para aplicação do SOM.....	43

3.4.2	Configuração do modelo	43
3.4.3	Estratégia de aplicação do SOM.....	45
3.4.4	Crerérios de avaliação e interpretação.....	46
4	RESULTADOS E DISCUSSÖES.....	49
4.1	ESTATÍSTICA DESCRITIVA	49
4.2	CORRELAÇÃO ENTRE VARIÁVEIS.....	52
4.3	ANÁLISE DE COMPONENTES PRINCIPAIS (PCA).....	54
4.4	MODELAGEM COM SOM – ABORDAGEM INTEGRADA	55
4.4.1	Treinoamento e desempenho do modelo.....	56
4.4.2	Estrutura do mapa SOM (U-Matrix)	56
4.4.3	Mapas de componentes.....	57
4.4.4	Clusterização do mapa (k-means).....	61
4.4.5	Interpretação ambiental dos agrupamentos	63
4.5	MODELAGEM COM SOM – APLICAÇÃO INDIVIDUAL POR RESERVATÓRIO.....	65
4.5.1	Treinoamento e desempenho dos modelos individuais	65
4.5.2	Estrutura dos mapas SOM individuais	66
4.5.3	Mapas de componentes individuais	71
4.5.3.1	Acauã e Camalaú	71
4.5.3.2	Cordeiro e Epitácio Pessoa	76
4.5.3.3	Poções e Sumé	81
4.5.3.4	Taperoá	85
4.5.4	Clusterização dos SOM individuais	88
4.6	COMPARAÇÃO ENTRE MODELAGEM INTEGRADA E A MODELAGEM INDIVIDUAL	100
5	CONCLUSÖES	103
5.1	RECOMENDAÇÕES.....	105
	REFERÊNCIAS BIBLIOGRÁFICAS	106

CAPÍTULO I

INTRODUÇÃO

Neste primeiro capítulo são introduzidos os principais elementos que fundamentam esta dissertação, estando ele estruturado em cinco seções que abordam o contexto do estudo, a justificativa da pesquisa, o objetivo geral e objetivos específicos, bem como a organização da dissertação.

1 INTRODUÇÃO

1.1 CONTEXTO GERAL

As mudanças climáticas têm provocado alterações expressivas nos padrões de temperatura e precipitação, afetando diretamente ecossistemas e recursos hídricos em diversas regiões do mundo. Esses impactos manifestam-se por meio de danos ambientais progressivos e, em alguns casos, irreversíveis. As modificações nos regimes pluviométricos, especialmente a redução dos volumes de chuva e as alterações no início e término das estações chuvosas, evidenciam a maior vulnerabilidade de regiões cuja dinâmica socioeconômica está diretamente vinculada à disponibilidade hídrica (IPCC, 2023).

Nas últimas décadas, tornou-se cada vez mais evidente a influência das mudanças climáticas sobre a dinâmica e a ocorrência de secas. Projeções indicam que, com a intensificação do aquecimento global, as secas tendem a se tornar mais frequentes, severas e prolongadas em diferentes partes do mundo (Gu et al., 2019; Naumann et al., 2018).

No Brasil, esses efeitos manifestam-se de forma particularmente intensa na região semiárida, caracterizada por regime pluviométrico irregular, associado a elevadas taxas de evapotranspiração potencial, fatores que contribuem para a instabilidade na disponibilidade hídrica (Alvalá et al., 2019). Embora diferentes eventos extremos possam ocorrer nessa região, a seca destaca-se como o fenômeno de maior recorrência e relevância. Definida como a redução prolongada da disponibilidade hídrica em relação às condições médias históricas, a seca provoca impactos que transcendem o aspecto climático, afetando os regimes hidrológicos, a produção agrícola, os ecossistemas e a dinâmica socioeconômica das populações (Brito et al., 2021).

A bacia hidrográfica do Rio Paraíba, por exemplo, insere-se nesse contexto de elevada sensibilidade hidrológica. Trata-se de uma das principais bacias do semiárido brasileiro, marcada por longos períodos de estiagem. Segundo a classificação de Köppen-Geiger, a região apresenta clima do tipo BSh (semiárido quente), caracterizado por precipitações escassas e irregulares, com média anual em torno de 400 mm (Paiva, 2023). Aproximadamente 65% da precipitação anual concentra-se entre os meses de janeiro e junho, enquanto a estação seca pode se estender de 8 a 10 meses (AESAs, 2025), configurando um cenário de forte variabilidade hidrológica e recorrente escassez hídrica.

Nesse contexto, a construção de reservatórios artificiais é uma alternativa importante para mitigar os problemas de abastecimento de água e assegurar a sobrevivência das populações

locais. Trata-se de um dos sistemas de engenharia mais antigos, consolidado historicamente como uma das primeiras estratégias políticas de enfrentamento e mitigação dos efeitos da seca, desempenhando relevante papel social, econômico e ambiental (Álvaro et al., 2023; Pereira Neto, 2017). Além disso, esses reservatórios são concebidos para atender a múltiplos usos, incluindo abastecimento humano, irrigação, dessedentação animal, pesca e produção aquícola, assumindo função estratégica no contexto da gestão integrada dos recursos hídricos e no desenvolvimento regional.

Apesar de essenciais, os reservatórios constituem sistemas ambientalmente complexos e vulneráveis, altamente suscetíveis às pressões antrópicas e às variações climáticas. Segundo Gomes et al. (2022), a gestão operacional desses sistemas não deve se restringir a parâmetros estritamente quantitativos, sendo indispensável incorporar a dimensão qualitativa da água nos processos decisórios, especialmente em regiões semiáridas. Nesses ambientes, períodos prolongados de estiagem e elevadas taxas de evaporação provocam oscilações expressivas nos volumes armazenados e alterações nas características morfométricas do reservatório, intensificando a concentração de solutos e modificando suas condições físico-químicas e biológicas.

Essas alterações refletem-se diretamente nos parâmetros que caracterizam a qualidade da água. Dessa forma, a dinâmica de nutrientes assume papel central nas pesquisas sobre a qualidade hídrica dos reservatórios em regiões semiáridas, com destaque para o fósforo, cujo aumento de concentração pode decorrer tanto de cargas externas quanto de fontes internas associadas aos sedimentos (Rocha Júnior et al., 2024; Lima Neto et al., 2022; Rocha et al., 2025; Santos et al., 2025). A elevação do aporte de nutrientes, seja de origem natural ou antropogênica, associada ao aumento das temperaturas globais, favorece a ocorrência de florações tóxicas de cianobactérias (Barros et al., 2019). Adicionalmente, a maior turbidez decorrente da biomassa algal pode reduzir a penetração de luz na coluna d'água, influenciando processos físicos e energéticos no sistema (Rodrigues et al., 2025).

O aumento da temperatura também interfere no gradiente vertical de densidade da água, promovendo períodos mais longos e estáveis de estratificação térmica. Essa estratificação limita a circulação e a transferência de massa entre as camadas superficiais e profundas, favorecendo, em conjunto com o acúmulo de matéria orgânica, a depleção de oxigênio dissolvido nas camadas mais profundas. Em situações extremas, o hipolímnio pode atingir condições anóxicas, o que estimula a liberação de fósforo do sedimento para a coluna d'água, retroalimentando o processo de eutrofização (Carneiro et al., 2023). Como consequência dessas interações, observam-se elevações nas concentrações de nutrientes, no pH, na condutividade elétrica, na

turbidez e na biomassa algal, evidenciando o comprometimento da qualidade da água em reservatórios submetidos a condições de estresse climático e pressões antrópicas (Cortez et al., 2022; Costa et al., 2016).

Freire et al. (2025) evidenciam que a governança eficaz dos recursos hídricos demanda uma abordagem integrada que considere simultaneamente a disponibilidade hídrica, a qualidade da água e os usos múltiplos. Nesse sentido, o monitoramento sistemático da qualidade da água assume papel estratégico e essencial, especialmente em sistemas destinados ao abastecimento humano e outros usos produtivos. A avaliação contínua de parâmetros físicos, químicos e biológicos possibilita verificar a conformidade com os padrões de qualidade estabelecidos pela legislação vigente, além de subsidiar a análise da condição ecológica dos ecossistemas aquáticos, a prevenção de riscos à saúde pública e o suporte aos diferentes usos da água.

Entretanto, apesar de amplamente empregadas, as abordagens convencionais de monitoramento apresentam limitações relevantes. A geração desses dados depende de procedimentos técnicos que envolvem mobilização de equipes especializadas, uso de equipamentos específicos e posterior processamento analítico, o que implica demanda significativa de recursos financeiros e operacionais. Além disso, a necessidade de campanhas periódicas nem sempre é compatível com restrições logísticas e orçamentárias, dificultando a construção de séries temporais contínuas e a adequada representação da variabilidade espacial e sazonal dos sistemas hídricos.

Diante desse cenário, torna-se fundamental incorporar tecnologias e ferramentas que aprimorem a eficiência da gestão dos recursos hídricos, ampliando a capacidade de análise e interpretação dos dados disponíveis (Santos et al., 2025). Sob essa perspectiva, abordagens computacionais e métodos inteligentes emergem como alternativas promissoras para integrar múltiplas variáveis, identificar padrões e apoiar processos decisórios.

Diante disso, o presente estudo tem como objetivo aplicar redes neurais do tipo Self-Organizing Map (SOM) para analisar a qualidade da água de sete reservatórios localizados na bacia do Rio Paraíba. Os mapas auto-organizáveis, uma técnica de rede neural não supervisionada, têm se mostrado promissores pela capacidade de estruturar e visualizar dados multivariados, permitindo revelar relações não lineares e organizar informações ambientais de alta dimensionalidade de forma intuitiva e interpretável. No entanto, seu uso ainda é pouco consolidado na área de recursos hídricos, especialmente em estudos voltados à qualidade da água. Diante dessa lacuna, a proposta deste trabalho é desenvolver uma ferramenta robusta de análise e apoio à interpretação dos dados, que auxilie na gestão dos recursos hídricos, permitindo intervenções mais eficazes e tempestivas no manejo da água na região.

1.2 JUSTIFICATIVA

A construção de reservatórios artificiais constitui uma alternativa fundamental para mitigar os efeitos da escassez hídrica em regiões sujeitas a períodos prolongados de seca e irregularidade do regime pluviométrico. Esses reservatórios desempenham funções ecológicas, sociais e econômicas significativas, fornecendo água para consumo humano, irrigação, recreação e atividades pesqueiras (Azevedo et al., 2017; Gutiérrez et al., 2014; Tundisi, 2008).

Em regiões semiáridas, a escassez de água ocorre não apenas em termos quantitativos, mas também qualitativos. Os reservatórios nessa região são fortemente influenciados pela variabilidade hidroclimática, uma vez que as características físico-químicas desses sistemas estão diretamente associadas ao regime hidrológico e ao fluxo e volume armazenado (Raulino; Silveira; Neto, 2021). Segundo Wiegand et al. (2021), diversos fatores contribuem para a degradação da qualidade da água, incluindo variações no volume dos reservatórios, padrões de precipitação, aportes externos de nutrientes e processos de enriquecimento interno relacionados a atividades antrópicas, como a aquicultura. Adicionalmente, condições típicas do semiárido, como elevadas temperaturas, interrupções no fluxo fluvial, múltiplos usos da água e longos períodos de residência hidrológica, favorecem a acumulação de nutrientes, bem como o aumento de componentes orgânicos e inorgânicos dissolvidos, resultando, por exemplo, em maior turbidez (Medeiros et al., 2025).

Embora existam métodos estatísticos convencionais amplamente utilizados na análise da qualidade da água, eles podem ser limitados na identificação de padrões complexos, não lineares e multivariados (Chou et al., 2018). Além disso, Lacerda (2024) destaca que, na obtenção de dados de qualidade da água, existem dificuldades associadas às análises laboratoriais, especialmente no monitoramento de contaminantes microbiológicos, que requerem o uso de substratos de maior custo e demandam períodos prolongados para a obtenção dos resultados, o que pode comprometer a rapidez e a eficiência da avaliação da qualidade da água.

Nesse contexto, a aplicação de ferramentas inteligentes, como as redes neurais artificiais, representa um avanço significativo no monitoramento da qualidade da água em reservatórios, permitindo um controle mais eficaz e proativo desses recursos. A escolha dos Mapas Auto-Organizáveis (Self-Organizing Maps), por sua vez, oferece uma abordagem eficiente para a identificação de padrões complexos, agrupamentos e tendências ocultas em grandes volumes de dados ambientais.

1.3 OBJETIVOS

1.3.1 Objetivo Geral

Desenvolver e implementar um modelo de rede neural do tipo Self-Organizing Map (SOM) para a modelagem da qualidade da água em sete reservatórios da bacia do Rio Paraíba, com o intuito de oferecer uma ferramenta eficaz para a análise da qualidade da água, complementando as abordagens estatísticas convencionais.

1.3.2 Objetivos Específicos

- Analisar as variáveis de qualidade da água por meio de técnicas estatísticas convencionais, visando caracterizar sua variabilidade e explorar as relações entre os parâmetros ambientais;
- Aplicar um modelo SOM integrado, treinado com dados combinados dos sete reservatórios, para permitir uma análise comparativa abrangente da qualidade da água entre os diferentes sistemas;
- Identificar e interpretar os padrões de qualidade da água em cada reservatório por meio de modelos SOM individuais;
- Comparar e avaliar os resultados obtidos pelos modelos SOM e pelas técnicas estatísticas convencionais, buscando examinar a complementaridade entre as abordagens na interpretação da qualidade da água.

1.4 ESTRUTURA DA DISSERTAÇÃO

Esta dissertação de mestrado é composta por cinco capítulos, de forma a proporcionar uma abordagem estruturada e clareza ao seu leitor. A seguir, são apresentadas a organização e o conteúdo de cada capítulo:

Capítulo I – Introdução: Este capítulo contextualiza o tema da pesquisa, destacando sua relevância para a área de Recursos Hídricos, com ênfase na qualidade da água. São apresentados a justificativa do estudo, os objetivos gerais e específicos e a estrutura da dissertação.

Capítulo II – Revisão de literatura: Apresenta uma revisão da literatura sobre o uso de redes SOM em estudos de recursos hídricos, com foco em sua aplicação na análise e avaliação da qualidade da água.

Capítulo III – Metodologia: Neste capítulo, descreve-se a área de estudo e discute-se a importância dos sete reservatórios analisados. Em seguida, são detalhados os procedimentos metodológicos adotados, incluindo a coleta e o tratamento dos dados, as técnicas estatísticas convencionais e a implementação do modelo SOM.

Capítulo IV – Resultados e Discussões: Este capítulo apresenta a aplicação de técnicas estatísticas descritivas e multivariadas aos dados de qualidade da água, bem como o uso do modelo de Mapas Auto-Organizáveis (SOM). A análise é conduzida em duas etapas: inicialmente, o SOM é treinado e aplicado de forma integrada a todos os reservatórios, o que possibilita uma comparação global. Posteriormente, o modelo é aplicado individualmente a cada reservatório, permitindo a identificação de padrões internos. Por fim, discutem-se os principais resultados obtidos, bem como as vantagens e limitações das análises estatísticas convencionais em comparação com o modelo SOM.

Capítulo V – Conclusões: Por fim, serão discutidas as contribuições da pesquisa, suas restrições e sugestões para trabalhos futuros.

CAPÍTULO II

REVISÃO DE LITERATURA

Este capítulo apresenta a revisão de literatura que fundamenta esta pesquisa, abordando conceitos relacionados à qualidade da água, ao aprendizado de máquina e aos Mapas Auto-Organizáveis (Self-Organizing Maps – SOM) aplicados a estudos de recursos hídricos. Inicialmente, são discutidos aspectos gerais sobre recursos hídricos e qualidade da água. Em seguida, são apresentados os fundamentos do aprendizado de máquina e os princípios do SOM. Por fim, são discutidas as aplicações do SOM em estudos de qualidade da água, com base em uma análise bibliométrica da literatura.

2 REVISÃO DE LITERATURA

2.1 RECURSOS HÍDRICOS E QUALIDADE DA ÁGUA

A qualidade da água é um dos principais indicadores da integridade ecológica dos sistemas aquáticos e desempenha papel fundamental na manutenção dos serviços ecossistêmicos e na segurança hídrica. Rios, lagos e reservatórios são ambientes altamente dinâmicos, nos quais a qualidade da água resulta da interação entre fatores naturais e pressões antrópicas, como uso do solo, atividades agrícolas, urbanização e processos biogeoquímicos internos.

A gestão sustentável dos recursos hídricos tem se consolidado como um dos grandes desafios ambientais em escala global. Esse desafio torna-se ainda mais evidente em regiões áridas e semiáridas, onde a elevada variabilidade hidrológica, associada ao aumento das pressões antrópicas sobre os sistemas aquáticos, tende a intensificar os problemas relacionados à disponibilidade e à qualidade da água (Salem; Pudza; Yihdego, 2022).

Diante desse cenário, o monitoramento da qualidade da água assume um papel fundamental para subsidiar estratégias de gestão e conservação dos recursos hídricos. A compreensão das características dos dados de qualidade da água, incluindo suas variabilidades espaciais e temporais, constitui etapa essencial para o desenvolvimento de diagnósticos ambientais consistentes e para a implementação de medidas eficazes de gestão e controle da poluição (Badrudeen et al., 2026).

Nesse contexto, a análise da qualidade da água frequentemente envolve conjuntos de dados complexos e multivariados, o que demanda métodos capazes de identificar padrões e relações entre diferentes variáveis ambientais. Assim, técnicas computacionais baseadas em aprendizado de máquina têm sido cada vez mais empregadas como ferramentas auxiliares para a interpretação e modelagem desses dados.

2.2 APRENDIZADO DE MÁQUINA APLICADO À MODELAGEM DE QUALIDADE DA ÁGUA

O aprendizado de máquina (Machine Learning – ML) é um dos principais ramos da Inteligência Artificial (IA) e baseia-se no desenvolvimento de algoritmos capazes de aprender padrões a partir de conjuntos de dados. Os modelos de aprendizado de máquina podem ser classificados de acordo com seus objetivos analíticos, podendo assumir caráter descritivo ou

preditivo. Modelos descritivos são utilizados para explorar e compreender os dados disponíveis, permitindo identificar padrões, dependências e estruturas presentes no conjunto de dados analisado. Já os modelos preditivos têm como finalidade estimar valores ou comportamentos futuros a partir de dados históricos, sendo amplamente aplicados em situações que envolvem previsão e suporte à tomada de decisão (Klompenburg; Kassahun; Catal, 2020).

A avaliação da qualidade da água, fundamentada exclusivamente em índices ou em abordagens estatísticas tradicionais, apresenta limitações na interpretação de sistemas ambientais complexos. Esses métodos geralmente simplificam a informação ao integrar múltiplos parâmetros em um único indicador ou ao pressupor relações lineares entre as variáveis, o que dificulta a identificação de padrões mais complexos e de interações não lineares em sistemas aquáticos (Gad et al., 2024). Ademais, tais abordagens demonstram limitações na discriminação de características relevantes ou na revelação de estruturas ocultas nos dados, especialmente em conjuntos extensos e multivariados.

Nesse contexto, técnicas baseadas em aprendizado de máquina têm se mostrado promissoras para a análise de dados ambientais. Essas abordagens permitem explorar grandes volumes de informações, identificar relações complexas entre variáveis e aprimorar o desempenho das análises preditivas e exploratórias. Assim, o uso de algoritmos de ML tem recebido crescente atenção em estudos voltados ao monitoramento e à modelagem da qualidade da água, contribuindo para o desenvolvimento de ferramentas mais robustas de análise ambiental (Bourhabi et al., 2026; Mengistu; Chung; Chang, 2026).

Os algoritmos de aprendizado de máquina podem ser classificados de acordo com o tipo de problema que buscam resolver e com a forma como utilizam os dados disponíveis para o processo de aprendizagem. De modo geral, destacam-se quatro principais abordagens: aprendizado supervisionado, aprendizado não supervisionado, aprendizado semi-supervisionado e aprendizado por reforço.

No aprendizado supervisionado, o algoritmo é treinado a partir de um conjunto de dados rotulados, no qual cada instância de entrada está associada a uma saída conhecida. O objetivo do modelo é aprender a relação entre essas variáveis para que, após o treinamento, seja capaz de prever resultados para novos dados ainda não observados. Essa abordagem é amplamente utilizada em tarefas de classificação e de regressão. Já no aprendizado não supervisionado, os dados não possuem rótulos associados e o objetivo do algoritmo é identificar padrões, estruturas ou agrupamentos presentes neles. Técnicas desse tipo são frequentemente empregadas em análises exploratórias, redução de dimensionalidade e identificação de clusters em conjuntos de dados complexos (Salah; Hayette, 2025).

O aprendizado semi-supervisionado representa uma abordagem intermediária, na qual o modelo é treinado utilizando uma pequena quantidade de dados rotulados combinada com um conjunto maior de dados não rotulados. Essa estratégia é particularmente útil em situações em que a obtenção de rótulos é custosa ou limitada, permitindo aproveitar melhor as informações disponíveis no conjunto de dados. Por fim, o aprendizado por reforço baseia-se em um processo de interação entre um agente e o ambiente, no qual o algoritmo aprende a tomar decisões por meio de tentativas sucessivas, recebendo recompensas ou penalidades conforme as ações realizadas (Alpaydin, 2020).

O enfoque desta pesquisa será direcionado às abordagens de aprendizado não supervisionado, que permitem identificar padrões e estruturas em conjuntos de dados ambientais sem a necessidade de rótulos previamente definidos.

2.3 MAPAS AUTO-ORGANIZÁVEIS

Os Mapas Auto-Organizáveis (Self-Organizing Maps – SOM), também conhecidos como Mapas de Kohonen, são um tipo de rede neural artificial baseada em aprendizado não supervisionado. Esse modelo foi proposto no início da década de 1980 por Teuvo Kohonen, como uma abordagem capaz de realizar projeções não lineares de dados de alta dimensionalidade em representações de menor dimensão, preservando as relações de similaridade entre as observações originais. Desde então, os SOM têm sido amplamente aplicados em diferentes áreas do conhecimento, como indústria, finanças, ciências naturais e linguística, especialmente em problemas relacionados à análise exploratória de dados e à identificação de padrões em conjuntos de dados complexos (Kohonen, 2013).

Do ponto de vista estrutural, o modelo SOM é composto por um conjunto de neurônios artificiais organizados em uma grade topológica, geralmente bidimensional, na qual cada neurônio está associado a um vetor de pesos que representa as características do conjunto de dados analisado. Durante o processo de treinamento, esses neurônios competem entre si para representar os padrões presentes nos dados de entrada, ajustando progressivamente seus vetores de peso com base nas similaridades entre as observações. Esse mecanismo adaptativo permite que o modelo organize automaticamente os dados, agrupando observações com características semelhantes em regiões próximas do mapa (Song et al., 2024).

Uma das principais características dos SOM é a capacidade de organizar os dados multivariados complexos em mapas bidimensionais que mantêm a proximidade entre observações semelhantes. Essa propriedade facilita a visualização e interpretação de padrões

presentes nos dados, tornando o método particularmente útil para análises exploratórias em bases de dados extensas (Melo et al., 2019). Em estudos ambientais, especialmente aqueles relacionados à qualidade da água, essa característica torna-se ainda mais relevante, uma vez que os programas de monitoramento frequentemente geram grandes volumes de dados compostos por diversas variáveis físicas, químicas e biológicas. Nesse contexto, os SOM têm sido amplamente utilizados como ferramentas capazes de representar relações não lineares entre variáveis e auxiliar na interpretação de conjuntos de dados multivariados (Hadjisolomou et al., 2018). Além disso, a capacidade do método de agrupar observações semelhantes e identificar padrões de variação em parâmetros ambientais torna o SOM uma ferramenta útil para investigações relacionadas à classificação da qualidade da água, à identificação de fontes de poluição e à análise de padrões espaciais e sazonais em sistemas aquáticos (Ahmadmoazzam et al., 2021).

2.4 APLICAÇÃO DO SOM EM ESTUDOS DE QUALIDADE DA ÁGUA

Diversos estudos têm demonstrado o potencial das redes neurais do tipo Self-Organizing Map (SOM) para a análise de dados ambientais, especialmente na avaliação da qualidade da água em diferentes sistemas aquáticos. Na literatura, o SOM tem sido aplicado em diferentes abordagens de monitoramento e análise da qualidade da água. Entre essas aplicações, destacam-se a classificação de estados tróficos em reservatórios (Lu; Lo, 2002), a investigação das relações entre características da paisagem e a qualidade da água (Park et al., 2014), a identificação de padrões espaciais de degradação em bacias hidrográficas (Chea; Grenouillet; Lek, 2016), bem como a análise de processos biogeoquímicos em reservatórios (Melo et al., 2019). Além disso, a técnica tem sido utilizada na avaliação de riscos associados a florações algais (Chen et al., 2014) e na análise da distribuição de elementos potencialmente tóxicos em sistemas fluviais (Cai et al., 2025). Outras aplicações incluem a investigação de padrões limnológicos em lagoas costeiras (Jeong et al., 2008), a identificação de fontes de matéria orgânica dissolvida em rios e lagos (Jin et al., 2024; Song et al., 2024) e a análise de fatores que influenciam a qualidade da água em áreas industriais e urbanizadas (Zhao et al., 2025). Esses estudos evidenciam a versatilidade do SOM como ferramenta de apoio à análise exploratória de dados ambientais.

Apesar do avanço dessas aplicações, ainda se observa uma lacuna na sistematização da produção científica sobre o uso de SOM em estudos de qualidade da água. Ainda não está claramente consolidado na literatura como esse campo de pesquisa tem evoluído ao longo do

tempo, quais são seus principais autores, periódicos científicos, temas investigados e tendências de desenvolvimento. Diante desse contexto, torna-se relevante compreender como a literatura científica tem se estruturado em torno da aplicação de SOM na análise da qualidade da água em sistemas hídricos.

2.4.1 Procedimento de busca e construção do portfólio bibliográfico

Com o objetivo de identificar e analisar os principais estudos sobre a aplicação de Self-Organizing Maps (SOM) na análise da qualidade da água em sistemas aquáticos, foi realizada uma revisão bibliográfica estruturada, associada a uma análise bibliométrica da produção científica. Para garantir maior rigor e transparência na seleção dos trabalhos, adotou-se o método ProKnow-C (Knowledge Development Process – Constructivist), amplamente utilizado na construção de portfólios bibliográficos relevantes em diferentes áreas do conhecimento.

Inicialmente, definiu-se a temática central da busca como a aplicação de Self-Organizing Maps na análise da qualidade da água em sistemas aquáticos continentais. A busca bibliográfica foi realizada nas bases de dados Scopus e Web of Science, consideradas importantes repositórios internacionais de literatura científica. A estratégia de busca resultou na recuperação inicial de 189 registros científicos, sendo 146 provenientes da base Scopus e 43 da base Web of Science.

Os registros obtidos foram posteriormente integrados e organizados no ambiente R, utilizando o pacote Bibliometrix, que permitiu a padronização das informações bibliográficas e a preparação da base de dados para as etapas subsequentes de análise. Após a remoção de registros duplicados, a base consolidada passou a ser composta por 100 artigos científicos únicos.

A partir dessa base inicial, foi realizada a etapa de filtragem do portfólio bibliográfico seguindo as etapas estruturadas do método ProKnow-C. Inicialmente, procedeu-se à análise dos títulos dos artigos para verificar o alinhamento temático com o escopo da pesquisa, identificando-se 33 estudos considerados aderentes ao tema. Em seguida, foi realizada a análise de reconhecimento científico dos estudos com base no número de citações, aplicando-se o princípio de Pareto, com ponto de corte mínimo de 24 citações.

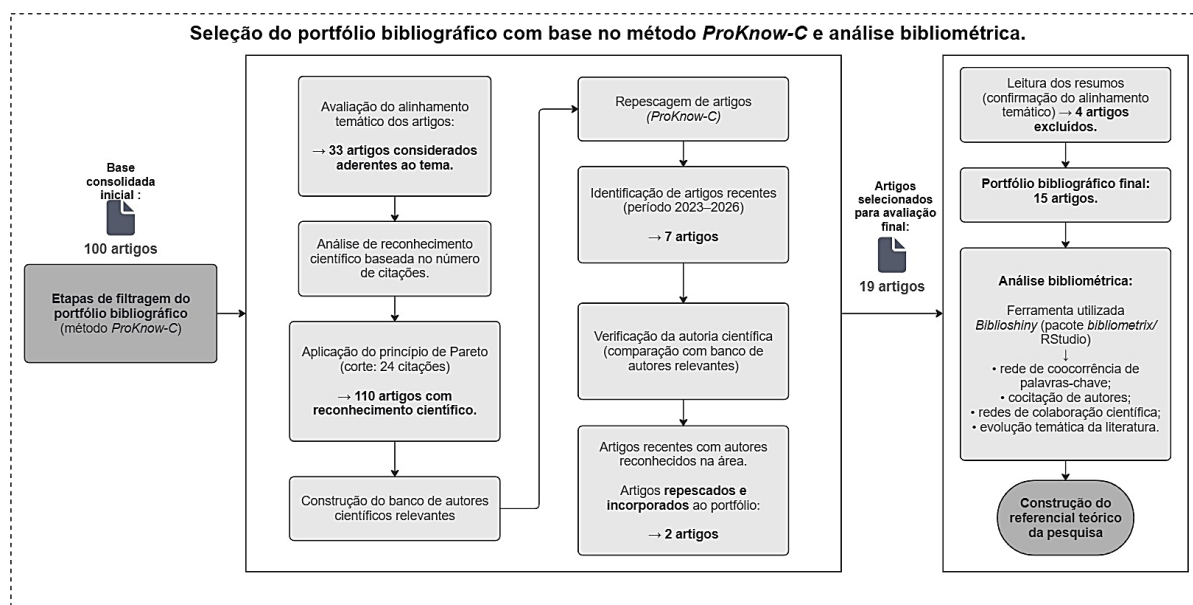
Com o objetivo de evitar viés temporal e incorporar contribuições recentes, ainda com baixo número de citações, foram também considerados artigos publicados entre 2023 e 2025, além da etapa de repescagem de estudos, baseada na identificação de autores relevantes na área.

Após a aplicação dessas etapas, foi obtido um portfólio preliminar composto por 19 artigos, que foram posteriormente submetidos à leitura dos resumos para verificação final do alinhamento temático.

Como resultado desse processo, foi constituído um portfólio bibliográfico final composto por 15 artigos científicos, que passaram a representar a base de referência utilizada na análise bibliométrica e na discussão das aplicações de SOM em estudos de qualidade da água.

O procedimento adotado para a construção do portfólio bibliográfico é apresentado na **Figura 1**, que sintetiza as principais etapas do processo de seleção dos estudos.

Figura 1 – Seleção do portfólio bibliográfico pelo método ProKnow-C e análise bibliométrica.



Fonte: Autora (2026)

A partir do portfólio bibliográfico selecionado, foram realizadas análises bibliométricas utilizando o pacote Bibliometrix, permitindo identificar padrões de publicação, redes de colaboração científica, coocorrência de palavras-chave e evolução temática da literatura. Os principais resultados dessa análise são apresentados e discutidos na seção seguinte.

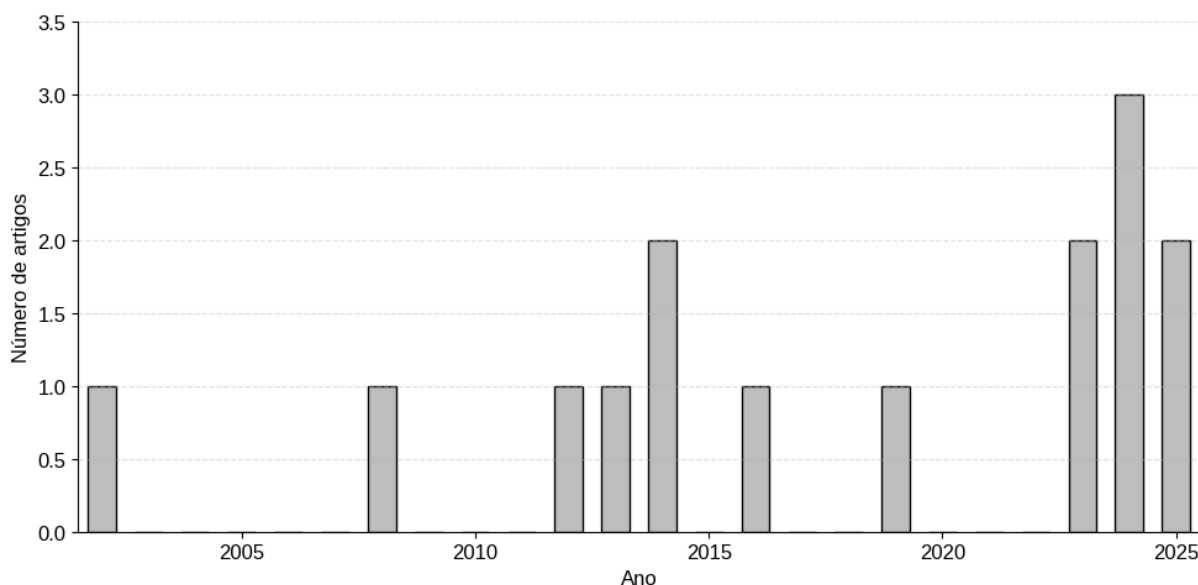
2.4.2 Panorama da produção científica sobre SOM e qualidade da água

Inicialmente, foi analisada a evolução temporal da produção científica relacionada à aplicação de Self-Organizing Maps (SOM) em estudos de qualidade da água, considerando o ano de publicação dos artigos que compõem o portfólio bibliográfico selecionado. Observa-se que a produção científica apresenta registros pontuais nas primeiras décadas analisadas, com

publicações esporádicas no início dos anos 2000 e ao longo da década de 2010.

No entanto, verifica-se um aumento mais consistente no número de estudos publicados a partir da década de 2020, com maior concentração de artigos nos anos mais recentes. Esse comportamento indica o recente fortalecimento do interesse científico pela utilização de SOM como ferramenta de análise de dados ambientais, especialmente em estudos voltados ao monitoramento, à classificação e à avaliação da qualidade da água em diferentes sistemas aquáticos.

Figura 2 – Evolução temporal da produção científica sobre aplicação de SOM na análise da qualidade da água



Fonte: Autora (2026)

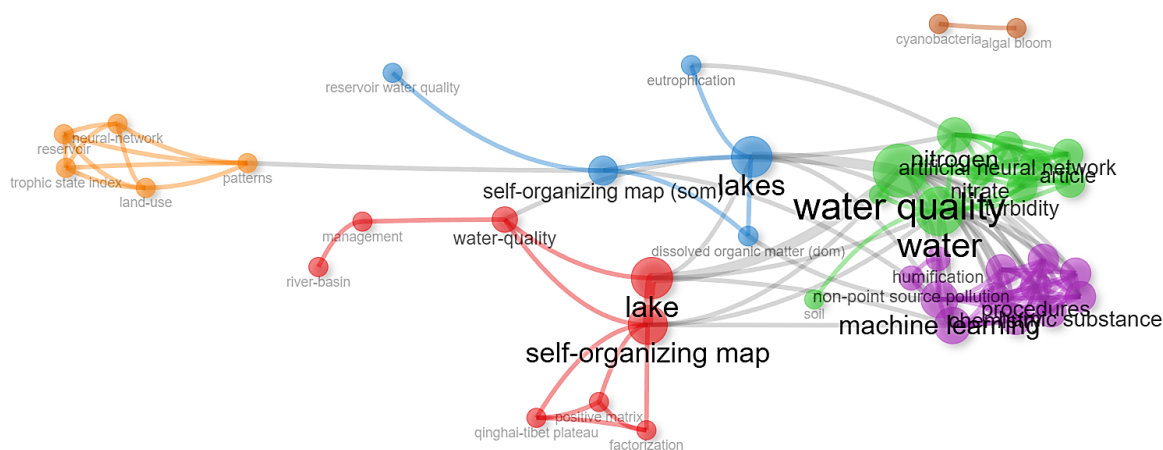
Além da evolução temporal das publicações, foi analisada a estrutura temática da literatura por meio da rede de coocorrência de palavras-chave identificadas nos estudos que compõem o portfólio bibliográfico. Esse tipo de análise permite visualizar os principais temas investigados na área e as relações entre os conceitos abordados nos trabalhos científicos.

Observa-se que os termos *water quality*, *water*, *lake*, *lakes* e *self-organizing map* ocupam posições centrais na rede, indicando sua relevância na estrutura temática da literatura analisada. Esses termos aparecem associados a variáveis frequentemente utilizadas em estudos de qualidade da água, como *nitrogen*, *nitrate*, *turbidity* e *eutrophication*, evidenciando o foco dos estudos na análise de parâmetros físico-químicos associados à dinâmica da qualidade da água.

Além disso, a presença de termos como *machine learning*, *artificial neural network* e

non-point source pollution indica a integração crescente entre técnicas computacionais e análises ambientais. Esses resultados sugerem que a aplicação de SOM tem sido amplamente empregada para identificar padrões em conjuntos de dados ambientais multivariados, contribuindo para a compreensão de processos ecológicos e para o monitoramento da qualidade da água em ambientes lacustres e reservatórios.

Figura 3 – Rede de coocorrência das palavras-chave associadas aos estudos sobre aplicação de SOM na análise da qualidade da água



Fonte: Autora (2026)

Complementarmente, foi analisado o mapa temático da literatura, que permite identificar os principais temas de pesquisa com base nos níveis de centralidade e de densidade das palavras-chave. Essa abordagem permite compreender o grau de desenvolvimento e a relevância dos diferentes tópicos investigados no campo científico.

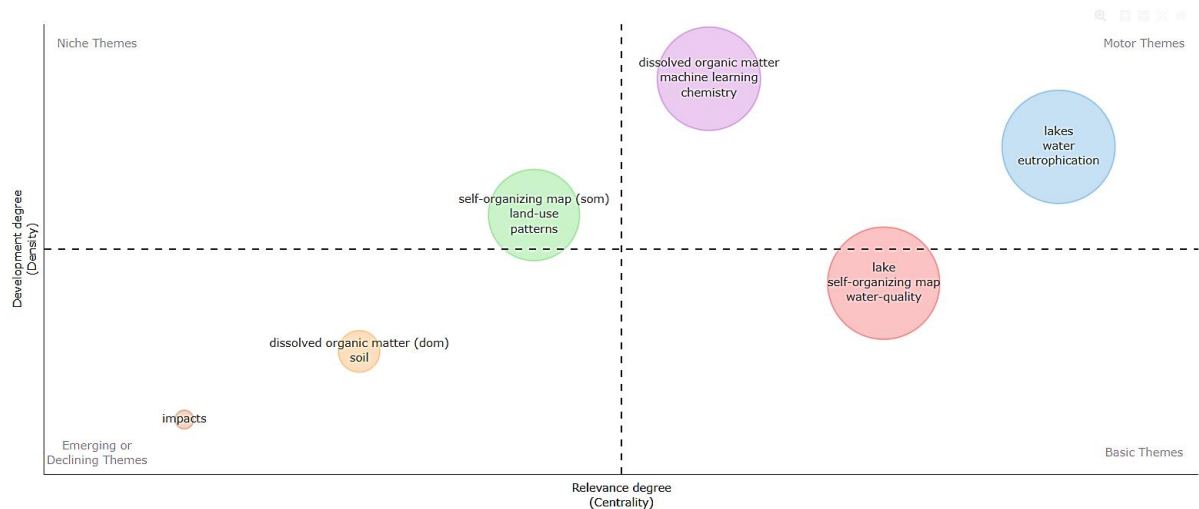
Observa-se que o cluster formado pelos termos lakes, water e eutrophication está localizado na região de temas motores, indicando elevada relevância e desenvolvimento na literatura científica. Esses temas representam linhas consolidadas de investigação voltadas à análise da qualidade da água em ambientes lacustres e reservatórios.

Por sua vez, o cluster composto por lake, self-organizing map e water-quality aparece como tema básico, evidenciando sua importância estrutural no campo de pesquisa. Esses termos refletem o núcleo conceitual das investigações que aplicam técnicas de aprendizado de máquina, em particular o SOM, na análise de dados ambientais.

Adicionalmente, temas relacionados a dissolved organic matter, machine learning e chemistry apresentam elevado grau de desenvolvimento, caracterizando-se como temas especializados na área. Finalmente, termos como soil, dissolved organic matter (DOM) e

impacts aparecem na região de temas emergentes ou em declínio, sugerindo linhas de investigação ainda em desenvolvimento ou com menor consolidação na literatura científica.

Figura 4 – Mapa temático da literatura científica sobre aplicação de SOM em estudos de qualidade da água.



Fonte: Autora (2026)

De forma geral, os resultados da análise bibliométrica indicam que a aplicação de Self-Organizing Maps em estudos de qualidade da água, especialmente em ambientes lacustres e reservatórios, constitui um eixo central de investigação na literatura científica recente. Observa-se, ainda, uma integração crescente entre abordagens de aprendizado de máquina e análises ambientais, reforçando o potencial dessas técnicas para identificar padrões complexos em conjuntos de dados limnológicos.

Nesse contexto, o uso de redes neurais do tipo Self-Organizing Map apresenta-se como uma ferramenta promissora para a análise exploratória de dados ambientais complexos e para a identificação de padrões na qualidade da água.

CAPÍTULO III

METODOLOGIA

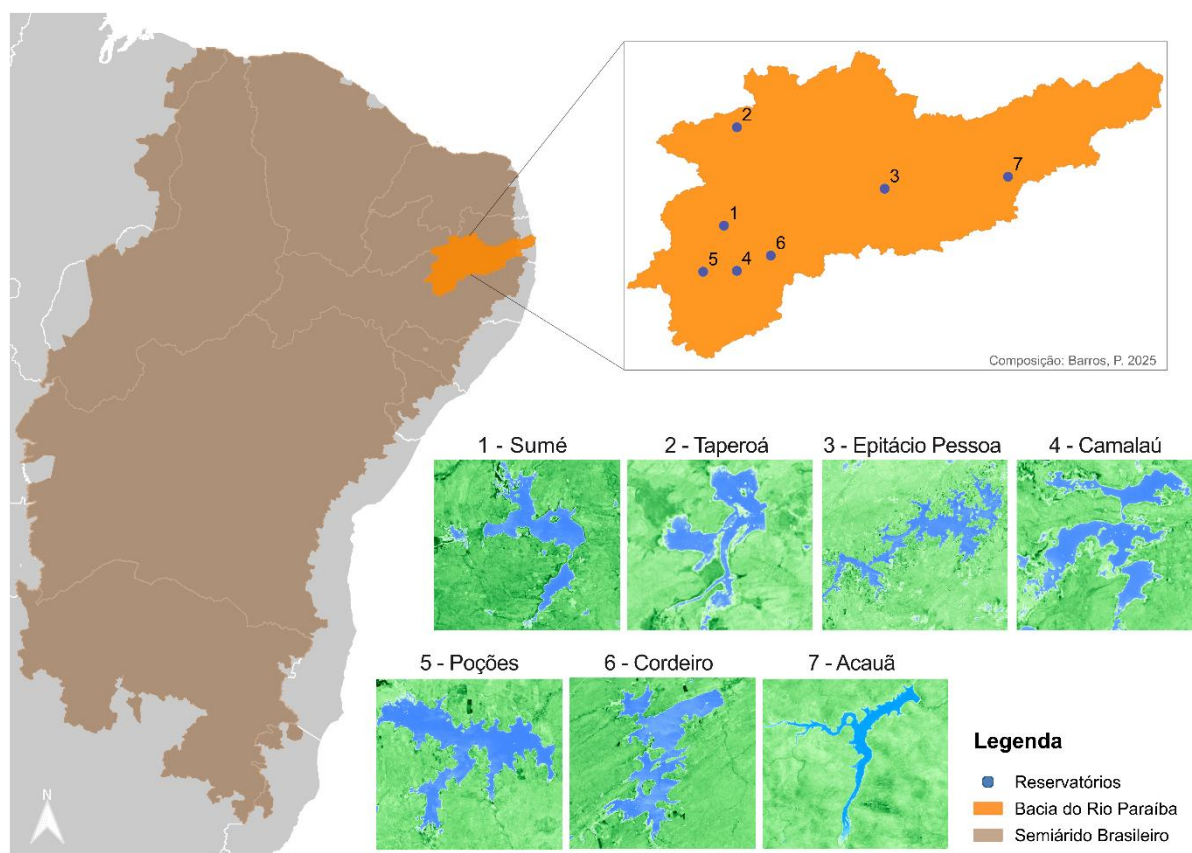
Neste capítulo, são descritos os procedimentos metodológicos adotados na pesquisa, incluindo a caracterização da área de estudo, a aquisição e o tratamento dos dados de qualidade da água, as análises estatísticas exploratórias e a modelagem por meio de redes neurais do tipo Self-Organizing Map (SOM). A aplicação do SOM foi realizada em duas abordagens complementares: individualmente, para cada reservatório, e de forma integrada, considerando o conjunto dos sistemas, com o objetivo de permitir análises comparativas e identificar padrões espaciais e sazonais na qualidade da água.

3 METODOLOGIA

3.1 ÁREA DE ESTUDO

Este estudo foi realizado em sete reservatórios situados no semiárido brasileiro, denominados Poções, Camalaú, Epitácio Pessoa, Sumé, Cordeiro, Taperoá e Acauã. Conforme ilustrado na **Figura 5**, todos esses reservatórios estão inseridos na Bacia do Rio Paraíba.

Figura 5 – Reservatórios selecionados para o estudo



Fonte: Autora (2025)

3.1.1 Caracterização da Bacia Hidrográfica

A Bacia do Rio Paraíba abrange uma área aproximada de 20.000 km², estendendo-se entre as latitudes 6°51'31" e 8°26'21" sul e as longitudes 34°48'35" e 37°2'15" oeste. Trata-se da segunda maior bacia hidrográfica do Estado da Paraíba, correspondendo a cerca de 38% do território estadual. Considerada uma das mais importantes do semiárido nordestino, é composta pela sub-bacia do Rio Taperoá e pelas regiões do Alto, Médio e Baixo cursos do Rio Paraíba (AESA, 2025). A bacia hidrográfica do Rio Paraíba desempenha papel fundamental nas atividades agrícolas, contribuindo significativamente para o desenvolvimento socioeconômico

em âmbito local e regional. A região caracteriza-se pela predominância da agropecuária, cujas práticas intensivas exercem pressão sobre os recursos naturais, aumentando sua vulnerabilidade (Silva et al., 2023).

3.1.2 Caracterização dos Reservatórios

3.1.2.1 Sumé

O Reservatório de Sumé está localizado nas coordenadas geográficas 07°40'18,96" S e 36°53'59,84" W (ANA, 2017). Implantado sobre o leito do rio Sucuru, sua construção foi iniciada em 1957 e concluída em 1962 pelo Departamento Nacional de Obras Contra as Secas (DNOCS), no contexto das políticas públicas voltadas ao enfrentamento das secas no semiárido nordestino.

Originalmente, o reservatório foi projetado para atender ao abastecimento público do município de Sumé e ao Perímetro Irrigado de Sumé (PIS). Entre as décadas de 1970 e 1980, o sistema desempenhou papel relevante no desenvolvimento agropecuário local, sobretudo com a implantação de um perímetro irrigado que totalizava 705 hectares, dos quais 273,65 hectares foram efetivamente estruturados. O método de irrigação predominante era por sulcos com escoamento por gravidade, favorecendo principalmente o cultivo intensivo de tomate. Contudo, com a posterior desativação do perímetro irrigado, observou-se declínio na dinâmica produtiva da região (Cantalice, 2010).

Em termos de capacidade de armazenamento, dados recentes indicam alteração no volume máximo acumulável. Em 2021, a Agência Nacional de Águas realizou levantamento batimétrico por meio de técnicas de sensoriamento remoto. E o estudo apontou redução da capacidade máxima de 44,86 hm³, valor originalmente informado pelo DNOCS, para 37,09 hm³, representando decréscimo de aproximadamente 17,3% no volume total do reservatório (ANA, 2021). Essa diminuição implica maior vulnerabilidade do sistema quanto à disponibilidade hídrica.

Atualmente, o reservatório mantém papel estratégico para a região, atendendo demandas relacionadas ao abastecimento humano, à agricultura e à pecuária, especialmente durante períodos de estiagem recorrentes no semiárido. Além disso, contribui para a regularização do regime hidrológico local, mitigando os efeitos das secas e garantindo suporte às atividades econômicas e às comunidades dependentes desse recurso (Filho, 2024).

3.1.2.2 Taperoá

O Reservatório Taperoá, oficialmente denominado Manoel Marcionilo, constitui o principal manancial de abastecimento do município de Taperoá – PB. Localiza-se nas coordenadas 07°12'29,77" S e 36°50'19,54" W e possui capacidade máxima de aproximadamente 15,15 milhões de metros cúbicos (ANA, 2017).

Sua construção ocorreu durante a gestão do governador Wilson Braga (1983–1986), período em que a política de açudagem foi amplamente difundida no estado da Paraíba como estratégia de convivência com a seca no semiárido. Conforme relatado por Araújo (2023), no âmbito do Projeto Canaã foi planejada a construção de 62 açudes de médio e grande porte no estado, entre os quais se inclui o reservatório Manoel Marcionilo.

Além de abastecer a população urbana de Taperoá, o reservatório também atende demandas da zona rural, sendo utilizado para dessedentação animal, consumo humano e atividades agrícolas, como o cultivo de capim (Farias, 2012).

Entretanto, a qualidade da água do reservatório pode ser influenciada por fatores externos. Segundo Fernandes (2022), a ausência de tratamento adequado de esgotos no município de Desterro – PB favorece o lançamento de efluentes sem tratamento nos corpos hídricos da região. Considerando que Desterro está localizado a montante de Taperoá dentro do sistema de drenagem local, os fluxos hídricos tendem a conduzir esses poluentes em direção às áreas situadas a jusante. Durante o período chuvoso, o aumento do escoamento superficial intensifica o transporte de matéria orgânica, microrganismos e resíduos para os cursos d'água, potencializando o transporte dessas cargas até o reservatório. Esse processo pode elevar a turbidez e a carga orgânica e microbiológica da água, ampliando o risco de comprometimento da sua qualidade.

3.1.2.3 Epitácio Pessoa

O reservatório Epitácio Pessoa, popularmente conhecido como Boqueirão, denominação associada ao município onde se encontra inserido, está inserido na região do Alto Curso do Rio Paraíba e constitui um dos principais reservatórios do estado. Localizado nas coordenadas 07°29'52,39" S e 36°08'30,74" W, o manancial foi construído na década de 1950 e é operado pelo Departamento Nacional de Obras Contra as Secas (DNOCS) (ANA, 2017).

Com capacidade máxima de armazenamento de aproximadamente 466,53 hm³, é o segundo maior reservatório da Paraíba em volume acumulado. Atualmente, encontra-se com cerca de 37,94% de sua capacidade total (ANA, 2026), conforme dados do Sistema de

Acompanhamento de Reservatórios.

O reservatório possui importância estratégica para o sistema de abastecimento regional, atendendo 18 municípios do estado da Paraíba. Entre as cidades beneficiadas destaca-se Campina Grande, reconhecida como a segunda mais populosa do estado, o que evidencia a relevância do manancial para a segurança hídrica e para a dinâmica socioeconômica da região.

O reservatório enfrentou, nas últimas décadas, dois períodos prolongados de escassez hídrica, compreendidos entre 1997–2003 e 2012–2017. Embora a baixa ocorrência de precipitações tenha sido o principal fator desencadeador dessas crises, limitações na gestão dos recursos hídricos contribuíram para intensificar seus impactos. O segundo episódio foi particularmente mais severo, chegando a colocar em risco o abastecimento público no ano de 2017. A situação somente foi estabilizada com a antecipação das obras do Projeto de Integração do Rio São Francisco (PISF), que representou, naquele momento, a principal alternativa emergencial para reduzir os efeitos da escassez (Silva; Ribeiro, 2023).

Além de desempenhar papel estratégico no abastecimento público, o reservatório atende diversas demandas das comunidades situadas em sua área de influência. Suas águas são destinadas ao consumo humano e também sustentam atividades produtivas, como a irrigação agrícola e a dessedentação de rebanhos. Ademais, o manancial possui relevância social ao possibilitar práticas recreativas e de lazer. Dessa forma, configura-se como recurso hídrico de uso múltiplo, fundamental para a dinâmica socioeconômica (Barbosa, 2025).

3.1.2.4 Camalaú

O reservatório Camalaú está inserido na região do Alto Paraíba, pertencente à bacia hidrográfica do Rio Paraíba. Localiza-se nas coordenadas 07°53'5,60" S e 36°50'20,50" W e é operado pela Companhia de Água e Esgotos da Paraíba (CAGEPA) (ANA, 2017). O reservatório possui capacidade máxima de 48,11 hm³ e, de acordo com os últimos dados do Sistema de Acompanhamento de Reservatórios, em fevereiro de 2026, apresenta volume correspondente a 74,81% de sua capacidade total (ANA, 2026).

Esse reservatório foi implantado em 1982 e, ao longo do tempo, passou por intervenções estruturais relevantes. Em 2017, foram executadas obras de requalificação que envolveram reforços nos maciços das estruturas principal e auxiliar, intervenções no sistema extravasor, melhorias na tomada d'água existente e a implantação de uma estrutura complementar destinada à integração com o Projeto de Integração do Rio São Francisco (PISF) (PARAÍBA, 2020).

De acordo com Vasconcelos (2013), o manancial atende a múltiplos usos, como

abastecimento humano, aquicultura de subsistência, dessedentação de rebanhos, irrigação e atividades recreativas. Além disso, integra o eixo leste da cascata de reservatórios beneficiados pela transposição das águas do Rio São Francisco, sendo o segundo a receber aporte hídrico nesse sistema.

3.1.2.5 Poções

O Açude Poções está localizado no município de Monteiro, inserido no Alto Curso do Rio Paraíba, integrante da Bacia Hidrográfica do Rio Paraíba. Situa-se nas coordenadas 07°53'20,47" S e 36°59'53,05" W e é operado pelo Departamento Nacional de Obras Contra as Secas (DNOCS). O manancial atende a múltiplos usos, incluindo abastecimento urbano e rural, irrigação e dessedentação animal, desempenhando papel relevante para o abastecimento e o desenvolvimento socioeconômico regional (ANA, 2017).

Construído na década de 1980, o reservatório possui capacidade máxima de 29,86 hm³. Em fevereiro de 2026, atingiu seu volume máximo, encontrando-se atualmente com 100% da capacidade armazenada (ANA, 2026).

O reservatório possui relevância significativa no contexto da bacia hidrográfica em que está inserido. Embora não figure entre os maiores mananciais do estado da Paraíba, desempenha função estratégica por compor o sistema integrado de reservatórios da região, conectando-se a estruturas como Camalaú, Epitácio Pessoa e Acauã. A partir de 2017, passou a ser abastecido pelas águas provenientes do Projeto de Integração do Rio São Francisco (PISF), por meio do eixo leste. Nesse arranjo, destaca-se como o segundo reservatório paraibano a receber esse aporte hídrico e assume papel fundamental na redistribuição desses recursos aos demais reservatórios do sistema (Silva et al., 2023).

3.1.2.6 Cordeiro

Localizado no município do Congo, na região do Alto Curso do Rio Paraíba, integrante da Bacia Hidrográfica do Rio Paraíba, o reservatório apresenta capacidade máxima de armazenamento de 69,97 hm³. Em fevereiro de 2026, encontra-se com 38,47% do volume total acumulado, conforme dados da ANA (2026).

Situado nas coordenadas 07°48'46,95" S e 36°40'44,18" W, o reservatório teve sua construção concluída em 1988. A operação e a gestão do manancial são de responsabilidade da Agência Executiva de Gestão das Águas do Estado da Paraíba (AESPA), órgão gestor dos recursos hídricos estaduais (ANA, 2017).

O reservatório desempenha função central no suprimento hídrico do município do Congo, onde está situado, e integra um sistema adutor responsável pela distribuição de água para outras cidades da região, como Monteiro, Sumé, Serra Branca e São João do Cariri. Além do atendimento ao abastecimento público, o manancial também sustenta atividades agrícolas por meio da irrigação, garante a dessedentação animal e possibilita práticas como a pesca, contribuindo para a manutenção das atividades produtivas e da subsistência local (Cruz et al., 2018).

3.1.2.7 Acauã

A Barragem Argemiro de Figueiredo, amplamente conhecida como Reservatório Acauã, está situado no município de Itatuba, nas coordenadas 07°26'28,73" S e 35°33'36,68" W. Concluído em 2002, é o mais recente entre os reservatórios analisados nesta pesquisa. O manancial possui capacidade máxima de armazenamento de 253,00 hm³ e tem como órgão responsável pela operação e gestão a AESA. Em fevereiro de 2026, o reservatório encontra-se com 51,47% de sua capacidade total armazenada (ANA, 2017).

Em termos de volume, destaca-se no contexto estadual por figurar entre os maiores reservatórios da Paraíba. Conforme Lemos (2019), sua capacidade o posiciona atrás apenas do Complexo Coremas Mae d'Água, do Açude Epitácio Pessoa e da Barragem Engenheiro Ávidos, ocupando, assim, a quarta colocação em volume máximo de armazenamento no estado.

O reservatório exerce funções múltiplas e estratégicas, além do atendimento ao abastecimento público e rural, suas águas sustentam atividades como irrigação agrícola e aquicultura, desempenhando papel relevante na produção de alimentos, na geração de renda e na dinâmica econômica local (AESAs, 2025). Dessa maneira, o Acauã constitui elemento essencial na regulação hídrica e no suporte às demandas sociais e produtivas da região onde está inserido.

3.2 AQUISIÇÃO DE DADOS

3.2.1 Base de Dados

Os dados de qualidade da água dos sete reservatórios estudados foram fornecidos pelo Projeto Ecológico de Longa Duração do Rio Paraíba Integrado (PELD-RIPA), integrante do Programa de Pesquisa Ecológica de Longa Duração coordenado e financiado pelo Conselho

Nacional de Desenvolvimento Científico e Tecnológico (CNPq). O Programa PELD tem como finalidade apoiar sítios de pesquisa ecológica em diferentes biomas brasileiros, promovendo a produção de informações que possibilitem compreender os efeitos de perturbações naturais e antrópicas sobre a composição, a dinâmica e o funcionamento dos ecossistemas (CNPq, 2020). Caracteriza-se pela atuação integrada de equipes interdisciplinares, envolvendo pesquisadores de diversas áreas e promovendo articulação com gestores e outros atores sociais, fortalecendo o diálogo entre ciência e sociedade.

No âmbito das atividades do PELD-RIPA, são desenvolvidas múltiplas linhas de investigação na Bacia do Rio Paraíba, incluindo o monitoramento contínuo de reservatórios, contemplando variáveis físicas, químicas e biológicas fundamentais para a avaliação dos processos ecológicos que estruturam esses ecossistemas. A utilização dessa base de dados assegura padronização metodológica, consistência analítica e representatividade temporal e espacial das informações, permitindo análises mais robustas acerca da variabilidade da qualidade da água nos sistemas estudados.

3.2.2 Variáveis Analisadas

Os dados utilizados neste estudo compreendem o período de dezembro de 2021 a agosto de 2024, sendo provenientes de duas campanhas de amostragem realizadas anualmente, representativas dos períodos chuvoso e seco da região semiárida. As coletas ocorreram em duas regiões distintas dos reservatórios: a região limnética e a região litorânea.

Na região limnética, as amostragens foram realizadas em três zonas longitudinais do reservatório: P1 – zona lântica (região próxima à barragem); P2 – zona de transição (entre a barragem e a entrada do rio); e P3 – zona lótica (próxima à entrada do rio). Em cada uma dessas zonas, as coletas ocorreram em três profundidades distintas: superfície (0,5 m), profundidade correspondente a 1% da incidência luminosa (limite inferior da zona eufótica, representando a profundidade máxima de penetração efetiva de luz) e fundo (profundidade máxima local).

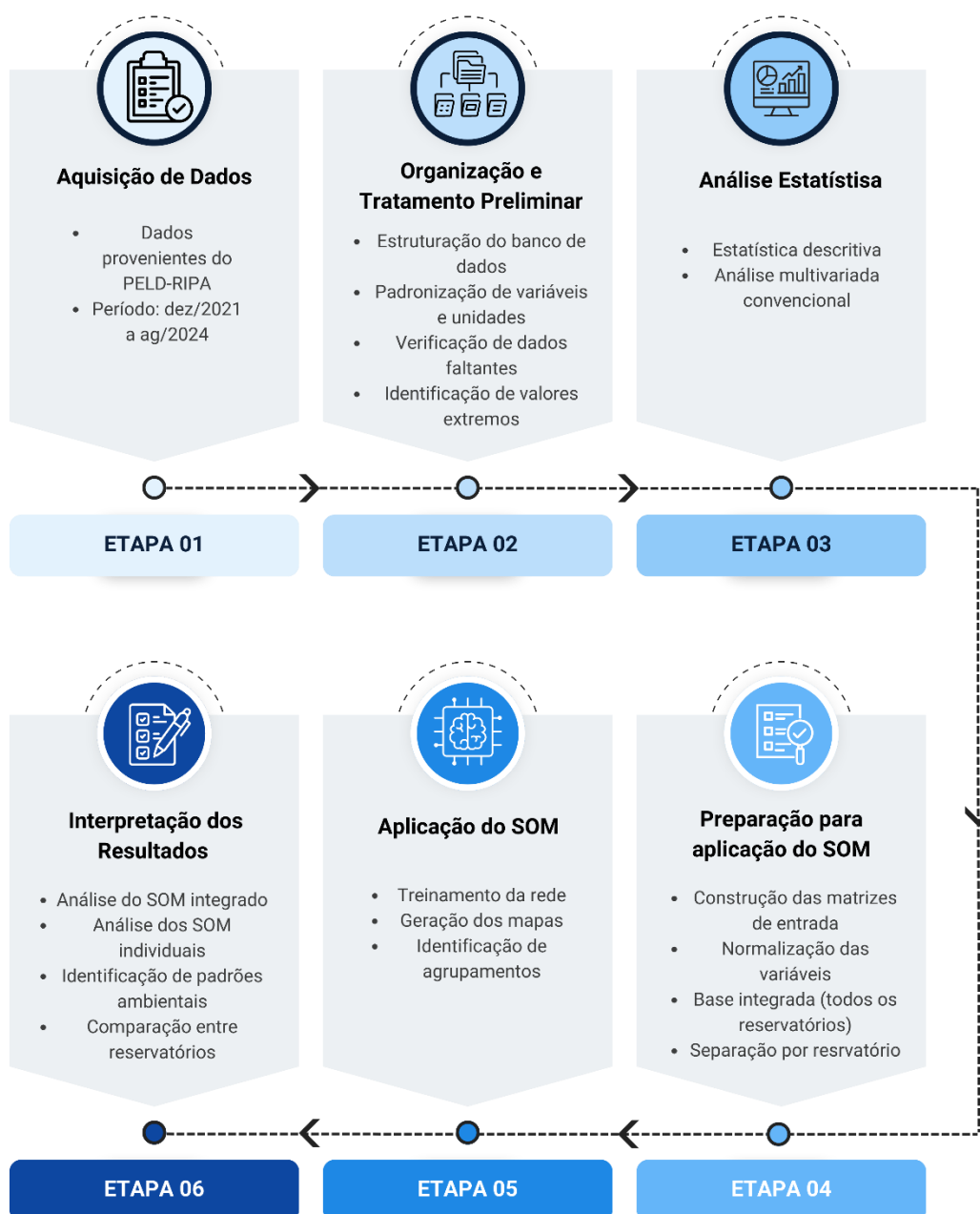
Na região litorânea, foram estabelecidos aproximadamente 15 pontos de amostragem, distribuídos de modo a abranger, tanto quanto possível, toda a extensão marginal dos reservatórios. Em alguns casos, o número de pontos variou entre 15 e 17, em função das condições de acesso e das características morfológicas de cada sistema.

As variáveis analisadas incluíram: temperatura da água, pH, potencial redox, condutividade elétrica, turbidez, oxigênio dissolvido, sólidos totais dissolvidos, salinidade, transparência da água, amônia, nitrito, nitrato, fósforo total e clorofila total.

3.2.3 Tratamento Preliminar

Com o objetivo de sistematizar as etapas analíticas desenvolvidas nesta pesquisa, a **Figura 6** apresenta o fluxograma metodológico adotado, desde a aquisição e organização dos dados até a aplicação do algoritmo SOM e interpretação dos resultados.

Figura 6 – Fluxograma das etapas metodológicas do estudo



Fonte: Autora (2026).

Conforme ilustrado na **Figura 6**, o desenvolvimento metodológico foi estruturado de

forma sequencial, iniciando-se pela organização e controle de qualidade do banco de dados, seguido das análises estatísticas e, posteriormente, da aplicação do SOM. As etapas referentes ao tratamento preliminar dos dados são detalhadas a seguir.

A etapa de tratamento preliminar teve como finalidade assegurar a consistência, confiabilidade e adequação do banco de dados às técnicas analíticas empregadas. Inicialmente, os dados brutos foram estruturados em planilhas eletrônicas, nas quais cada linha representou uma unidade amostral (reservatório, ponto de coleta, profundidade e campanha), e cada coluna correspondeu a uma variável físico-química ou biológica. Nessa fase, procedeu-se à padronização da nomenclatura das variáveis, das unidades de medida e da formatação numérica, garantindo uniformidade ao conjunto de dados.

Em seguida, realizou-se a seleção das variáveis de interesse, considerando os objetivos da pesquisa e a aplicabilidade das técnicas estatísticas multivariadas. Algumas informações de caráter mais específico, especialmente relacionadas à composição detalhada do fitoplâncton (como a discriminação entre grupos de algas verdes, pardas e cianobactérias), não foram incluídas na análise. Optou-se pela utilização da clorofila total, por representar um indicador sintético da biomassa fitoplanctônica. De forma semelhante, foi selecionado o fósforo total em detrimento do fósforo solúvel reativo, considerando-se sua maior abrangência como indicador do potencial de enriquecimento nutricional do sistema.

Posteriormente, foi conduzida a verificação de dados faltantes. Registros incompletos foram analisados individualmente e, quando necessário, excluídos da base final utilizada nas análises, de modo a evitar distorções nos resultados estatísticos e no treinamento da rede neural. Valores iguais a zero foram mantidos quando compatíveis com a natureza da variável. Valores negativos, especialmente associados ao potencial redox, foram considerados fisicamente aceitáveis quando coerentes com as condições ambientais observadas. Nos casos em que inconsistências analíticas foram identificadas, procedeu-se à exclusão pontual dos registros.

A análise de completude do banco de dados indicou elevada frequência de registros ausentes para a variável Transparência da Água (185 valores faltantes), o que implicaria redução expressiva do conjunto amostral caso fosse mantida na modelagem multivariada. Considerando a presença de outras variáveis indicadoras das condições ópticas e tróficas do sistema, optou-se pela exclusão dessa variável da etapa de modelagem com SOM, preservando maior representatividade amostral.

Após a exclusão da variável Transparência da Água e a remoção dos registros com dados faltantes nas demais variáveis selecionadas, a base final utilizada na modelagem integrada foi composta por 854 amostras e 13 variáveis ambientais.

Por fim, realizou-se análise exploratória com o objetivo de identificar valores extremos (outliers), por meio de inspeção gráfica e análise estatística descritiva. Considerando a elevada variabilidade natural dos sistemas aquáticos em regiões semiáridas, valores extremos ecologicamente plausíveis foram mantidos na base de dados, sendo descartados apenas aqueles inconsistentes ou incompatíveis com as características do sistema, quando identificados.

3.3 ANÁLISE ESTATÍSTICA

A análise estatística foi conduzida com o objetivo de caracterizar o comportamento das variáveis de qualidade da água e identificar padrões de variação, constituindo uma etapa exploratória anterior à aplicação do algoritmo Self-Organizing Map (SOM). As análises foram realizadas em ambiente computacional, utilizando a linguagem de programação Python no Google Colaboratory (GOOGLE, 2023), ferramenta que possibilita o processamento e a análise de dados ambientais multivariados.

3.3.1 Estatística Descritiva

Inicialmente, realizou-se análise estatística descritiva das variáveis selecionadas, com o cálculo de medidas de tendência central (média e mediana) e de dispersão (desvio padrão, coeficiente de variação, valores mínimos e máximos). O coeficiente de variação foi calculado conforme a Equação (1):

$$CV = (\sigma/\mu) \times 100 \quad \text{Equação (1)}$$

Onde σ representa o desvio padrão e μ a média da variável analisada.

Essa etapa permitiu caracterizar o comportamento geral dos parâmetros físico-químicos e biológicos, bem como avaliar a amplitude de variação espacial e sazonal entre os reservatórios. Adicionalmente, foi calculado o coeficiente de correlação de Pearson para avaliar a intensidade e direção da relação linear entre as variáveis, conforme a Equação (2):

$$r = \Sigma[(x_i - \bar{x})(y_i - \bar{y})] / \sqrt{[\Sigma(x_i - \bar{x})^2 \Sigma(y_i - \bar{y})^2]} \quad \text{Equação (2)}$$

A matriz de correlação foi representada graficamente por meio de um mapa de calor,

permitindo visualizar padrões de associação entre os parâmetros de qualidade da água. Também foram elaborados histogramas e boxplots para examinar a distribuição dos dados e possíveis assimetrias.

3.3.2 Análise Multivariada Convencional

Foi aplicada a Análise de Componentes Principais (Principal Component Analysis – PCA), com o objetivo de reduzir a dimensionalidade do conjunto de dados e identificar padrões de variabilidade. Previamente, à aplicação da PCA, os dados foram padronizados por meio da transformação z-score, conforme a Equação (3):

$$Z = (x - \mu) / \sigma \quad \text{Equação (3)}$$

Onde x representa o valor observado, μ a média e o σ o desvio padrão da variável.

A PCA baseia-se na decomposição da matriz de correlação, gerando componentes principais ortogonais que representam combinações lineares das variáveis originais:

$$PC_1 = a_{11}X_1 + a_{12}X_2 + \dots + a_{1p}X_p \quad \text{Equação (4)}$$

Os componentes com maiores autovalores explicam a maior proporção da variância total do conjunto de dados, sendo considerados relevantes aqueles com autovalores superiores a 1, conforme o critério de Kaiser, além da variância acumulada explicada. Nesta pesquisa, a PCA foi utilizada como ferramenta exploratória complementar à modelagem com SOM, permitindo evidenciar relações entre variáveis e identificar os principais gradientes ambientais. Enquanto a PCA fornece uma representação global baseada em relações lineares, o SOM permite a identificação de padrões não lineares e de estruturas complexas, de modo que ambas as abordagens se complementam na interpretação dos resultados.

3.4 MODELAGEM COM SELF-ORGANIZING MAPS (SOM)

A modelagem foi realizada por meio de Mapas Auto-Organizáveis, propostos por Teuvo Kohonen. O SOM consiste em uma rede neural artificial não supervisionada capaz de projetar dados multidimensionais em uma grade bidimensional, preservando as relações topológicas

entre as amostras. Diferentemente das técnicas estatísticas convencionais, o SOM não pressupõe linearidade entre as variáveis, sendo particularmente adequado para a análise de sistemas ambientais complexos e multivariados.

3.4.1 Preparação dos dados para aplicação do SOM

A matriz de entrada foi estruturada de modo que cada linha representasse uma unidade amostral (reservatório, ponto, profundidade e campanha) e cada coluna correspondesse a uma variável físico-química ou biológica selecionada. O conjunto completo foi inicialmente composto por 1033 amostras. Após a etapa de tratamento preliminar dos dados e exclusão de registros com valores ausentes, optou-se pela retirada da variável Transparência da água, em virtude do elevado número de dados faltantes, resultando em uma base final estruturada com 854 amostras e 13 variáveis ambientais, utilizada tanto no modelo integrado quanto nos modelos individuais por reservatório.

Cada vetor de entrada $X \in \mathbb{R}^p$ foi formado por p variáveis ambientais, sendo que cada neurônio da rede foi representado por um vetor de pesos de mesma dimensionalidade. Considerando as diferentes escalas e unidades de medida, os dados foram padronizados por meio da transformação z-score, conforme a Equação (5):

$$Z_{ij} = \frac{X_{ij} - \mu_j}{\sigma_j} \quad \text{Equação (5)}$$

Onde:

- X_{ij} = valor observado
- μ_j = média da variável
- σ_j = desvio padrão

A padronização assegurou contribuição equivalente das variáveis durante o treinamento da rede.

3.4.2 Configuração do modelo

A implementação do SOM foi realizada em ambiente Python, utilizando o Google Colaboratory (GOOGLE, 2023) e a biblioteca MiniSom, amplamente empregada em aplicações

de aprendizado não supervisionado. O SOM foi estruturado como uma grade bidimensional com topologia hexagonal, selecionada por apresentar melhor preservação das relações de vizinhança entre neurônios.

Cada neurônio k foi representado por um vetor de pesos:

$$w_k = [w_{k1}, w_{k2}, \dots, w_{kp}] \quad \text{Equação (6)}$$

A determinação do número de neurônios da grade bidimensional constitui etapa fundamental na modelagem com SOM, uma vez que influencia diretamente a capacidade de representação da estrutura dos dados e o nível de generalização do modelo (Mia et al., 2023). Neste estudo, o número de neurônios foi estimado com base na heurística proposta na literatura por Vesanto et al. (2000), comumente utilizada (Eid et al., 2025; Farias; Bezerra; Filho, 2015; García; González; Noval, 2003) expressa por:

$$m = 5\sqrt{N} \quad \text{Equação (7)}$$

Onde m corresponde ao número total de neurônios da grade e N representa o número de amostras do conjunto de dados.

Para o modelo integrado, considerando $N = 854$, obteve-se aproximadamente 146 neurônios, sendo adotada uma grade hexagonal 12×12 (144 neurônios), valor próximo ao estimado e considerado adequado para representar a estrutura do conjunto de dados. Nos modelos individuais, o número de neurônios foi determinado separadamente para cada reservatório, aplicando-se a mesma heurística com base no respectivo número de amostras.

Para cada vetor de entrada x , o neurônio vencedor (BMU) foi definido como aquele que minimiza a distância euclidiana:

$$c = \arg \min_k \|x - w_k\| \quad \text{Equação (8)}$$

A atualização dos pesos foi realizada segundo:

$$w_k(t+1) = w_k(t) + \alpha(t)h_{ck}(t)[x(t) - w_k(t)] \quad \text{Equação (9)}$$

A função de vizinhança gaussiana foi definida como:

$$h_{ck}(t) = \exp\left(-\frac{\|r_c - r_k\|^2}{2\sigma^2(t)}\right) \quad \text{Equação (10)}$$

Onde $\sigma(t)$ e $\alpha(t)$ decrescem ao longo das iterações.

3.4.3 Estratégia de aplicação do SOM

Foram adotadas duas abordagens complementares, considerando-se a base final estruturada após o tratamento preliminar dos dados e definição das variáveis utilizadas em cada abordagem:

- (i) Aplicação individual por reservatório, permitindo a análise da variabilidade interna de cada sistema e a identificação de padrões associados às condições ambientais e à dinâmica dos dados em cada um deles.
- (ii) Aplicação integrada, considerando o conjunto completo de 854 amostras após a etapa de tratamento preliminar, excluindo a variável Transparência da água, possibilitando análise comparativa e identificação de similaridades estruturais entre os reservatórios.

As amostras provenientes das zonas limnética e litorânea foram analisadas conjuntamente no processo de treinamento do SOM, sem distinção prévia entre as regiões. Essa abordagem permitiu que a organização dos dados no mapa fosse conduzida exclusivamente pelas similaridades entre os padrões das variáveis ambientais, independentemente da localização espacial das amostras.

No modelo integrado, aplicou-se, adicionalmente, o algoritmo k-means aos vetores de pesos dos neurônios, com o objetivo de delimitar agrupamentos estruturais no mapa. O algoritmo busca particionar os dados em k grupos, minimizando a variabilidade interna de cada cluster e maximizando a separação entre grupos. A clusterização foi realizada com base na minimização da soma das distâncias quadráticas intra-grupo, conforme apresentado na Equação (11).

$$J = \sum_{i=1}^k \sum_{x \in C_i} \|x - \mu_i\|^2 \quad \text{Equação (11)}$$

Onde:

k representa o número de clusters;

C_i corresponde ao conjunto de elementos pertencentes ao cluster i ;

x representa cada vetor de dados (no caso, os vetores de pesos dos neurônios do SOM);

μ_i corresponde ao centroide do cluster i ;

$\|x - \mu_i\|^2$ representa a distância euclidiana quadrática entre o vetor x e o centroide do cluster i .

O número ótimo de clusters foi definido a partir da análise conjunta do Método do Cotovelo (Elbow Method) e do Índice de Silhueta (Silhouette Score), visando identificar a solução que melhor equilibrasse a separação entre os grupos e a coesão interna. A escolha final do número de clusters considerou não apenas os critérios quantitativos, mas também a coerência ambiental dos agrupamentos obtidos, garantindo que os clusters representassem padrões interpretáveis do ponto de vista da qualidade da água. Essa abordagem combinada torna a definição dos agrupamentos mais robusta e defensável, reduzindo a subjetividade associada ao processo de clusterização.

3.4.4 Critérios de avaliação e interpretação

Inicialmente, utilizou-se o Erro de Quantização (Quantization Error – QE) como métrica de desempenho do modelo. O QE mede a distância média entre cada vetor de entrada e o vetor de pesos do neurônio vencedor (Best Matching Unit – BMU), sendo expresso por:

$$QE = \frac{1}{n} \sum_{i=1}^n \|x_i - w_{c(i)}\| \quad \text{Equação (12)}$$

Onde:

- n = número total de amostras
- x_i = vetor de entrada da amostra i
- $w_{c(i)}$ = vetor de pesos do neurônio vencedor

Valores menores de QE indicam uma melhor capacidade de representação do conjunto de dados pelo mapa.

A análise estrutural do SOM foi realizada por meio da U-Matrix (Unified Distance

Matrix), que permite visualizar as distâncias entre neurônios adjacentes. Regiões com maiores distâncias indicam possíveis fronteiras entre agrupamentos distintos, enquanto áreas homogêneas sugerem similaridade entre as amostras projetadas.

A interpretação ambiental foi conduzida a partir dos mapas de componentes (component planes), nos quais cada variável é representada individualmente na grade do SOM. Essa análise possibilitou identificar gradientes ambientais dominantes, relações entre variáveis e padrões associados à dinâmica espacial e sazonal dos reservatórios.

No modelo integrado, os agrupamentos identificados pelo algoritmo k-means foram analisados quanto à coerência ambiental, sendo interpretados como possíveis perfis estruturais de qualidade da água. A distribuição dos reservatórios e das zonas de amostragem dentro desses agrupamentos foi considerada na análise comparativa entre os sistemas.

CAPÍTULO IV

RESULTADOS E DISCUSSÕES

Este capítulo apresenta e discute os principais resultados obtidos a partir das análises realizadas com os dados de qualidade da água dos reservatórios estudados. Inicialmente, apresentam-se as análises estatísticas exploratórias, que fornecem uma visão geral do comportamento das variáveis analisadas. Em seguida, são discutidos os resultados da modelagem por meio das redes neurais do tipo Self-Organizing Map (SOM), considerando a aplicação integrada ao conjunto de reservatórios e, posteriormente, individual para cada sistema, permitindo a identificação de padrões, similaridades e diferenças nas condições ambientais dos reservatórios estudados.

4 RESULTADOS E DISCUSSÕES

4.1 ESTATÍSTICA DESCRITIVA

A análise estatística descritiva foi realizada com o objetivo de caracterizar o comportamento geral das variáveis de qualidade da água consideradas no estudo, com base em 854 amostras válidas. Foram calculadas medidas de tendência central (média e mediana) e medidas de dispersão (desvio padrão, coeficiente de variação, valores mínimos e máximos), conforme descrito na metodologia. Os resultados obtidos estão apresentados na **Tabela 1**, que sintetiza os principais parâmetros estatísticos das variáveis físico-químicas e biológicas analisadas.

Tabela 1 – Estatística descritiva das variáveis de qualidade da água utilizadas na modelagem integrada

Variável	Média	Mediana	Desvio Padrão	Coef. Variação	Mínimo	Máximo
Temperatura da água (°C)	26.98	26.76	1.53	5.68	22.2	36.53
pH	8.08	8.59	1.61	19.87	0.08	10.59
Potencial redox	180.63	185.5	100.16	55.45	-390	566
Condutividade elétrica (mS cm ⁻¹)	1.48	0.5	11.6	781.17	0.14	256
Turbidez (NTU)	43.5	15.95	107.23	246.5	0	1000
Oxigênio dissolvido (mg/L)	7.48	6.68	21.59	288.78	0	630
Sólidos totais dissolvidos (gL ⁻¹)	1.79	0.32	32.09	1793.65	0.05	927
Salinidade (ppt)	0.11	0.04	0.15	147.14	0.01	1.1
Amônia (µg L ⁻¹)	60.74	31.14	157.99	260.11	2.95	1317.05
Nitrito (µg L ⁻¹)	4.56	1.5	14.91	327.09	0	195.91
Nitrato (µg L ⁻¹)	43.24	16.86	143.49	331.83	0	1669.71
Fósforo total (µg L ⁻¹)	90.61	33.67	113.69	125.48	0.33	622
Clorofila total (µg L ⁻¹)	41.27	17.3	55.07	133.46	4.2	796.7

Fonte: Autora (2026)

De modo geral, observou-se baixa variabilidade na temperatura da água (CV = 5,68%) e no pH (CV = 19,87%), indicando relativa estabilidade desses parâmetros ao longo dos reservatórios e do período amostral. Essa menor dispersão sugere que tais variáveis apresentaram comportamento mais homogêneo em comparação com os demais parâmetros avaliados. Por outro lado, a maioria das variáveis apresentou elevada variabilidade (CV > 50%), destacando-se sólidos totais dissolvidos (CV = 1793,65%), condutividade elétrica (CV = 781,17%), nitrito (CV = 327,09%), nitrato (CV = 331,83%) e oxigênio dissolvido (CV = 288,78%).

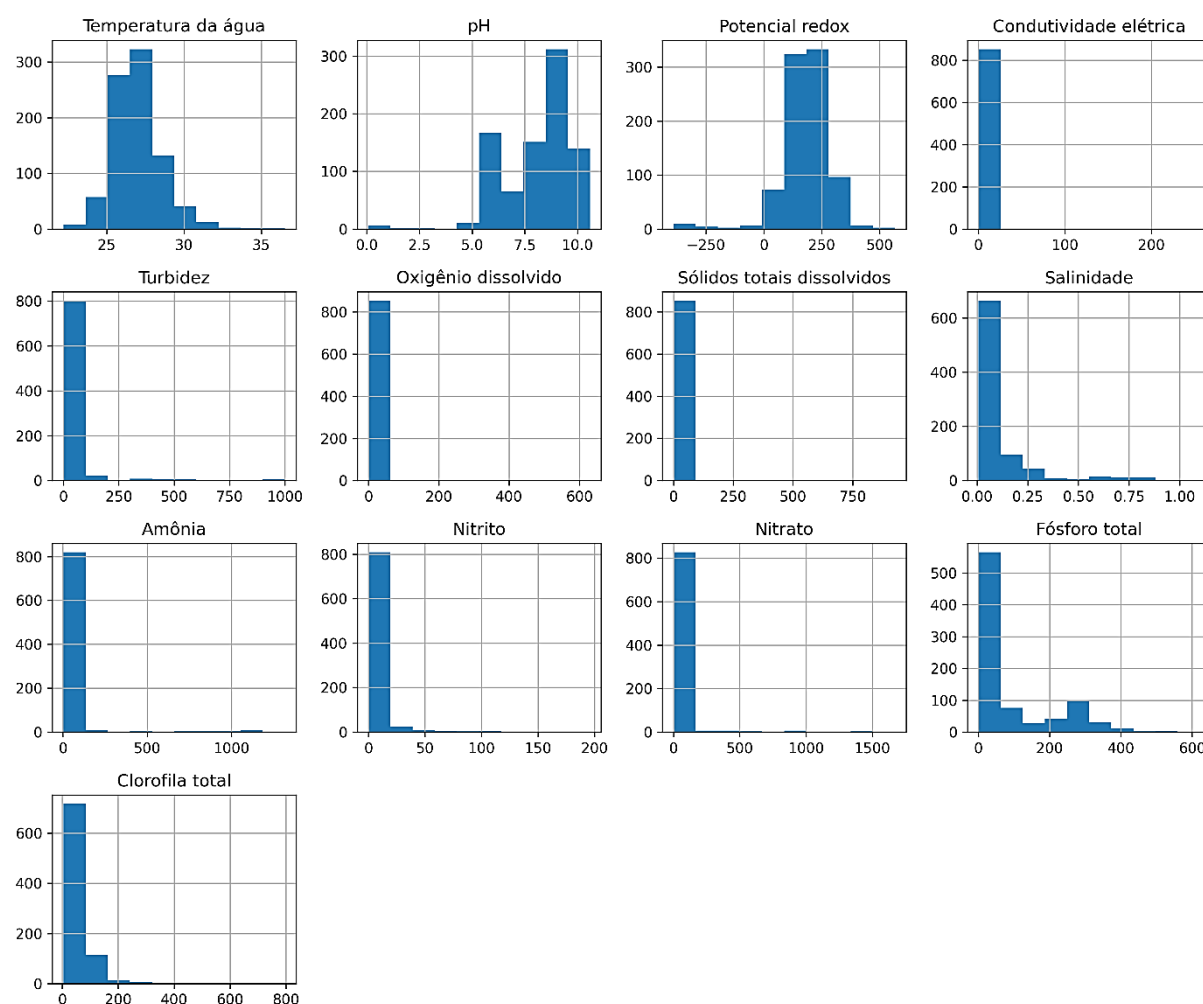
Esses resultados indicam forte heterogeneidade espacial e/ou sazonal, possivelmente

50

associada a diferenças nas variações hidrológicas, no aporte de nutrientes e em processos como nitrificação, mineralização da matéria orgânica e dinâmica do fósforo nos reservatórios. Observa-se ainda diferença expressiva entre a média e a mediana em diversas variáveis, sugerindo distribuições assimétricas e a presença de valores extremos, especialmente nos parâmetros relacionados à carga de nutrientes.

A distribuição dos dados foi analisada por meio de histogramas (**Figura 7**), o que permitiu observar diferentes padrões de comportamento entre os parâmetros analisados.

Figura 7 – Histogramas das variáveis analisadas



Fonte: Autora (2026)

De modo geral, verificou-se que várias variáveis apresentaram distribuições assimétricas, com forte concentração de valores em faixas mais baixas e caudas longas associadas a valores elevados. Esse comportamento foi particularmente evidente em variáveis relacionadas à carga de nutrientes e de material particulado, como turbidez, amônia, nitrato,

51

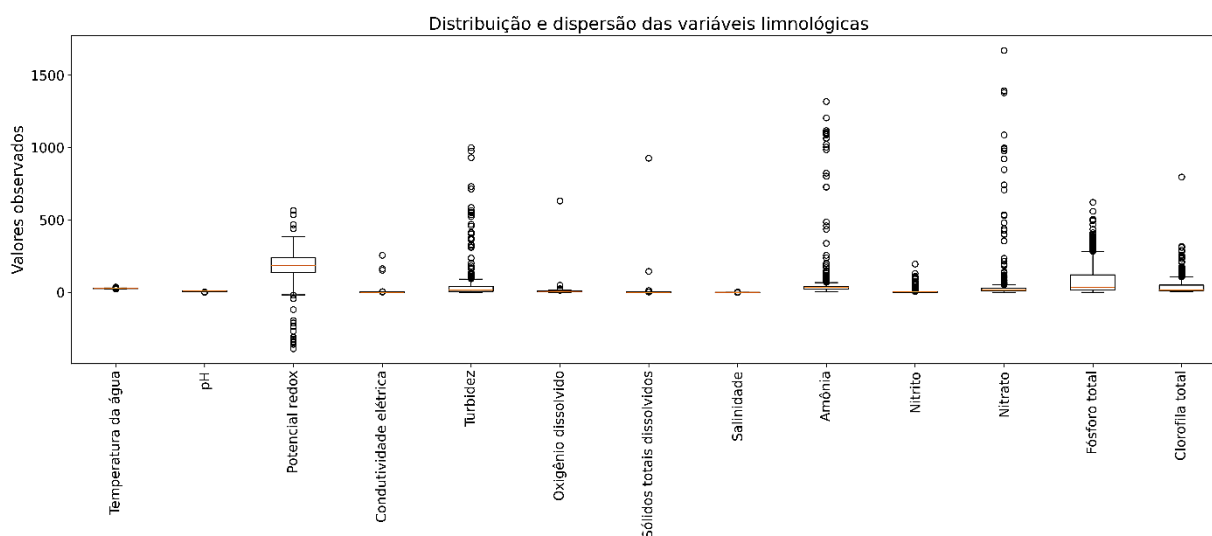
fósforo total e clorofila total, indicando a ocorrência de eventos pontuais de enriquecimento nutricional e elevada variabilidade ambiental nos reservatórios analisados.

Adicionalmente, observou-se a presença de valores extremos em algumas variáveis, destacando-se o oxigênio dissolvido, cuja distribuição apresentou poucos valores significativamente elevados em comparação com o conjunto geral de dados. Esses valores extremos contribuem para o aumento da dispersão estatística e para a diferença observada entre a média e a mediana em algumas variáveis, refletindo a elevada variabilidade espacial e temporal das condições ambientais nos reservatórios analisados.

De maneira geral, os histogramas confirmam os padrões identificados pela análise estatística descritiva, evidenciando elevada variabilidade em diversos parâmetros limnológicos. Além disso, a predominância de distribuições assimétricas indica que as variáveis analisadas não seguem, em sua maioria, uma distribuição normal, característica frequentemente observada em dados ambientais. Esse comportamento reforça a pertinência da aplicação de técnicas multivariadas capazes de lidar com relações não lineares e com conjuntos de dados complexos, como os Mapas Auto-Organizáveis (Self-Organizing Maps – SOM), utilizados nas etapas subsequentes desta pesquisa.

A dispersão dos dados e a presença de valores extremos foram avaliadas por meio de boxplots (**Figura 8**).

Figura 8 – Boxplot das variáveis limnológicas utilizadas na análise estatística descritiva



Fonte: Autora (2026)

Observa-se elevada variabilidade em diversas variáveis limnológicas, especialmente na turbidez, na amônia, no nitrato e no fósforo total. Essas variáveis apresentaram ampla dispersão

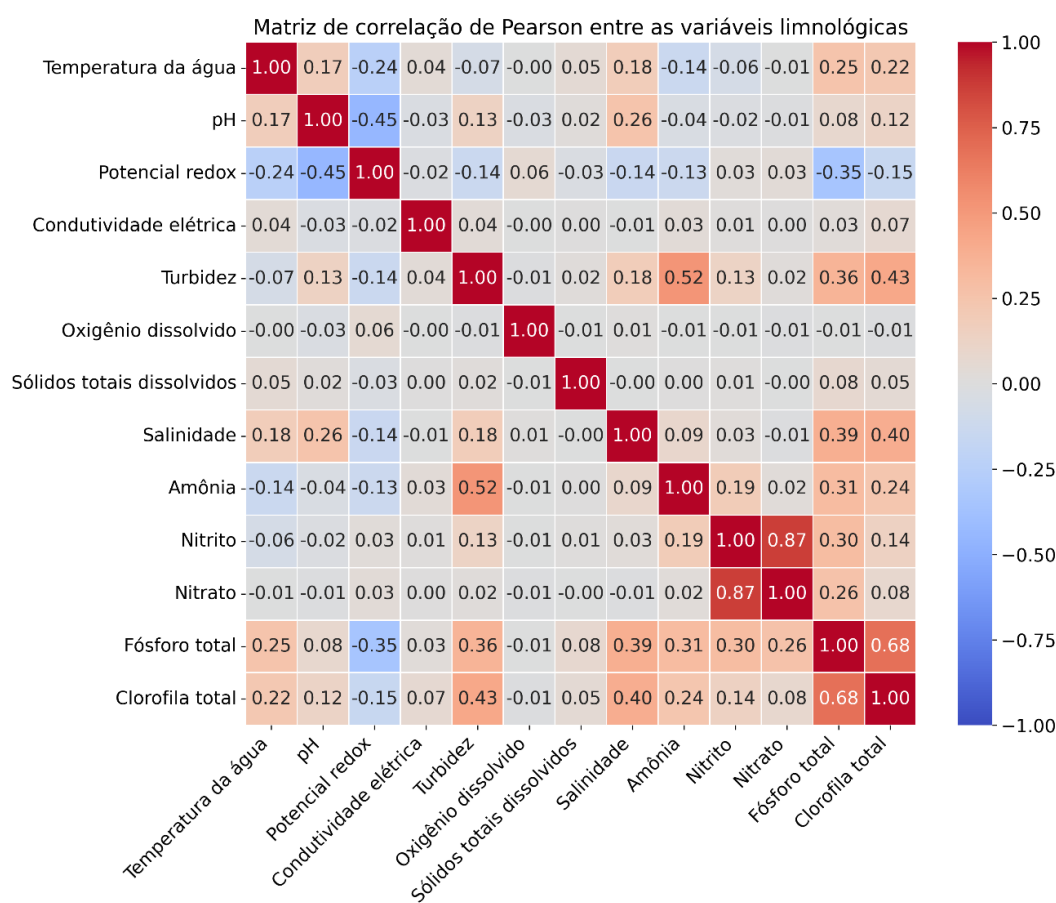
52

e maior frequência de valores extremos, o que indica forte heterogeneidade nas condições ambientais entre as amostras analisadas. Em contraste, variáveis como temperatura da água, pH e salinidade apresentaram menor dispersão relativa, sugerindo maior estabilidade ao longo do conjunto de dados. De modo geral, os boxplots confirmam os padrões identificados na análise estatística descritiva e nos histogramas, evidenciando elevada variabilidade em vários parâmetros de qualidade da água.

4.2 CORRELAÇÃO ENTRE VARIÁVEIS

A correlação de Pearson foi calculada para avaliar a intensidade e a direção das relações lineares entre as variáveis de qualidade da água analisadas. Sua representação gráfica é apresentada na **Figura 9**, em mapa de calor, o que facilita a visualização da intensidade e da direção das associações entre os parâmetros limnológicos.

Figura 9 – Mapa de calor da matriz de correlação de Pearson entre as variáveis de qualidade da água analisadas



Fonte: Autora (2026)

53

Observou-se predominância de correlações fracas a moderadas entre as variáveis, indicando que grande parte dos parâmetros analisados apresenta comportamento relativamente independente no conjunto de dados. Esse resultado sugere a atuação simultânea de múltiplos processos ambientais, reforçando a complexidade da dinâmica da qualidade da água nos reservatórios estudados.

Entretanto, algumas associações relevantes foram identificadas. A correlação mais elevada foi observada entre nitrito e nitrato ($r = 0,87$), indicando uma forte associação entre esses compostos nitrogenados. Essa relação pode estar associada aos processos de nitrificação no ambiente aquático, nos quais a amônia é oxidada a nitrito e, posteriormente, a nitrato, refletindo a dinâmica do ciclo do nitrogênio nos reservatórios analisados.

Outra correlação expressiva foi observada entre fósforo total e clorofila total ($r = 0,68$), sugerindo a influência do fósforo sobre a biomassa fitoplanctônica. Essa relação é frequentemente associada a processos de enriquecimento nutricional, uma vez que o fósforo atua como nutriente limitante ao crescimento do fitoplâncton. Correlação moderada também foi observada entre turbidez e amônia ($r = 0,52$), possivelmente relacionada à ressuspensão de sedimentos ou à presença de matéria orgânica em suspensão. Além disso, a turbidez apresentou correlação moderada com clorofila total ($r = 0,43$), indicando possível contribuição da biomassa fitoplanctônica para o aumento da turbidez da água.

No que se refere às correlações negativas, destacam-se a relação entre pH e potencial redox ($r = -0,45$), indicando uma possível interação entre as condições de acidez/alcalinidade e os processos redox no ambiente aquático. Também foi observada correlação negativa entre temperatura da água e potencial redox ($r = -0,24$), sugerindo que variações térmicas podem influenciar os processos de oxidação e redução. Adicionalmente, a relação negativa entre fósforo total e potencial redox ($r = -0,35$) pode indicar condições redutoras associadas à liberação de fósforo a partir dos sedimentos, especialmente em ambientes com baixa oxigenação.

De modo geral, os resultados evidenciam que as relações entre as variáveis limnológicas são complexas e não predominantemente lineares, sendo influenciadas por múltiplos fatores físicos, químicos e biológicos. Esse padrão, caracterizado pela predominância de correlações fracas a moderadas, reforça a limitação das análises univariadas ou bivariadas na compreensão da dinâmica do sistema. Nesse contexto, esses resultados justificam a aplicação de técnicas multivariadas na etapa subsequente, como a Análise de Componentes Principais (PCA), que sintetiza a variabilidade do conjunto de dados e identifica os principais gradientes ambientais

54

que estruturam o sistema estudado.

4.3 ANÁLISE DE COMPONENTES PRINCIPAIS (PCA)

A Análise de Componentes Principais (PCA) permitiu identificar os principais padrões de variabilidade entre as variáveis limnológicas analisadas. Os resultados da decomposição da matriz de correlação são apresentados na **Tabela 2**, que mostra os autovalores, a variância explicada e a variância acumulada de cada componente.

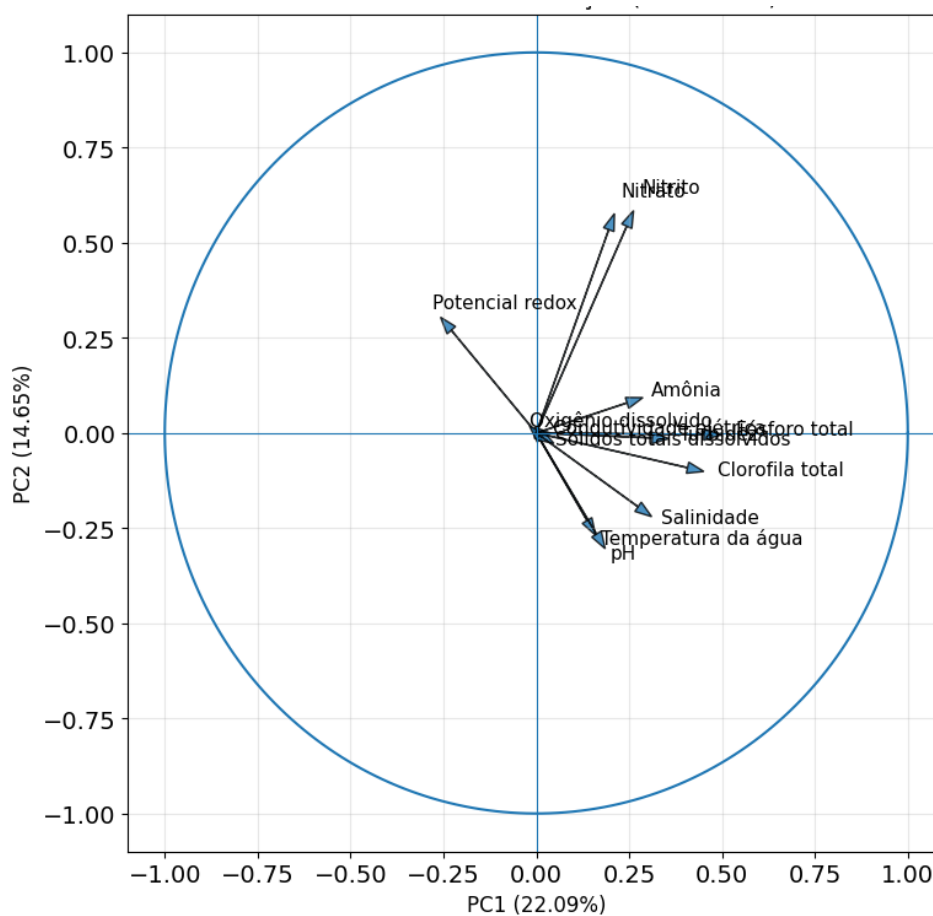
Tabela 2 – Variância explicada pelos componentes principais

Componente	Autovalor	Variância explicada (%)	Variância acumulada (%)
PC1	2.88	22.09	22.09
PC2	1.91	14.65	36.75
PC3	1.46	11.18	47.93
PC4	1.12	8.61	56.54
PC5	1.02	7.84	64.38
PC6	0,99	7.68	72.06
PC7	0.97	7.45	79.51
PC8	0.80	6.14	85.64
PC9	0.60	4.59	90.23
PC10	0.56	4.32	94.55
PC11	0.38	2.9	97.45
PC12	0.22	1.73	99.18
PC13	0.11	0.82	100

Fonte: Autora (2026)

De acordo com o critério de Kaiser, os componentes com autovalores superiores a 1 foram considerados relevantes, resultando na retenção de cinco componentes principais. O primeiro componente (PC1) explicou 22,09% da variância total dos dados, enquanto o segundo (PC2) explicou 14,65%. Os cinco primeiros componentes principais explicaram, conjuntamente, 64,37% da variabilidade total do conjunto de dados, indicando que grande parte da estrutura das relações entre as variáveis pode ser representada por esse subconjunto de componentes.

A projeção das variáveis no espaço definido pelos dois primeiros componentes principais é apresentada na **Figura 10**.

Figura 10 – Círculo de correlação (PC1 × PC2)

Fonte: Autora (2026)

O PC1 apresentou maior associação com fósforo total, clorofila total, amônia e salinidade, variáveis orientadas em direção semelhante no plano PC1 × PC2. Essa configuração sugere um gradiente relacionado ao enriquecimento nutricional e à produtividade biológica do sistema, evidenciando a influência da disponibilidade de nutrientes no desenvolvimento da biomassa fitoplanctônica nos reservatórios analisados.

O PC2 foi principalmente influenciado pelas variáveis nitrito e nitrato, que apresentam forte associação entre si e orientação semelhante no plano da PCA. Esse padrão reflete a dinâmica do ciclo do nitrogênio no ambiente aquático. De modo geral, a PCA evidenciou que a qualidade da água nos reservatórios estudados é influenciada por múltiplos processos ambientais, destacando-se gradientes associados ao enriquecimento trófico e às transformações biogeoquímicas do nitrogênio. Embora a técnica permita reduzir a dimensionalidade dos dados e identificar relações entre as variáveis, sua representação linear pode limitar a identificação de estruturas mais complexas presentes no conjunto de dados.

4.4 MODELAGEM COM SOM – ABORDAGEM INTEGRADA

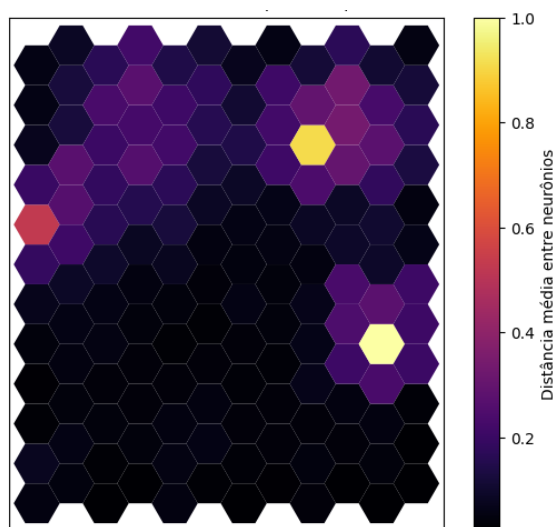
4.4.1 Treinamento e desempenho do modelo

O SOM foi treinado com base em 854 amostras e 13 variáveis limnológicas, previamente padronizadas por z-score. A rede foi estruturada em uma grade bidimensional hexagonal de 12×12 neurônios (144 unidades), conforme definido na etapa metodológica. O desempenho do modelo foi avaliado por meio do Erro de Quantização (Quantization Error – QE), que representa a distância média entre cada vetor de entrada e o vetor de pesos do neurônio vencedor (BMU). O modelo apresentou **QE = 0,7642**, indicando que o mapa tem capacidade adequada para representar a estrutura do conjunto de dados utilizado na modelagem.

4.4.2 Estrutura do mapa SOM (U-Matrix)

A estrutura do mapa auto-organizável foi analisada por meio da matriz unificada de distâncias (U-Matrix), apresentada na **Figura 11**. No mapa, cada hexágono representa um neurônio do SOM, ao qual estão associadas amostras com características semelhantes. A escala de cores indica a distância média entre neurônios adjacentes, em que tons mais escuros representam distâncias menores, enquanto os mais claros indicam distâncias maiores.

Figura 11 – U-Matrix do SOM (12×12)



Fonte: Autora (2026)

Observa-se que grande parte do mapa apresenta tonalidades mais escuras, especialmente nas regiões central e inferior, indicando baixa distância entre os neurônios e, conseqüentemente,

57

elevada similaridade entre as amostras agrupadas nessas áreas. Esse padrão sugere a presença de um conjunto significativo de amostras com características semelhantes, refletindo uma relativa homogeneidade nas condições de qualidade da água em parte dos dados analisados.

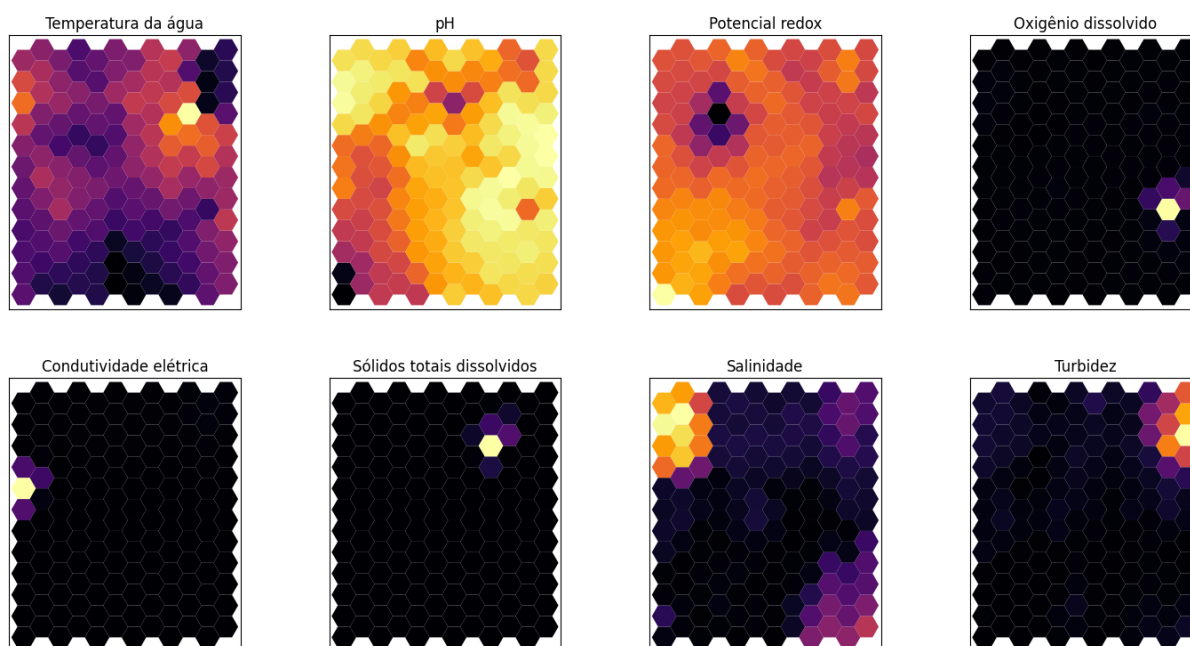
Por outro lado, algumas regiões do mapa apresentam tonalidades mais claras, especialmente em pontos isolados distribuídos ao longo do mapa, indicando maiores distâncias entre neurônios vizinhos. Essas áreas representam possíveis zonas de transição entre grupos de dados com características distintas, atuando como limites naturais entre agrupamentos.

4.4.3 Mapas de componentes

Esses mapas permitem visualizar a distribuição espacial das variáveis analisadas ao longo do SOM, possibilitando não apenas a identificação de padrões de comportamento, mas também a interpretação dos processos ambientais que estruturam a qualidade da água nos reservatórios. Para facilitar a visualização, os mapas de componentes foram divididos em 2 figuras: a primeira, com oito variáveis físico-químicas, e a segunda, com as variáveis químicas e biológicas, correspondentes às **Figuras 12 e 13**, respectivamente. Cada mapa de componente corresponde a uma variável específica e é representado pela mesma estrutura de neurônios do SOM. Dessa forma, posições equivalentes nos diferentes mapas correspondem aos mesmos neurônios, possibilitando a análise comparativa entre as variáveis e a identificação de possíveis relações entre os parâmetros analisados. A escala de cores representa a magnitude relativa de cada variável nos neurônios do SOM. Tons mais escuros indicam valores menores, enquanto tons mais claros indicam valores maiores da variável analisada, permitindo visualizar sua distribuição espacial no mapa.

Para compreender a contribuição das variáveis físico-químicas na organização do mapa auto-organizável, foram analisados os mapas de componentes apresentados na **Figura 12**. Esses mapas representam a distribuição espacial de cada variável sobre o SOM, permitindo visualizar como os diferentes parâmetros influenciam a formação dos padrões observados no mapa.

58

Figura 12 – Mapas de componentes: variáveis físico-químicas (abordagem integrada)

Fonte: Autora (2026)

A distribuição da temperatura da água indica a presença de gradientes térmicos associados a diferentes condições ambientais, possivelmente relacionados à variação da profundidade, à incidência solar e à dinâmica de mistura da coluna d'água. Essas diferenças podem influenciar diretamente os processos metabólicos e a solubilidade de gases, afetando a qualidade da água.

O pH apresenta predominância de valores mais elevados, indicando condições levemente alcalinas, típicas de reservatórios em regiões semiáridas, onde a evaporação intensa e a concentração de sais contribuem para o aumento da alcalinidade. Esse comportamento pode influenciar a disponibilidade de nutrientes e a dinâmica biogeoquímica do sistema.

O potencial redox evidencia variações espaciais associadas às condições de oxidação e redução, refletindo a intensidade dos processos biogeoquímicos, como a decomposição da matéria orgânica. Regiões com menores valores de potencial redox podem indicar ambientes mais redutores, potencialmente associados à liberação de nutrientes, como fósforo, a partir dos sedimentos.

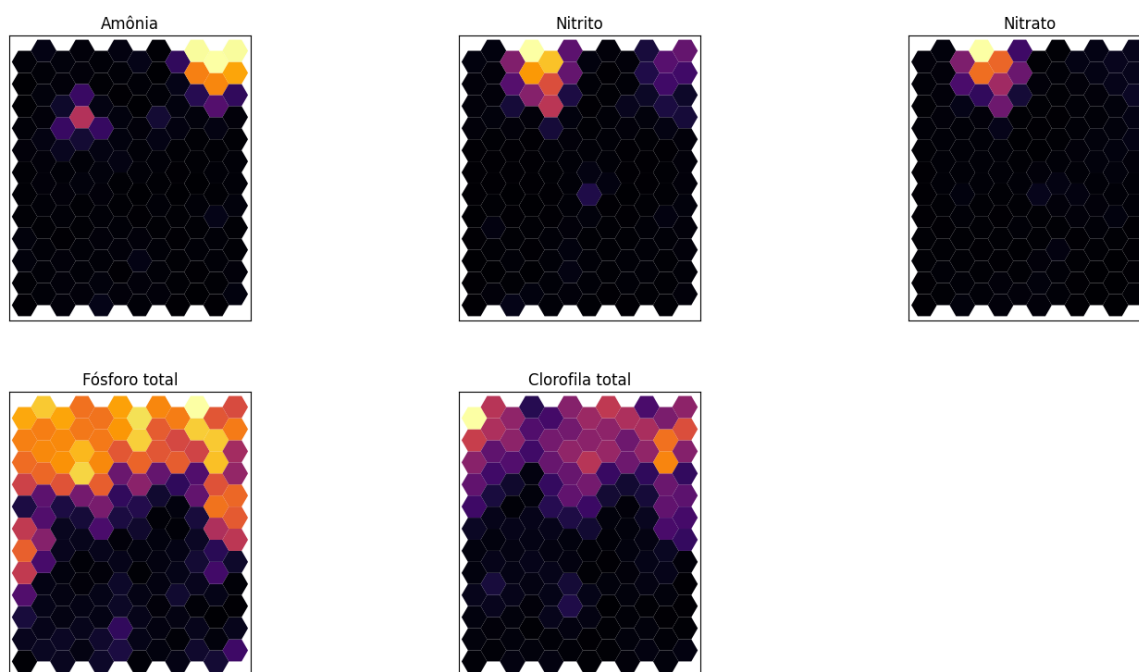
A distribuição do oxigênio dissolvido, predominantemente baixa, sugere condições de oxigenação limitada em grande parte das amostras, possivelmente relacionadas à elevada carga orgânica ou à estratificação da coluna d'água. As regiões com maiores concentrações indicam áreas com maior renovação hídrica ou atividade fotossintética, evidenciando heterogeneidade nos processos ecológicos.

59

A similaridade espacial entre a condutividade elétrica, os sólidos totais dissolvidos e a salinidade indica que esses parâmetros são controlados por um mesmo fator dominante, associado à concentração de íons dissolvidos. Esse padrão sugere a influência de processos de evaporação e de concentração salina, característicos de ambientes semiáridos, além de possíveis aportes externos de sais.

A turbidez apresenta regiões com valores elevados, indicando maior presença de partículas em suspensão, podendo estar associadas à ressuspensão de sedimentos, ao aporte de material particulado ou à intensificação da biomassa fitoplanctônica. Esse padrão evidencia a influência de processos físicos e biológicos na estrutura da qualidade da água. Além das variáveis físico-químicas, foram analisados os mapas de componentes referentes a nutrientes e indicadores de biomassa fitoplanctônica, apresentados na **Figura 13**. Esses mapas permitem identificar a distribuição espacial das concentrações de compostos nitrogenados, de fósforo total e de clorofila no SOM.

Figura 13 – Mapas de componentes: variáveis químicas e biológicas (abordagem integrada)



Fonte: Autora (2026)

De modo geral, observa-se que amônia ($\text{NH}_3/\text{NH}_4^+$), nitrito (NO_2^-) e nitrato (NO_3^-) apresentam valores baixos na maior parte do mapa, com regiões pontuais de maior intensidade concentradas principalmente nas áreas superiores do SOM. Esse padrão sugere que, em grande parte das amostras, as formas dissolvidas de nitrogênio ocorreram em níveis reduzidos,

60

enquanto determinados conjuntos específicos de amostras apresentaram concentrações mais elevadas, caracterizando condições diferenciadas quanto ao aporte e/ou à transformação do nitrogênio.

A amônia apresenta uma região bem definida com maiores concentrações na porção superior direita do mapa, indicando um agrupamento de amostras com maiores concentrações dessa forma reduzida de nitrogênio. Do ponto de vista ambiental, esse comportamento pode estar relacionado à maior influência da matéria orgânica e de processos de mineralização, além de um possível aporte de efluentes. Isso ocorre porque a amônia tende a aumentar em ambientes com elevada carga orgânica e baixa eficiência na oxidação do nitrogênio.

O nitrito apresenta valores elevados, principalmente na região superior central, enquanto o nitrato apresenta intensificação em área semelhante, predominantemente na região superior. A presença de áreas com maiores concentrações de nitrito e nitrato indica variações na predominância das formas oxidadas do nitrogênio, podendo refletir diferenças nos processos de nitrificação e nas condições, bem como nas condições de oxigenação e de mistura da coluna d'água. Em sistemas aquáticos, o nitrito geralmente atua como intermediário no ciclo do nitrogênio. Portanto, concentrações elevadas desse composto podem indicar condições transitórias associadas à transformação do nitrogênio, enquanto o nitrato tende a representar a forma mais estável em ambientes oxigenados.

O fósforo total apresenta um padrão distinto dos compostos nitrogenados, com valores mais elevados em uma faixa relativamente ampla na porção superior do mapa e mais baixos em grande parte das porções centrais e inferiores. O gradiente espacial indica a presença de um conjunto de amostras com maior enriquecimento em fósforo, variável frequentemente associada ao aumento do potencial de eutrofização em ambientes lênticos. A distribuição observada sugere que o fósforo total desempenhou papel relevante na diferenciação dos grupos formados pelo SOM, contribuindo para a separação de regiões do mapa com condições tróficas contrastantes.

A clorofila total, utilizada como indicador da biomassa fitoplanctônica, apresenta valores mais elevados majoritariamente nas regiões superiores do SOM, com intensificações localizadas, o que indica que determinados agrupamentos estão associados a maior biomassa algal. A semelhança parcial entre os padrões de clorofila e de fósforo total reforça a relação esperada entre a disponibilidade de nutrientes e o crescimento fitoplanctônico. Assim, regiões do SOM com maiores valores de fósforo total tendem a apresentar, de forma concomitante, maior clorofila, sugerindo condições mais favoráveis à produção primária e potencialmente

61

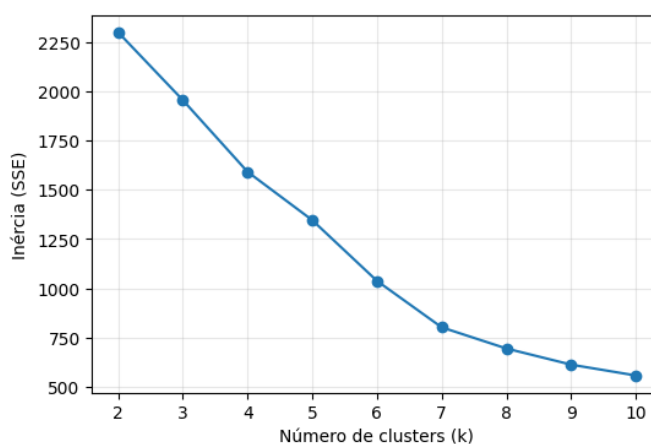
associadas a estados tróficos mais elevados.

De forma integrada, os mapas de componentes associados a nutrientes e à biomassa evidenciam que o SOM foi capaz de identificar regiões do mapa com maior enriquecimento de nutrientes e maior biomassa fitoplanctônica, em contraste com áreas onde os valores dessas variáveis permanecem mais baixos e homogêneos.

4.4.4 Clusterização do mapa (k-means)

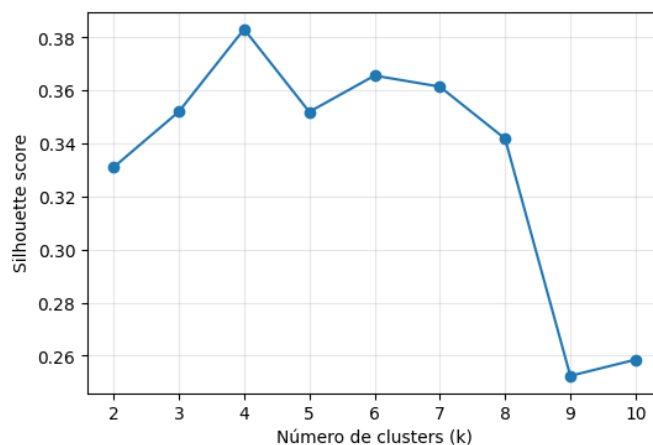
A determinação do número ideal de clusters foi realizada com base na análise conjunta do Método do Cotovelo (Elbow Method) e do Índice de Silhueta (Silhouette Score). O gráfico do método do cotovelo, apresentado na **Figura 14**, mostra a variação da inércia em função do número de clusters testados ($k = 2$ a 10). Observa-se uma redução acentuada da inércia até aproximadamente $k = 4$, a partir do qual os ganhos adicionais tornam-se progressivamente menores, indicando a estabilização da variabilidade intra-grupo.

Figura 14 – Método do Cotovelo (Elbow Method)



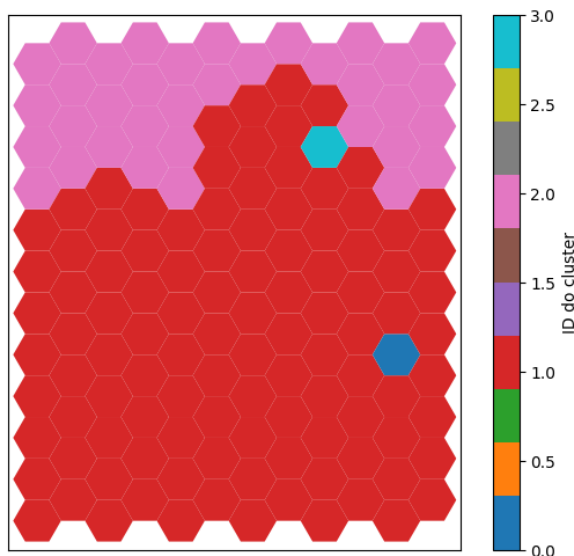
Fonte: Autora (2026)

Complementarmente, o índice de silhueta foi calculado para avaliar a qualidade da separação entre os agrupamentos. A **Figura 15** apresenta a variação do índice de silhueta para diferentes números de clusters ($k = 2$ a 10). O maior valor foi observado para $k = 4$ (0,383), indicando maior coesão interna dos clusters e melhor separação entre os grupos identificados.

Figura 15 – Índice de Silhueta

Fonte: Autora (2026)

A análise conjunta desses critérios, indicou $k = 4$ como número ideal de agrupamentos para a segmentação do mapa auto-organizável. A partir desse valor, foi aplicado o algoritmo K-means, resultando na divisão dos neurônios do mapa em quatro clusters distintos no mapa auto-organizável. Cada cluster é representado por uma cor diferente na **Figura 16**.

Figura 16 – Distribuição dos clusters no mapa SOM

Fonte: Autora (2026)

Observa-se que o Cluster 2 (vermelho) ocupa a maior parte da área do mapa, indicando que uma parcela significativa das amostras apresenta padrões relativamente semelhantes de qualidade da água. Esse agrupamento representa, portanto, a condição predominante observada no conjunto de dados analisado. O Cluster 3 (rosa) encontra-se concentrado principalmente na porção superior do mapa, formando uma região contínua que se distingue do agrupamento predominante. Esse padrão sugere a presença de amostras com características ambientais

63

distintas, possivelmente associadas a variações na concentração de nutrientes ou de determinadas variáveis físico-químicas. Já os Clusters 4 (verde) e 1 (azul) aparecem de forma mais localizada no mapa, sendo representados por um número reduzido de neurônios. Esses agrupamentos podem refletir condições ambientais específicas ou atípicas no conjunto de dados analisado.

4.4.5 Interpretação ambiental dos agrupamentos

A identificação dos agrupamentos obtidos por meio da clusterização K-means aplicada ao mapa auto-organizável permite compreender com mais clareza os padrões ambientais presentes no conjunto de dados analisado. A interpretação desses agrupamentos foi realizada de forma integrada aos mapas de componentes apresentados na seção anterior, que permitem identificar as variáveis que mais contribuem para a diferenciação entre os clusters.

Cada cluster representa um conjunto de neurônios do SOM com características semelhantes quanto às variáveis físico-químicas e biológicas analisadas. Dessa forma, os agrupamentos identificados refletem diferentes condições ambientais presentes nos reservatórios estudados.

Observa-se que um dos clusters ocupa a maior parte da área do mapa auto-organizável, indicando que grande parte das amostras apresenta características relativamente semelhantes em relação aos parâmetros analisados. Por outro lado, alguns clusters aparecem de forma mais localizada no mapa, representando regiões específicas do SOM que apresentam padrões de qualidade da água distintos. A análise dos mapas de componentes sugere que essas diferenças estão associadas principalmente a variações na concentração de nutrientes, na biomassa fitoplanctônica e no grau de mineralização da água.

Entre as variáveis que mais contribuem para a diferenciação entre os agrupamentos, destacam-se aquelas relacionadas à mineralização da água, como condutividade elétrica, salinidade e sólidos totais dissolvidos, bem como as variáveis associadas à disponibilidade de nutrientes e à produtividade biológica, como fósforo total, compostos nitrogenados e clorofila.

De modo geral, observa-se que regiões do mapa associadas a maiores concentrações de fósforo total tendem também a apresentar maiores valores de clorofila, indicando possível relação entre a disponibilidade de nutrientes e o aumento da biomassa fitoplanctônica. De forma semelhante, padrões espaciais observados entre a condutividade elétrica, a salinidade e os sólidos totais dissolvidos indicam associação entre essas variáveis e refletem processos relacionados à concentração de sais dissolvidos na água.

A caracterização quantitativa dos agrupamentos pode ser melhor compreendida a partir da análise das médias das variáveis em cada cluster, apresentada na **Tabela 3**.

Tabela 3 – Média das variáveis analisadas em cada cluster identificado no SOM

	Cluster 1	Cluster 2	Cluster 3	Cluster 4
Temperatura da água (°C)	26.51	26.89	28.36	29.14
pH	7.10	8.03	8.40	9.14
Potencial redox	386.00	185.52	137.82	121.00
Condutividade elétrica (mS cm ⁻¹)	0.50	1.15	1.50	1.46
Turbidez (NTU)	2.60	24.93	55.59	61.40
Oxigênio dissolvido (mg/L)	6.84	8.98	6.65	4.40
Sólidos totais dissolvidos (g L ⁻¹)	0.32	0.36	2.20	927.00
Salinidade (ppt)	0.02	0.08	0.28	0.07
Amônia (µg L ⁻¹)	37.5	42.97	37.80	38.41
Nitrito (µg L ⁻¹)	0.00	3.21	10.77	9.15
Nitrato (µg L ⁻¹)	4.00	35.42	88.55	18.29
Fósforo total (µg L ⁻¹)	15.33	58.25	254.22	297.00
Clorofila total (µg L ⁻¹)	11.1	27.84	99.94	110.00

Fonte: Autora (2026)

O Cluster 1 apresenta os menores valores de turbidez, fósforo total, nitrito, nitrato e clorofila entre os agrupamentos. Esses resultados indicam condições de menor disponibilidade de nutrientes e menor biomassa fitoplanctônica, sugerindo ambientes aquáticos com menor influência de processos de eutrofização. Além disso, a baixa turbidez observada nesse cluster indica menor concentração de partículas em suspensão, associada a maior transparência da água e menor aporte de sedimentos ou matéria orgânica. Dessa forma, o Cluster 1 pode ser interpretado como representando condições ambientais de menor enriquecimento trófico.

O Cluster 2 apresenta aumento moderado nas concentrações de fósforo total, nitrato e clorofila em relação ao Cluster 1, indicando maior disponibilidade de nutrientes e maior biomassa fitoplanctônica. Além disso, observa-se aumento da turbidez e da condutividade elétrica, o que pode indicar maior presença de partículas em suspensão e maior concentração de substâncias dissolvidas na água. Essas características indicam que o Cluster 2 representa uma condição intermediária de qualidade da água, caracterizada por maior disponibilidade de nutrientes em relação ao Cluster 1, porém ainda inferior aos valores observados nos clusters mais enriquecidos.

O Cluster 3 apresenta valores elevados de fósforo total, nitrato, nitrito e clorofila, além de turbidez significativamente maior em comparação com os clusters anteriores. Esses resultados indicam ambientes aquáticos com maior disponibilidade de nutrientes e maior

65

biomassa fitoplanctônica, sugerindo condições de maior produtividade biológica. A elevada turbidez observada nesse cluster pode estar associada tanto ao aumento da biomassa algal quanto à maior presença de partículas em suspensão no ambiente aquático. Dessa forma, o Cluster 3 pode ser interpretado como representando condições de maior enriquecimento trófico, possivelmente associadas a maior aporte de nutrientes nos reservatórios analisados.

O Cluster 4 apresenta os maiores valores de turbidez, fósforo total e clorofila entre todos os agrupamentos, indicando ambientes com elevada disponibilidade de nutrientes e intensa produção fitoplanctônica. Além disso, observa-se redução significativa dos valores de oxigênio dissolvido em comparação aos demais clusters, o que pode refletir maior consumo de oxigênio associado à decomposição de matéria orgânica ou a intensa atividade biológica. Essas características sugerem que o Cluster 4 representa condições de maior enriquecimento trófico e produtividade biológica, podendo estar associado a ambientes com maior influência de processos de eutrofização.

De forma geral, a análise integrada dos resultados do mapa auto-organizável e da clusterização K-means evidencia a presença de um gradiente ambiental associado ao enriquecimento em nutrientes e ao aumento da biomassa fitoplanctônica nos reservatórios analisados. Observa-se uma transição entre condições de menor enriquecimento trófico, representadas pelo Cluster 1, e ambientes com maiores concentrações de nutrientes, maior turbidez e maior biomassa algal, representados principalmente pelos Clusters 3 e 4. Esses resultados demonstram a capacidade do SOM de identificar padrões complexos em dados multivariados de qualidade da água, permitindo distinguir diferentes condições ambientais nos reservatórios estudados e contribuindo para a compreensão da variabilidade espacial das características limnológicas desses sistemas.

4.5 MODELAGEM COM SOM – APLICAÇÃO INDIVIDUAL POR RESERVATÓRIO

4.5.1 Treinamento e desempenho dos modelos individuais

Com o objetivo de investigar os padrões da qualidade da água de cada sistema hídrico, foram ajustados modelos de Mapas Auto-Organizáveis individualmente para os reservatórios Acauã, Camalaú, Cordeiro, Epitácio Pessoa, Poções, Sumé e Taperoá. Em todos os casos, foram utilizadas as 13 variáveis limnológicas previamente selecionadas, padronizadas por meio da transformação z-score, conforme descrito na metodologia, a fim de eliminar diferenças de

66

escala entre as variáveis e garantir a comparabilidade entre os modelos.

Considerando que cada reservatório possui aproximadamente 140 a 160 amostras, foi estimado um número de neurônios do SOM próximo de 60 para a estrutura do mapa. Com isso, foi adotada uma grade bidimensional de 8×8 (64 unidades) em todos os SOM individuais. Essa configuração apresenta um número de neurônios próximo ao recomendado pela heurística e, ao mesmo tempo, permite uma melhor visualização dos padrões espaciais no mapa. Além disso, a utilização da mesma estrutura de mapa em todos os reservatórios garante consistência metodológica e comparabilidade direta entre os modelos individuais e o modelo integrado.

O desempenho dos modelos foi avaliado por meio do Erro de Quantização (QE). Os valores obtidos para cada reservatório são apresentados na **Tabela 4**.

Tabela 4 – Erro de Quantização (QE) dos modelos SOM individuais

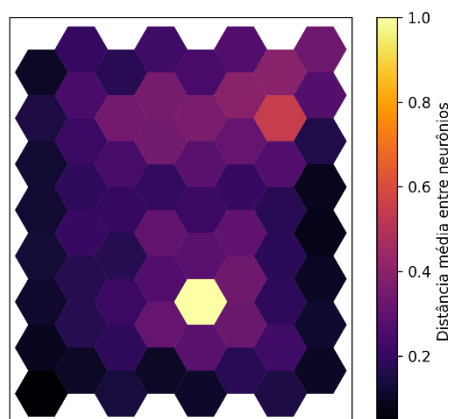
Reservatório	QE
Acauã	0,996
Camalaú	1,078
Cordeiro	0,974
Epitácio Pessoa	1,191
Poções	1,040
Sumé	1,001
Taperoá	0,726

Fonte: Autora (2026)

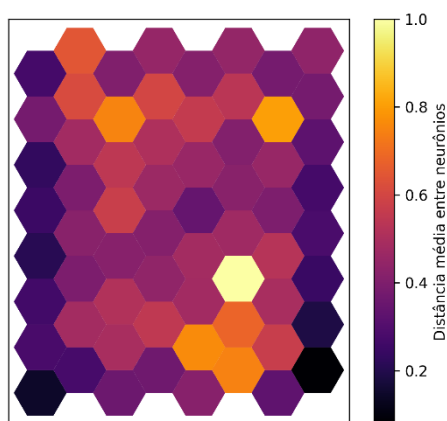
De modo geral, os valores observados indicam um desempenho satisfatório dos modelos SOM individuais, uma vez que os erros de quantização permaneceram próximos de 1 ou inferiores a 1 na maioria dos casos. O menor valor foi observado no reservatório Taperoá, sugerindo maior homogeneidade relativa na estrutura dos dados desse sistema. Em contraste, o maior valor foi registrado para Epitácio Pessoa, o que indica maior complexidade ou heterogeneidade ambiental interna. Esses resultados sugerem que, embora os reservatórios pertençam à mesma bacia hidrográfica, cada sistema apresenta uma estrutura limnológica distinta, refletindo diferenças nas condições ambientais e nos processos que influenciam a qualidade da água.

4.5.2 Estrutura dos mapas SOM individuais

A estrutura dos mapas SOM individuais foi analisada por meio da U-Matrix. As **Figuras 17 e 18** apresentam as U-Matrix dos reservatórios Acauã e Camalaú, respectivamente.

Figura 17 – U-Matrix Acauã

Fonte: Autora (2026)

Figura 18 – U-Matrix Camalaú

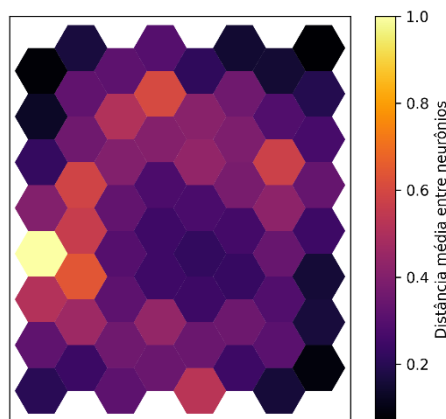
Fonte: Autora (2026)

Nesses reservatórios, observam-se regiões com diferentes níveis de distância entre neurônios, indicando a presença de padrões limnológicos distintos em cada sistema. No reservatório Acauã, predominam cores mais escuras, distribuídas ao longo da maior parte do mapa, com poucas áreas isoladas, com maior distância entre neurônios. Esse padrão sugere que as amostras se organizam em regiões relativamente contínuas no espaço do mapa, indicando transições graduais entre diferentes condições limnológicas. Em contraste, o reservatório Camalaú apresenta maior presença de células com elevada distância entre neurônios, distribuídas em diferentes regiões do mapa. Esse comportamento indica maior fragmentação da estrutura do SOM, sugerindo que as amostras se agrupam em subconjuntos mais bem definidos, separados por zonas de transição mais abruptas. Assim, os dados de Camalaú revelam maior separação entre padrões limnológicos, enquanto, em Acauã, as amostras se distribuem de forma mais gradual ao longo do espaço de características ambientais.

68

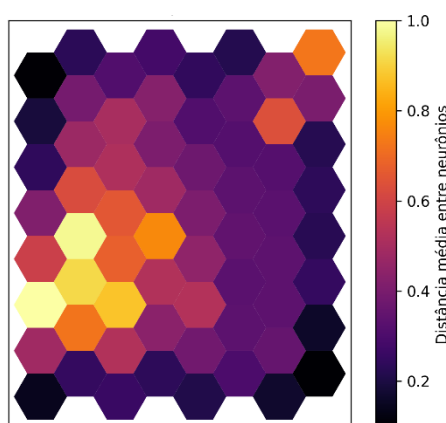
As **Figuras 19 e 20** apresentam os mapas U-Matrix para os reservatórios Cordeiro e Epitácio Pessoa.

Figura 19 – U-Matrix Cordeiro



Fonte: Autora (2026)

Figura 20 – U-Matrix Epitácio Pessoa



Fonte: Autora (2026)

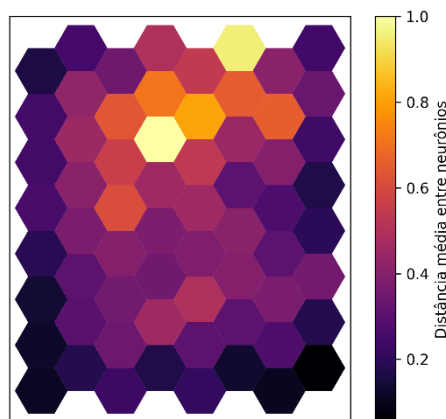
No reservatório Cordeiro, a estrutura do mapa apresenta algumas regiões isoladas com maiores distâncias entre neurônios, mas a maior parte da grade é composta por valores moderados. Esse padrão indica a presença de subconjuntos de amostras diferenciados, porém conectados por transições relativamente suaves no espaço do SOM. Por outro lado, o reservatório Epitácio Pessoa apresenta regiões mais amplas, com distâncias elevadas entre neurônios, formando áreas de contraste mais evidentes no mapa. Esse comportamento sugere que esse reservatório apresenta maior heterogeneidade limnológica, possivelmente relacionada à maior complexidade hidrológica e às variações espaciais e temporais nas condições ambientais.

As **Figuras 21 e 22** mostram as U-Matrices dos reservatórios Sumé e Poções, nas quais

69

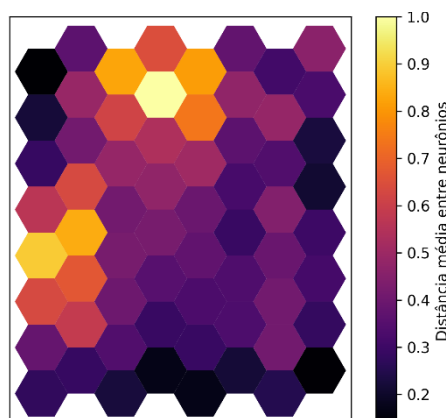
se observa uma distribuição relativamente equilibrada das distâncias entre neurônios. Ressalta-se que esses reservatórios correspondem aos menores sistemas do conjunto analisado.

Figura 21 – U-Matrix Sumé



Fonte: Autora (2026)

Figura 22 – U-Matrix Poções



Fonte: Autora (2026)

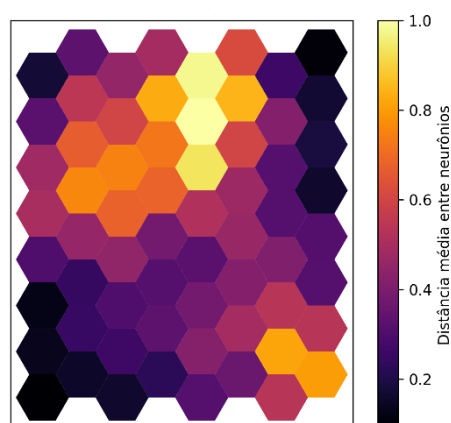
No reservatório Sumé, observa-se uma região central do mapa com valores elevados de distância entre neurônios, representada pelas cores mais claras. Essa área de maior contraste é cercada por zonas com valores intermediários, indicando a presença de uma região de transição bem definida no espaço do SOM. Essa configuração sugere que uma parcela significativa das amostras de Sumé apresenta condições limnológicas relativamente próximas entre si, enquanto um subconjunto menor se destaca por apresentar características ambientais diferenciadas. O reservatório Poções, por sua vez, apresenta um padrão espacial semelhante, porém com distribuição mais alongada das regiões de maior distância entre neurônios. Observa-se uma faixa vertical no lado esquerdo do mapa, caracterizada por valores elevados de distância,

70

indicando a presença de um gradiente mais evidente entre os conjuntos de amostras ao longo dessa região. Além disso, a área central do mapa apresenta valores elevados, sugerindo a presença de mais de um conjunto de amostras com características diferenciadas, o que contribui para a formação de zonas de contraste mais extensas. Comparando os dois reservatórios, observa-se que Sumé apresenta uma região de diferenciação mais concentrada no centro do mapa, enquanto, em Poções, as zonas de contraste aparecem distribuídas ao longo de uma faixa mais extensa da grade do SOM. Esse padrão sugere que, em Poções, as variações nas características limnológicas podem ocorrer de forma mais gradual ao longo do conjunto de amostras, enquanto, em Sumé, a diferenciação parece ocorrer de forma mais localizada. Essas diferenças na organização do mapa indicam que, embora ambos os reservatórios apresentem conjuntos de amostras com características distintas, a forma como essas variações se distribuem no espaço do SOM difere entre os sistemas.

Por fim, a **Figura 23** apresenta o mapa do reservatório Taperoá, no qual se observa um contraste mais acentuado entre algumas regiões, indicando a presença de conjuntos de amostras com características limnológicas bastante distintas.

Figura 23 – U-Matrix Taperoá



Fonte: Autora (2026)

Nesse reservatório, observa-se a presença de regiões com distâncias elevadas entre neurônios distribuídas em diferentes áreas do mapa, formando zonas de contraste relativamente bem definidas. Esse padrão sugere que as condições limnológicas em Taperoá apresentam maior variabilidade interna, possivelmente refletindo diferenças espaciais entre áreas do reservatório ou variações temporais nas características da água.

De forma geral, a análise da U-Matrix evidencia que os diferentes reservatórios apresentam estruturas distintas de organização no espaço do SOM, refletindo variações na

71

distribuição das amostras de acordo com suas características limnológicas. Enquanto alguns sistemas apresentam transições mais graduais entre regiões do mapa, outros exibem zonas mais delimitadas, sugerindo maior separação entre padrões ambientais. Esses resultados indicam que cada sistema apresenta uma configuração própria de variabilidade ambiental, refletindo diferenças nas condições hidrológicas, nos aportes de nutrientes e nos processos físicos e biogeoquímicos que controlam a qualidade da água.

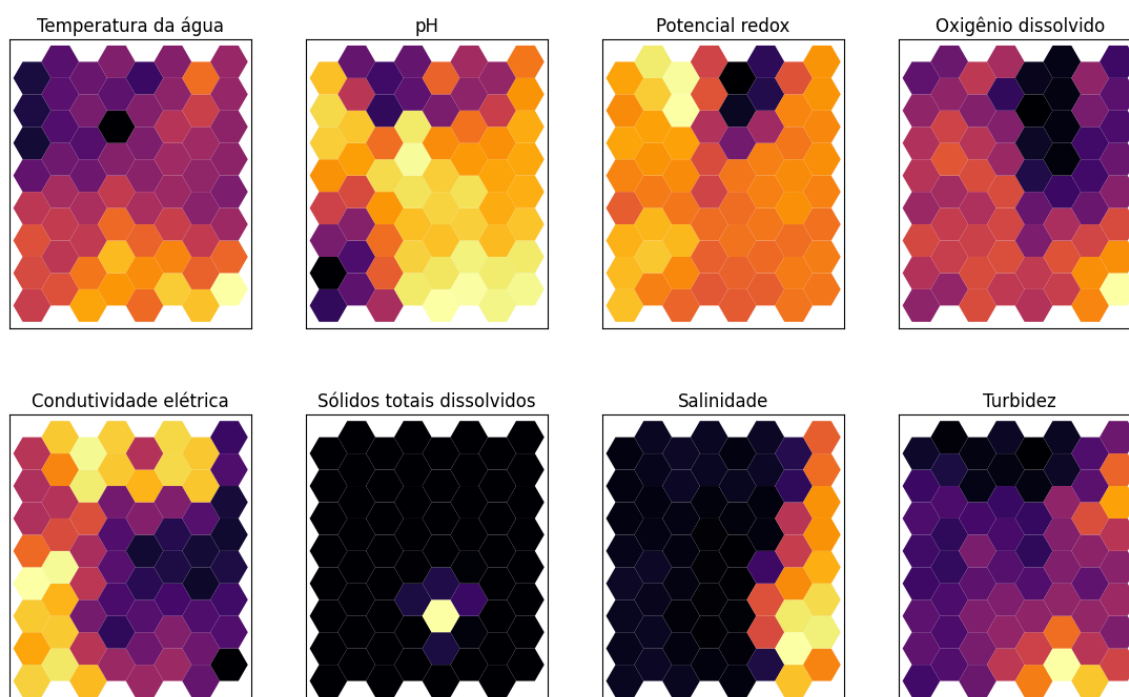
4.5.3 Mapas de componentes individuais

A interpretação dos mapas de componentes de cada reservatório segue o mesmo princípio aplicado aos mapas de componentes da abordagem integrada: cores mais claras representam valores mais elevados das variáveis, enquanto cores mais escuras indicam concentrações mais baixas. A análise conjunta desses mapas permite identificar gradientes ambientais e possíveis associações entre os parâmetros analisados.

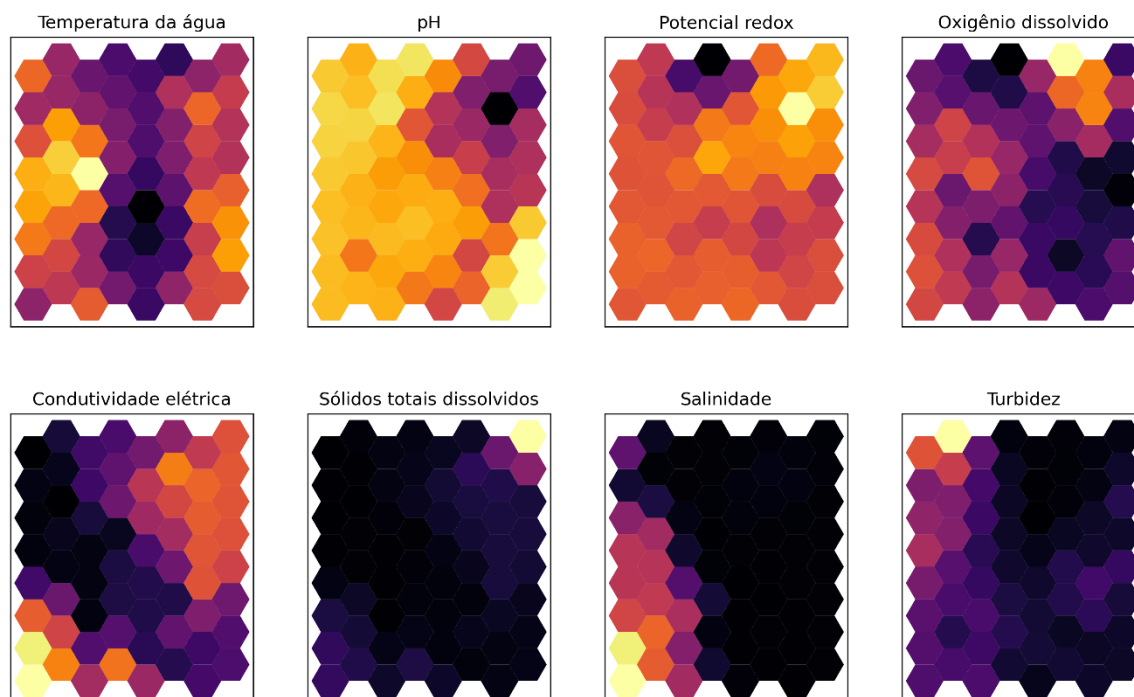
4.5.3.1 Acauã e Camalaú

A distribuição espacial das variáveis físico-químicas nos mapas auto-organizáveis dos reservatórios de Acauã e Camalaú é apresentada nas **Figuras 24 e 25**.

Figura 24 – Mapas de Componentes: variáveis físico-químicas de Acauã



Fonte: Autora (2026)

Figura 25 – Mapas de Componentes: variáveis físico-químicas de Camalaú

Fonte: Autora (2026)

No reservatório de Acauã (**Figura 24**), observam-se gradientes espaciais que indicam heterogeneidade nas condições limnológicas. A variação da temperatura da água sugere a ocorrência de diferenças na dinâmica térmica do reservatório, possivelmente associadas à estratificação da coluna d'água ou à influência de fatores sazonais. O predomínio de valores elevados de potencial redox indica condições majoritariamente oxidantes, sugerindo boa disponibilidade de oxigênio e menor tendência à liberação interna de nutrientes a partir dos sedimentos.

A distribuição da condutividade elétrica e da salinidade evidencia a influência de processos de concentração de sais, característicos de ambientes semiáridos, possivelmente associados à evaporação e à acumulação de íons dissolvidos. No entanto, os gradientes distintos entre essas variáveis indicam que a dinâmica iônica não é homogênea, sugerindo a atuação de diferentes fontes de aporte ou processos de diluição e concentração ao longo do reservatório.

A turbidez apresenta regiões com valores mais elevados, indicando maior presença de material particulado em suspensão, possivelmente associada à ressuspensão de sedimentos ou a eventos hidrológicos. Esse comportamento sugere maior instabilidade em determinadas áreas do reservatório, podendo impactar a transparência da água e os processos ecológicos associados.

No reservatório de Camalaú (**Figura 25**), observa-se um padrão mais contrastante entre

73

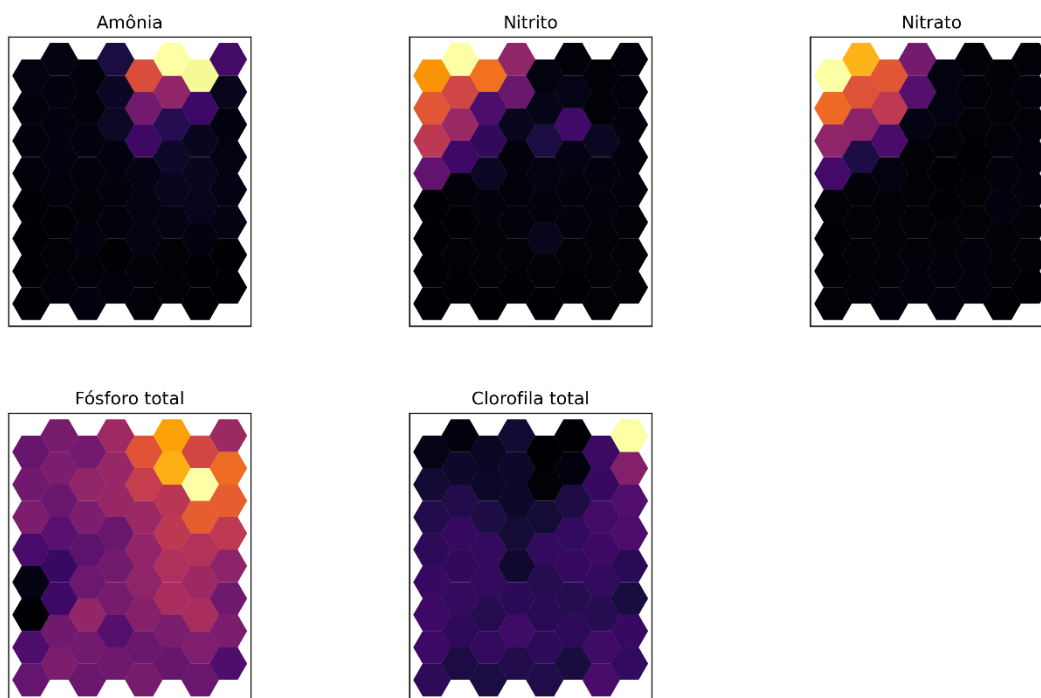
as regiões do mapa, indicando maior diferenciação entre os agrupamentos de amostras. O predomínio de valores elevados de pH reforça a característica alcalina do sistema, comum em reservatórios do semiárido, e pode influenciar a disponibilidade de nutrientes e a atividade biológica.

O potencial redox apresenta variações espaciais mais evidentes, indicando a coexistência de condições com diferentes estados de oxidação e redução, o que pode refletir variações na decomposição da matéria orgânica e na dinâmica do oxigênio dissolvido. A condutividade elétrica, por sua vez, apresenta regiões com maiores concentrações de sais dissolvidos, sugerindo influência de processos de evaporação e possíveis aportes externos, além de indicar maior mineralização da água em determinados agrupamentos.

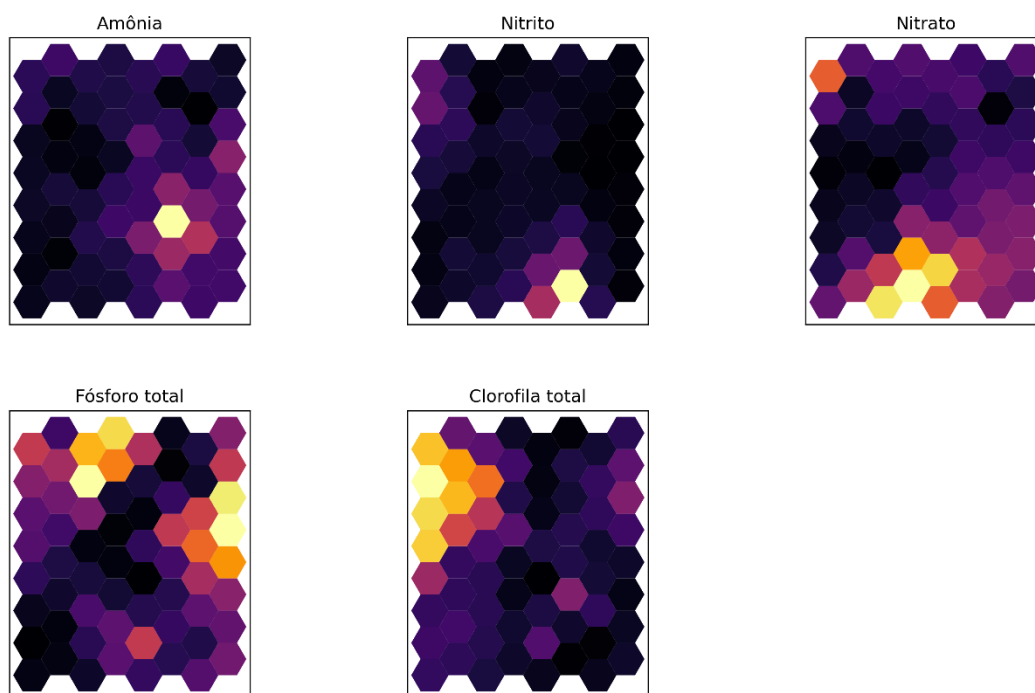
A turbidez, predominantemente baixa em Camalaú, sugere maior estabilidade da coluna d'água e menor influência de material particulado em suspensão quando comparado ao reservatório de Acauã, indicando condições potencialmente mais favoráveis à penetração de luz e ao desenvolvimento de processos fotossintéticos.

De forma comparativa, observa-se que, embora ambos os reservatórios apresentem características típicas de ambientes semiáridos, como tendência à alcalinidade e influência da salinização, o reservatório de Acauã apresenta maior variabilidade associada à presença de material particulado e heterogeneidade espacial, enquanto Camalaú evidencia maior estabilidade físico-química em relação à turbidez. Esses resultados indicam que os reservatórios, apesar de inseridos na mesma bacia, apresentam dinâmicas distintas, influenciadas por suas características hidrológicas, morfológicas e pelas condições ambientais locais. Os mapas de componentes relacionados a nutrientes e à biomassa fitoplanctônica dos reservatórios de Acauã e Camalaú são apresentados nas **Figuras 26 e 27**, respectivamente.

74

Figura 26 – Mapas de Componentes: variáveis químicas e biológicas de Acauã

Fonte: Autora (2026)

Figura 27 – Mapas de Componentes: variáveis químicas e biológicas de Camalaú

Fonte: Autora (2026)

No reservatório de Acauã (**Figura 26**), observa-se que os compostos nitrogenados

75

apresentam distribuição espacial bastante localizada, com regiões específicas do SOM associadas a maiores concentrações. A concentração de amônia em áreas pontuais sugere maior influência de matéria orgânica e processos de mineralização, possivelmente associados à decomposição de biomassa ou a aportes externos. Já as regiões com maiores concentrações de nitrito e nitrato indicam a atuação de processos de nitrificação, refletindo condições favoráveis à oxidação do nitrogênio e à atividade microbiana.

O fósforo total apresenta distribuição mais ampla ao longo do mapa, indicando que o enriquecimento por esse nutriente não ocorre de forma isolada, mas está presente em diferentes agrupamentos de amostras. Esse comportamento sugere uma influência mais difusa de fontes de fósforo, podendo estar associado tanto ao aporte externo quanto à liberação interna a partir dos sedimentos.

A clorofila total apresenta concentrações mais elevadas em regiões específicas que, em parte, coincidem com áreas de maior disponibilidade de fósforo, indicando que a produtividade fitoplanctônica está diretamente relacionada ao enriquecimento nutricional. Esse padrão evidencia a ocorrência de processos de eutrofização em determinados agrupamentos do reservatório, nos quais o aumento da disponibilidade de nutrientes favorece o crescimento do fitoplâncton.

No reservatório de Camalaú (**Figura 27**), observa-se uma organização espacial mais definida dos compostos nitrogenados, indicando maior diferenciação entre os processos que controlam a dinâmica do nitrogênio. O nitrato apresenta concentrações elevadas em regiões específicas do SOM, sugerindo condições mais oxigenadas e maior estabilidade na forma oxidada do nitrogênio. O nitrito, por sua vez, apresenta valores elevados em áreas mais localizadas, indicando possíveis condições transitórias no processo de nitrificação.

A amônia apresenta elevação pontual, o que pode indicar fontes localizadas de matéria orgânica ou baixa eficiência na oxidação do nitrogênio em determinadas condições. Esse padrão sugere a ocorrência de processos diferenciados dentro do reservatório, refletindo variações na dinâmica biogeoquímica.

O fósforo total apresenta elevada variabilidade espacial, indicando que o enriquecimento por esse nutriente ocorre de forma heterogênea no reservatório. A associação espacial entre regiões com maior fósforo e maiores concentrações de clorofila total indica que a disponibilidade de fósforo exerce controle direto sobre a produtividade fitoplanctônica, evidenciando a ocorrência de áreas com maior potencial eutrófico.

De forma comparativa, observa-se que o reservatório de Acauã apresenta padrões mais

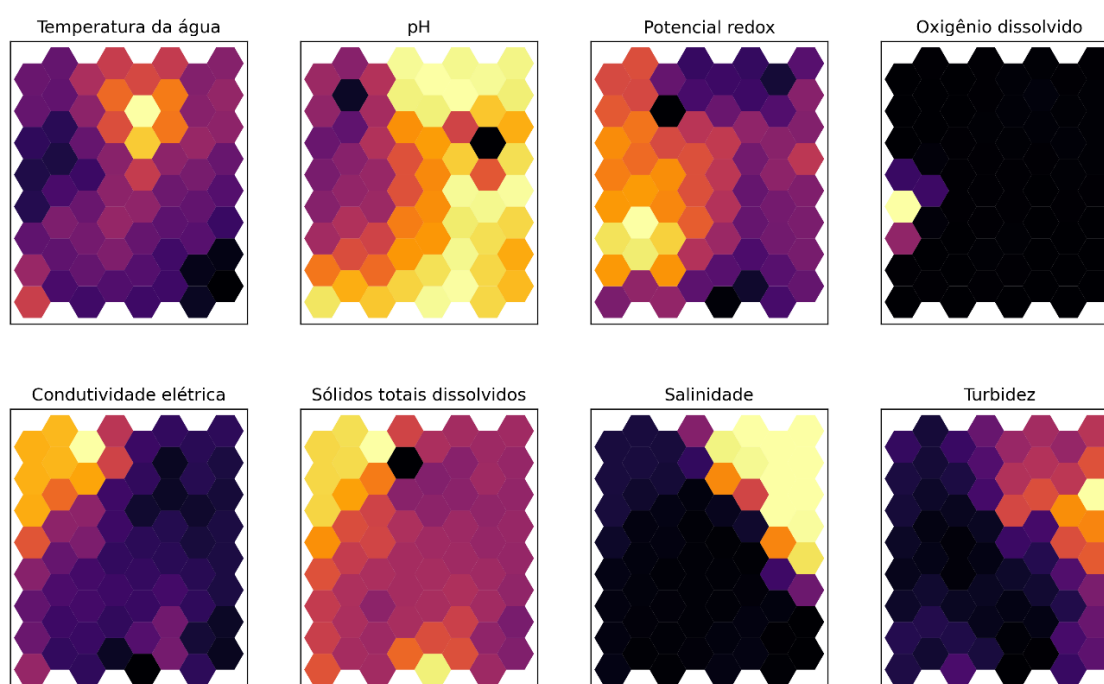
76

localizados de enriquecimento por nutrientes, sugerindo a influência de fontes pontuais ou processos específicos, enquanto Camalaú apresenta uma organização espacial mais estruturada dos nutrientes e da biomassa, indicando maior diferenciação entre estados tróficos. Esses resultados evidenciam que, embora ambos os reservatórios estejam sujeitos à influência de nutrientes, a dinâmica do enriquecimento e da resposta biológica ocorre de maneira distinta, refletindo as características ambientais e os processos internos de cada sistema.

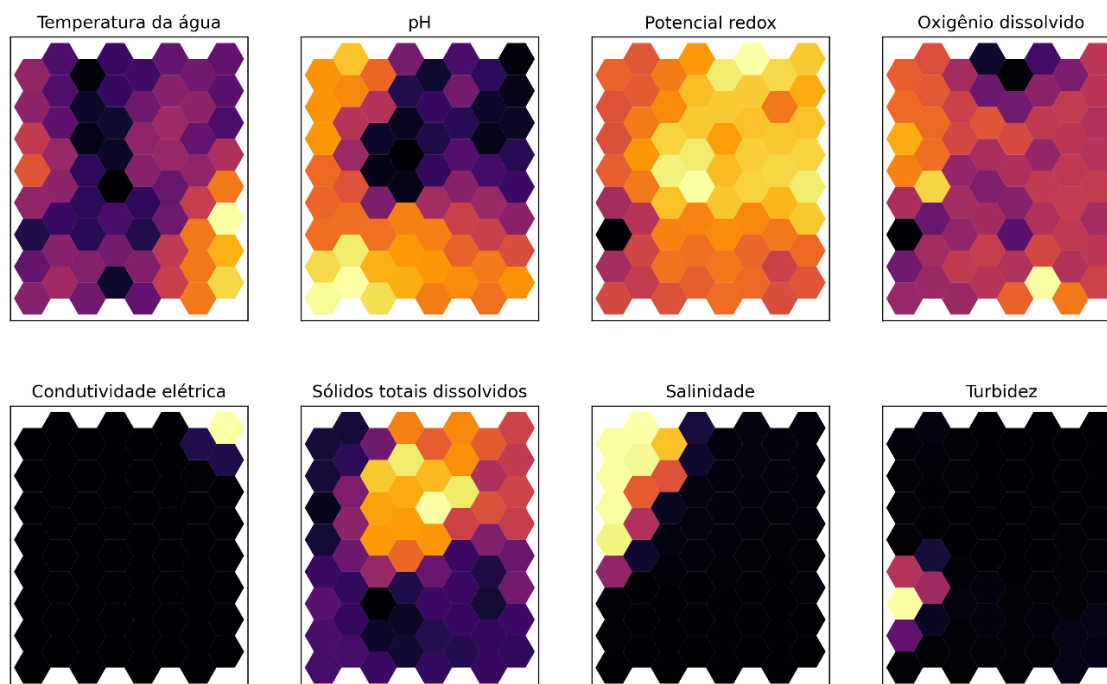
4.5.3.2 Cordeiro e Epitácio Pessoa

A distribuição espacial das variáveis físico-químicas nos reservatórios de Cordeiro e Epitácio Pessoa (**Figuras 28 e 29**) evidencia diferenças importantes na organização dos processos ambientais que controlam a qualidade da água nesses sistemas.

Figura 28 – Mapas de Componentes: variáveis físico-químicas de Cordeiro



Fonte: Autora (2026)

Figura 29 – Mapas de Componentes: variáveis físico-químicas de Epitácio Pessoa

Fonte: Autora (2026)

No reservatório de Cordeiro (**Figura 28**), observa-se uma organização espacial relativamente menos definida, com variações mais localizadas entre as variáveis, indicando menor heterogeneidade entre os agrupamentos de amostras. A concentração de valores mais elevados de temperatura na região central do mapa sugere a existência de condições térmicas específicas associadas a determinados conjuntos de amostras, possivelmente relacionadas à baixa profundidade ou maior exposição à radiação solar.

A condutividade elétrica e a salinidade apresentam elevações em regiões específicas do SOM, indicando variações locais na concentração de sais dissolvidos, possivelmente associadas à evaporação e à concentração iônica, características típicas de reservatórios em regiões semiáridas. No entanto, a distribuição mais localizada dessas variáveis sugere que esses processos não atuam de forma homogênea em todo o sistema.

O potencial redox apresenta valores mais elevados em determinadas regiões, indicando condições predominantemente oxidantes, o que sugere menor acúmulo de matéria orgânica e menor tendência à liberação interna de nutrientes a partir dos sedimentos. De forma geral, os padrões observados em Cordeiro indicam um sistema com menor complexidade espacial e maior homogeneidade relativa nas condições físico-químicas.

No reservatório Epitácio Pessoa (**Figura 29**), observa-se uma organização espacial mais estruturada, com a presença de gradientes bem definidos ao longo do mapa SOM, indicando

78

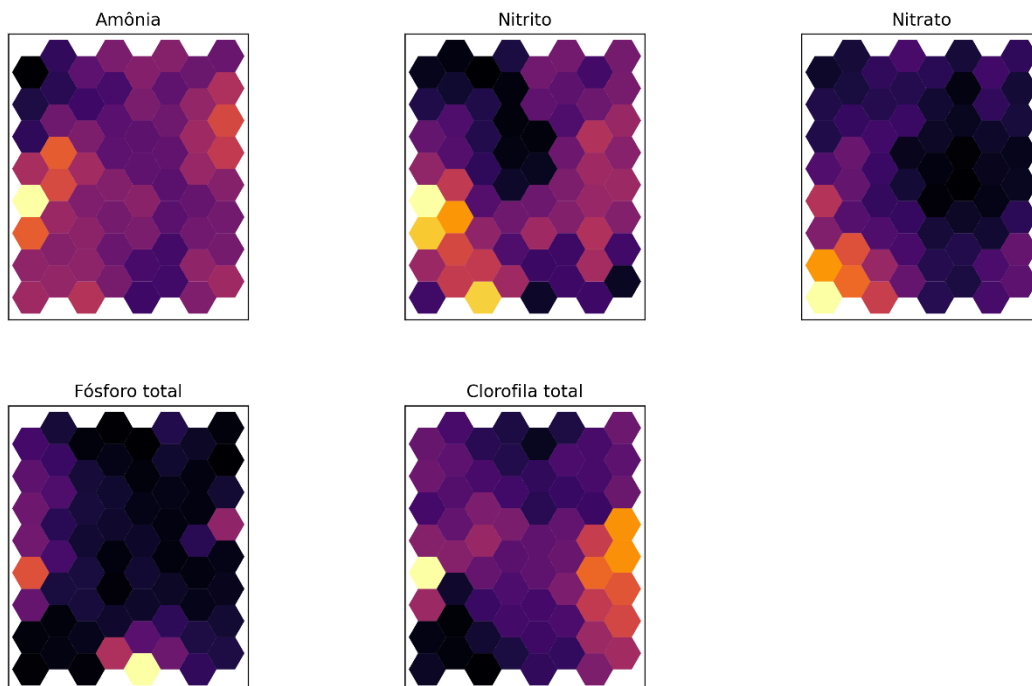
maior heterogeneidade ambiental. A variação da temperatura e do pH ao longo do mapa evidencia a existência de diferentes massas de água ou condições ambientais distintas, possivelmente associadas à dinâmica hidrológica do reservatório, como mistura, renovação e influência de aportes.

O potencial redox apresenta valores elevados em grande parte do mapa, indicando condições predominantemente oxidantes, enquanto a variação do oxigênio dissolvido sugere diferenças na intensidade dos processos biológicos e na renovação da coluna d'água entre os agrupamentos de amostras. A presença de regiões com maior oxigenação pode estar associada à maior atividade fotossintética ou a áreas com maior circulação hídrica.

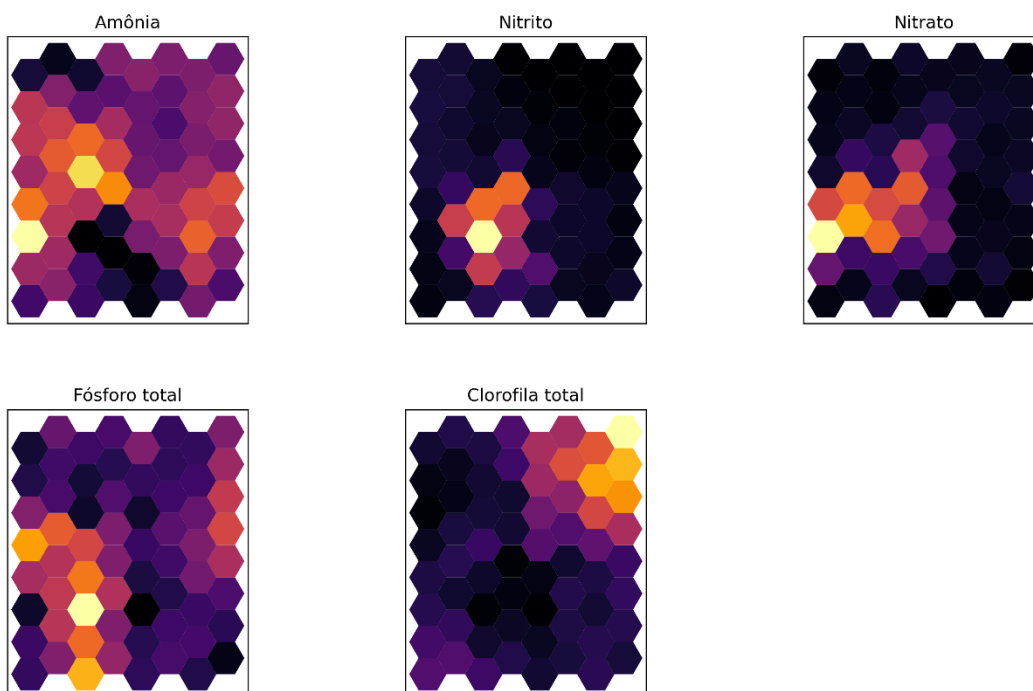
A condutividade elétrica, a salinidade e a turbidez apresentam padrões espaciais bem definidos, indicando a atuação de processos distintos de concentração de sais e de material particulado. A ocorrência de regiões com maiores valores dessas variáveis sugere a influência de aportes externos, processos de evaporação e possíveis zonas de maior ressuspensão de sedimentos.

De forma comparativa, observa-se que o reservatório de Eptácio Pessoa apresenta maior complexidade ambiental e maior heterogeneidade espacial quando comparado ao reservatório de Cordeiro, evidenciada pela formação de gradientes mais contínuos e estruturados no SOM. Enquanto Cordeiro apresenta padrões mais localizados e maior homogeneidade relativa, Eptácio Pessoa revela a atuação simultânea de múltiplos processos físicos, químicos e biológicos, indicando um sistema mais dinâmico e ambientalmente complexo. Os mapas de componentes de nutrientes e de biomassa fitoplanctônica são apresentados nas **Figuras 30 e 31**.

79

Figura 30 – Mapas de Componentes: variáveis químicas e biológicas de Cordeiro

Fonte: Autora (2026)

Figura 31 – Mapas de Componentes: variáveis químicas e biológicas de Epitácio Pessoa

Fonte: Autora (2026)

No reservatório de Cordeiro (**Figura 30**), observa-se que as regiões com maiores concentrações de nitrito e nitrato estão relativamente próximas no mapa, indicando a

80

predominância de formas oxidadas de nitrogênio. Esse padrão sugere a atuação de processos de nitrificação, evidenciando condições favoráveis à oxidação do nitrogênio e à atividade microbiana. A presença de valores moderados de amônia nessas regiões reforça a ideia de que o sistema apresenta uma dinâmica ativa de transformação do nitrogênio, com a amônia sendo convertida para formas mais oxidadas.

O fósforo total apresenta concentrações predominantemente baixas ao longo do mapa, com ocorrência pontual de valores mais elevados, indicando que o enriquecimento por fósforo é limitado e localizado. Essa condição sugere que o sistema pode estar sob limitação por fósforo, restringindo o desenvolvimento da biomassa fitoplanctônica.

A clorofila total apresenta aumento em uma região específica do SOM, indicando a existência de um agrupamento com maior produtividade primária. No entanto, a ausência de uma associação espacial direta com o fósforo total sugere que a biomassa fitoplanctônica pode estar sendo influenciada por fatores adicionais, como condições de luz, estabilidade da coluna d'água ou disponibilidade de outras formas de nutrientes.

No reservatório de Epitácio Pessoa (**Figura 31**), observa-se uma concentração espacial mais definida dos nutrientes, com regiões do mapa que apresentam simultaneamente maiores valores de amônia, nitrito e nitrato. Esse padrão indica a ocorrência de enriquecimento por nitrogênio e sugere a presença de áreas com maior aporte de nutrientes ou maior intensidade de processos biogeoquímicos.

A associação espacial entre os compostos nitrogenados e o fósforo total indica que, em determinadas regiões, ocorre enriquecimento conjunto por nutrientes, caracterizando condições potencialmente favoráveis ao desenvolvimento de processos de eutrofização. No entanto, a distribuição da clorofila total apresenta maior concentração em uma região parcialmente distinta, sugerindo que a resposta biológica não ocorre de forma imediata ou diretamente proporcional à disponibilidade de nutrientes. Esse descompasso entre nutrientes e biomassa fitoplanctônica pode indicar a influência de fatores limitantes adicionais, como condições hidrodinâmicas, tempo de residência da água ou disponibilidade de luz, evidenciando maior complexidade na resposta ecológica do sistema.

De forma comparativa, observa-se que o reservatório de Cordeiro apresenta um sistema mais limitado em termos de disponibilidade de fósforo e com dinâmica de nutrientes mais distribuída, enquanto Epitácio Pessoa evidencia maior enriquecimento nutricional e maior complexidade nos processos que controlam a produtividade fitoplanctônica. Esses resultados reforçam a diferença entre um sistema mais homogêneo e possivelmente limitado (Cordeiro) e

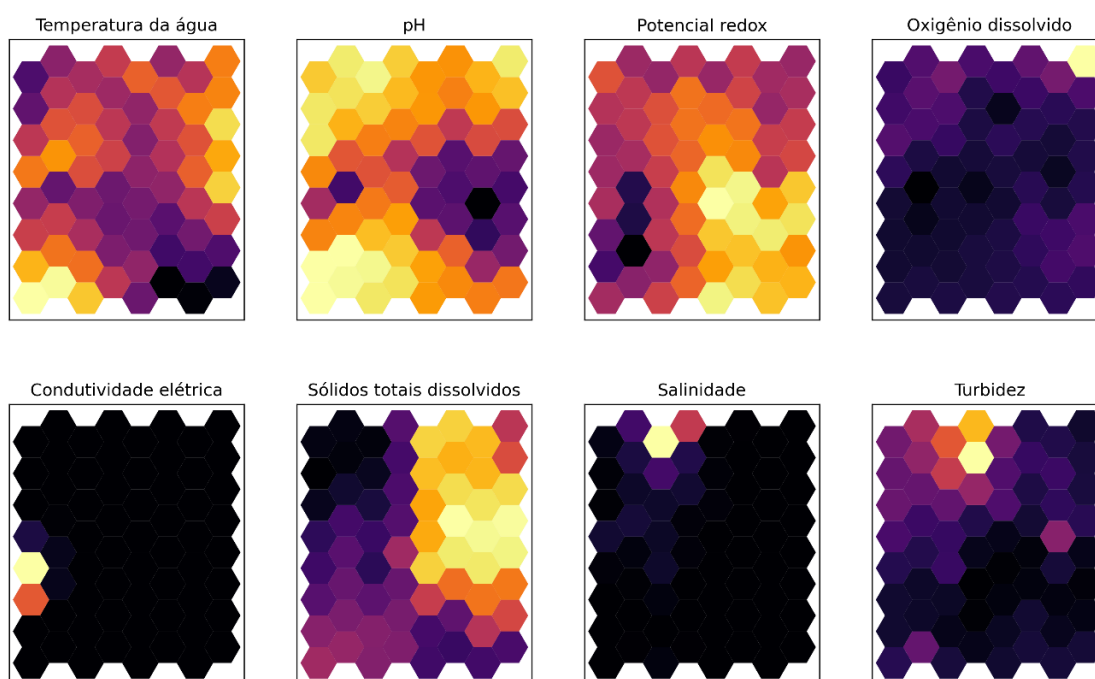
81

um sistema mais dinâmico e sujeito a processos de eutrofização (Epitácio Pessoa).

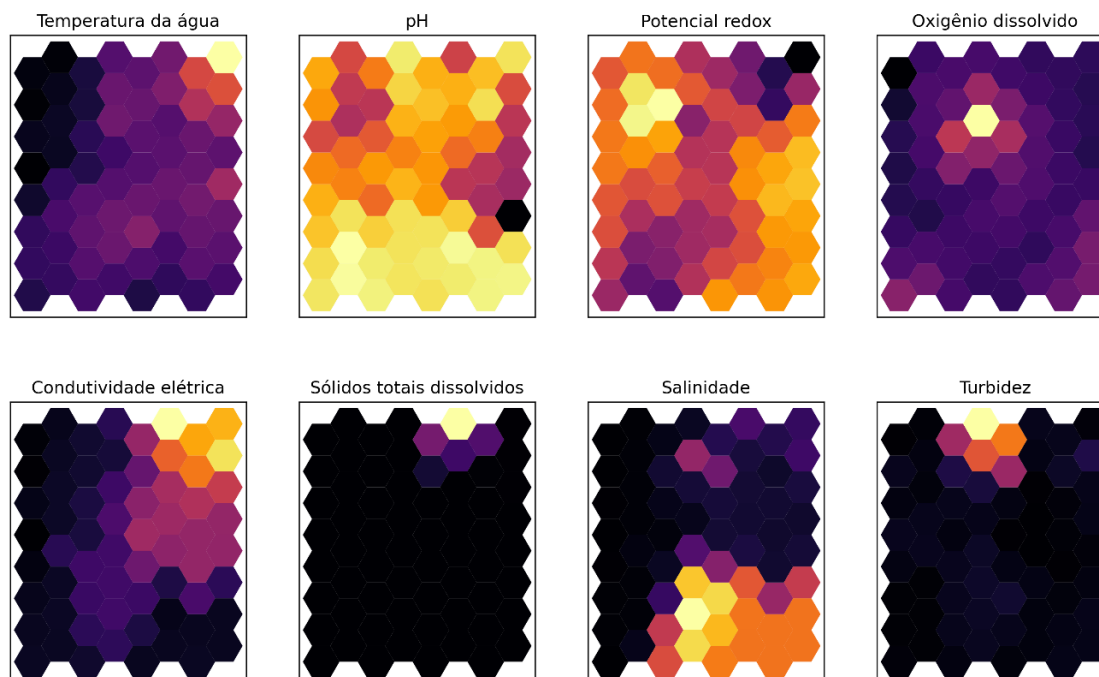
4.5.3.3 Poções e Sumé

Os mapas de componentes físico-químicos dos reservatórios de Poções e Sumé (**Figuras 32 e 33**) evidenciam diferenças importantes nos processos que controlam a qualidade da água nesses sistemas, especialmente quanto à dinâmica de sais dissolvidos.

Figura 32 – Mapas de Componentes: variáveis físico-químicas de Poções



Fonte: Autora (2026)

Figura 33 – Mapas de Componentes: variáveis físico-químicas de Sumé

Fonte: Autora (2026)

No reservatório de Poções (**Figura 32**), observa-se uma variação espacial moderada nas variáveis físico-químicas, indicando a presença de agrupamentos com características distintas, porém sem a formação de gradientes bem definidos. A variação da temperatura, do pH e do potencial redox sugere diferenças locais nas condições ambientais, possivelmente associadas à dinâmica de mistura da coluna d'água e à influência de fatores sazonais.

Os sólidos totais dissolvidos apresentam um gradiente mais evidente, indicando a existência de regiões com maior concentração de material dissolvido, o que pode estar associado à evaporação e à concentração de íons, processos típicos de reservatórios em regiões semiáridas. Por outro lado, a condutividade elétrica e a salinidade apresentam valores predominantemente baixos e pouco variáveis, sugerindo que a salinização não constitui um fator dominante na diferenciação das amostras nesse reservatório.

No reservatório de Sumé (**Figura 33**), observa-se um padrão distinto, caracterizado por gradientes bem definidos de condutividade elétrica e salinidade, indicando maior influência de processos de concentração de sais dissolvidos. Esse comportamento sugere forte atuação da evaporação e maior mineralização da água, características frequentemente associadas a sistemas com menor renovação hídrica. A distribuição espacial dessas variáveis indica a existência de zonas com maior acúmulo de sais, o que pode impactar diretamente a qualidade da água e influenciar a disponibilidade de nutrientes e a atividade biológica. Em contraste, os

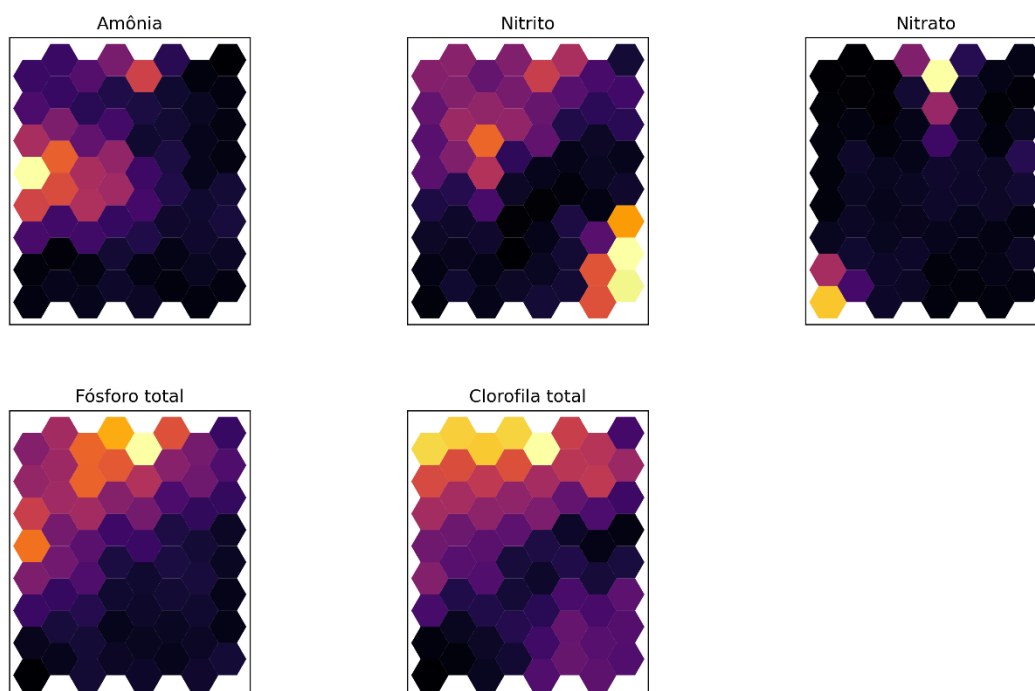
83

sólidos totais dissolvidos apresentam menor variabilidade, indicando que a composição da fração dissolvida pode estar mais associada à natureza iônica do que à quantidade total de sólidos.

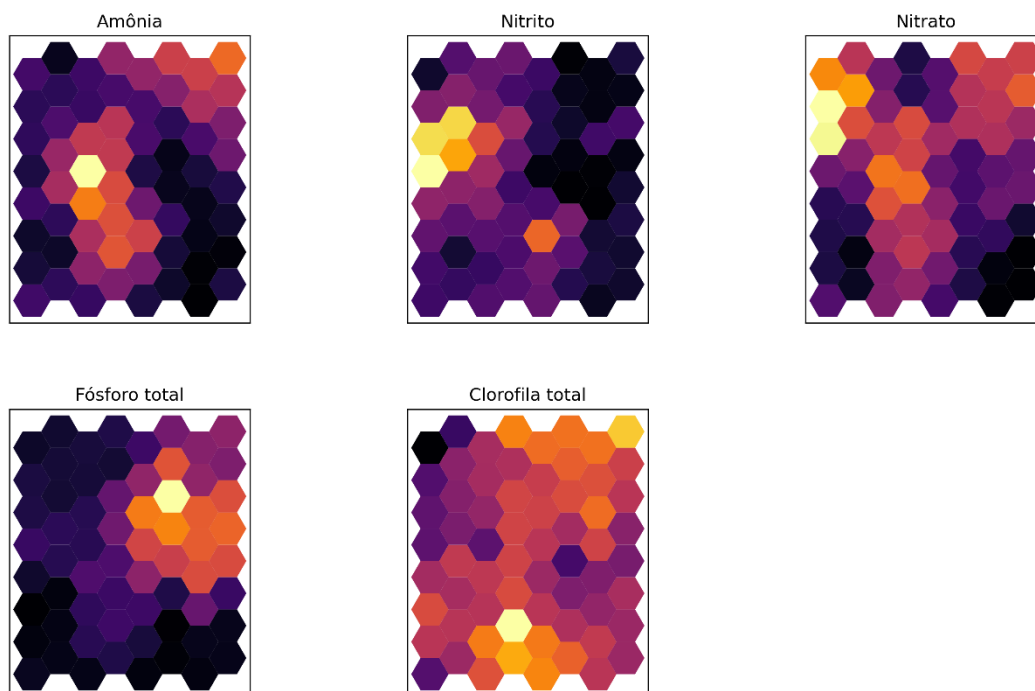
De forma comparativa, observa-se que Poções apresenta maior homogeneidade físico-química, com variações mais difusas, enquanto Sumé evidencia maior estruturação espacial e maior influência de processos de salinização, indicando um sistema mais condicionado pela dinâmica de concentração de sais.

Os mapas de nutrientes e de biomassa fitoplanctônica dos reservatórios de Poções e de Sumé são apresentados nas **Figuras 34 e 35**.

Figura 34 – Mapas de Componentes: variáveis químicas e biológicas de Poções



Fonte: Autora (2026)

Figura 35 – Mapas de Componentes: variáveis químicas e biológicas de Sumé

Fonte: Autora (2026)

No reservatório de Poções (**Figura 34**), observa-se que as diferentes formas de nitrogênio apresentam distribuição relativamente dispersa no mapa, indicando que o enriquecimento por esse nutriente não ocorre de forma concentrada em regiões específicas. Esse padrão sugere uma dinâmica mais difusa do nitrogênio, possivelmente associada à mistura da coluna d'água ou à ausência de fontes pontuais dominantes.

A amônia apresenta elevação em algumas regiões, indicando a influência de processos de decomposição da matéria orgânica e mineralização, enquanto o nitrato apresenta concentrações mais localizadas, sugerindo a atuação de processos de nitrificação em condições específicas. O fósforo total apresenta um gradiente bem definido, com regiões associadas a maiores concentrações que coincidem com áreas de maior clorofila total. Esse padrão evidencia a influência direta do fósforo sobre o crescimento fitoplanctônico, indicando que o reservatório apresenta áreas com maior potencial de eutrofização, nas quais o aumento da disponibilidade de nutrientes favorece o desenvolvimento da biomassa algal.

No reservatório de Sumé (**Figura 35**), observa-se maior organização espacial dos nutrientes, especialmente do nitrato, que apresenta gradiente bem definido ao longo do mapa, indicando regiões com maior disponibilidade dessa forma oxidada de nitrogênio. A amônia também apresenta elevação em determinadas áreas, enquanto o nitrito possui distribuição mais

85

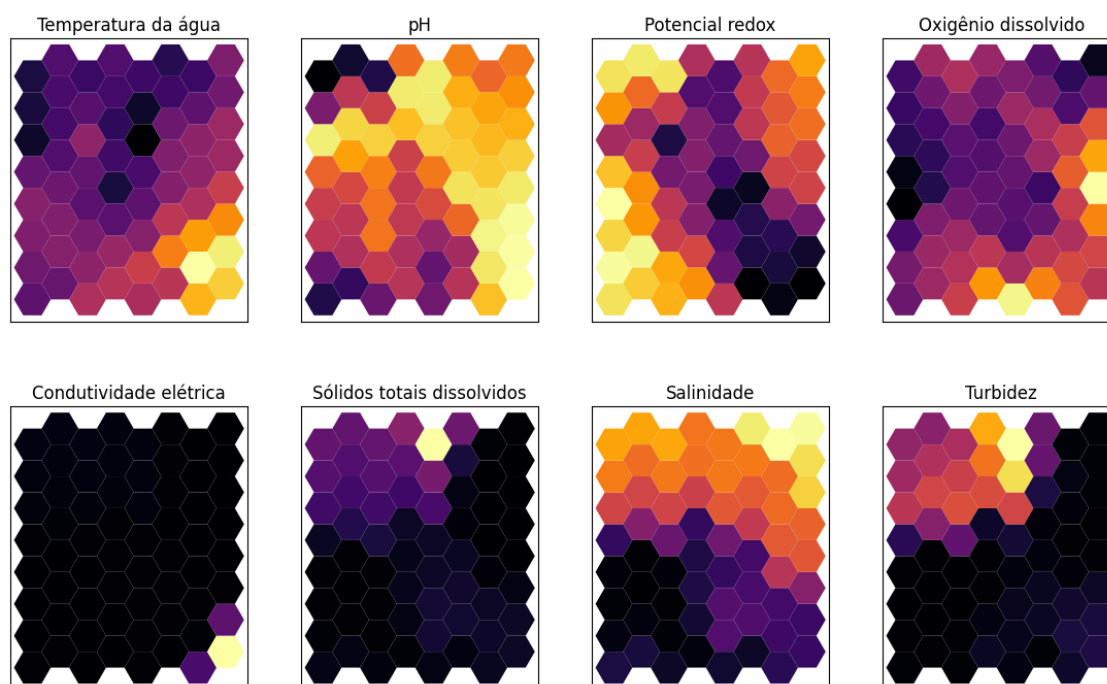
localizada, sugerindo a ocorrência de diferentes estágios do ciclo do nitrogênio no sistema. O fósforo total apresenta aumento em uma região central do SOM, indicando áreas com maior enriquecimento nutricional. A clorofila total, por sua vez, apresenta valores elevados distribuídos em grande parte do mapa, sugerindo que a biomassa fitoplanctônica está amplamente presente no reservatório.

Esse padrão indica que, diferentemente de Poções, onde a resposta biológica está mais associada a regiões específicas, em Sumé a biomassa fitoplanctônica apresenta distribuição mais generalizada, sugerindo um sistema com maior disponibilidade de nutrientes e maior propensão a processos de eutrofização em escala mais ampla.

4.5.3.4 Taperoá

Os mapas dos componentes físico-químicos do reservatório de Taperoá são apresentados na **Figura 36** e evidenciam uma organização espacial relativamente bem definida de algumas variáveis, apesar da menor complexidade geral do sistema, quando comparado a outros reservatórios analisados.

Figura 36 – Mapas de Componentes: variáveis físico-químicas de Taperoá



Fonte: Autora (2026)

O pH apresenta um gradiente claro ao longo do SOM, indicando a presença de

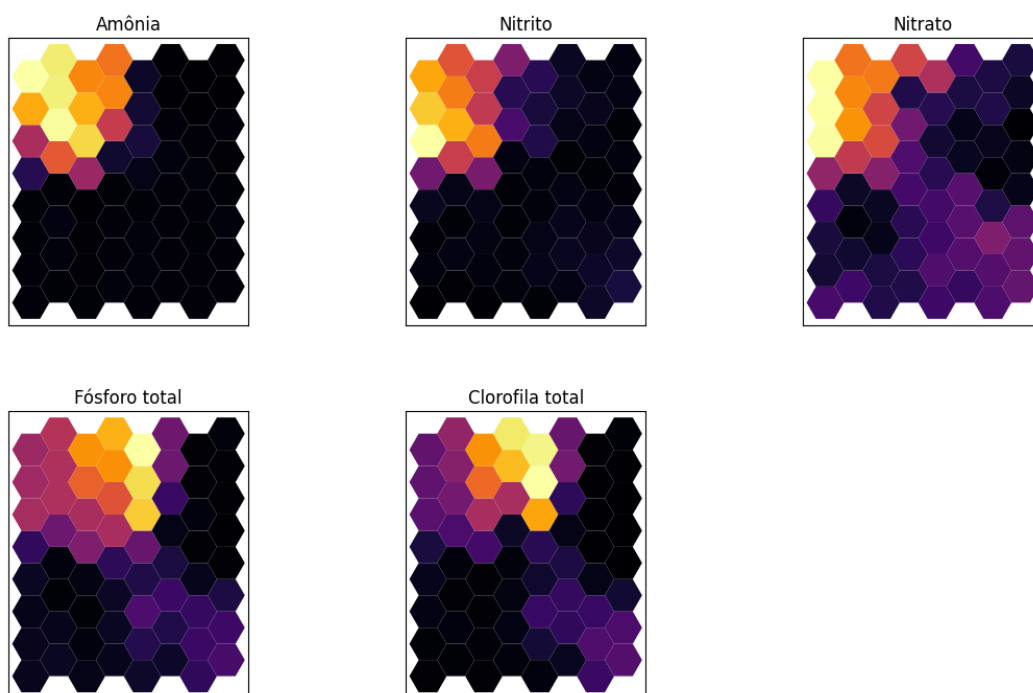
86

agrupamentos associados a condições mais alcalinas, característica comum em ambientes do semiárido, onde processos de evaporação contribuem para o aumento da alcalinidade da água. O potencial redox exhibe variações espaciais contrastantes, indicando a coexistência de regiões com diferentes condições de oxidação e redução, o que pode refletir variações locais na disponibilidade de oxigênio e na intensidade dos processos biogeoquímicos.

A salinidade apresenta um gradiente bem definido, com maiores concentrações em uma porção do mapa, sugerindo influência de processos de concentração de sais dissolvidos, associados à evaporação e à dinâmica hidrológica do reservatório. No entanto, a condutividade elétrica apresenta valores predominantemente baixos e pouca variação, indicando que, embora exista concentração de sais, essa não se apresenta como um fator dominante na diferenciação global das amostras.

Os sólidos totais dissolvidos apresentam elevação pontual, sugerindo que a presença de material dissolvido ocorre de forma localizada e não generalizada. A turbidez, por sua vez, apresenta aumento em uma região específica do mapa, indicando a ocorrência localizada de material particulado em suspensão, possivelmente associada à ressuspensão de sedimentos ou a eventos hidrológicos pontuais.

De forma geral, os padrões observados indicam um sistema com organização espacial moderada, porém com menor heterogeneidade quando comparado a reservatórios mais complexos, sugerindo maior estabilidade relativa nas condições físico-químicas. Os mapas de componentes relativos a nutrientes e à biomassa fitoplanctônica do reservatório de Taperoá são apresentados na **Figura 37**.

Figura 37 – Mapas de Componentes: variáveis químicas e biológicas de Taperoá

Fonte: Autora (2026)

Observa-se que as maiores concentrações das formas de nitrogênio (amônia, nitrito e nitrato) estão concentradas em uma mesma região do SOM, indicando a ocorrência de enriquecimento localizado por esse nutriente. A presença simultânea dessas diferentes formas sugere a atuação conjunta de processos de mineralização e nitrificação, evidenciando uma dinâmica ativa do ciclo do nitrogênio nessas áreas específicas do reservatório. O fósforo total apresenta um padrão semelhante, com maiores concentrações concentradas na mesma região do mapa, indicando enriquecimento nutricional localizado. A coincidência espacial entre fósforo e formas de nitrogênio sugere a presença de fontes comuns de nutrientes ou a atuação de processos internos que favorecem a acumulação desses elementos.

A clorofila total apresenta elevação nas mesmas regiões de maior disponibilidade de nutrientes, indicando uma resposta direta da biomassa fitoplanctônica ao enriquecimento nutricional. Esse padrão evidencia a ocorrência de processos de eutrofização localizada, nos quais o aumento da disponibilidade de nutrientes favorece o crescimento do fitoplâncton em áreas específicas do reservatório.

De forma integrada, os resultados obtidos a partir dos mapas de componentes dos diferentes reservatórios evidenciam que a qualidade da água na bacia do Rio Paraíba é controlada por um conjunto de processos físicos, químicos e biológicos que atuam de maneira distinta em cada sistema. Observa-se a coexistência de reservatórios com maior homogeneidade

88

e estabilidade ambiental, como Taperoá e Cordeiro, e sistemas mais complexos e heterogêneos, como Eptácio Pessoa e Sumé, nos quais há maior influência de processos como salinização, enriquecimento nutricional e eutrofização.

4.5.4 Clusterização dos SOM individuais

Para a etapa de clusterização dos mapas auto-organizáveis individuais, adotou-se o mesmo número de clusters do modelo SOM integrado, ou seja, quatro agrupamentos. A adoção desse mesmo número de clusters teve como objetivo manter a consistência metodológica entre as diferentes modelagens, permitindo uma comparação direta entre os padrões identificados no modelo integrado e os observados nos modelos individuais de cada reservatório. Esse número também foi suficiente para representar os principais gradientes ambientais observados nos mapas SOM.

O algoritmo k-means utiliza numeração iniciada em zero para identificar os agrupamentos (0–3). Para facilitar a apresentação e interpretação dos resultados, esses rótulos foram apresentados ao longo do texto como Clusters 1 a 4, mantendo correspondência direta com a numeração original do modelo (0 → Cluster 1, 1 → Cluster 2, 2 → Cluster 3, 3 → Cluster 4).

A distribuição das amostras entre os clusters identificados nos SOMs individuais de cada reservatório é apresentada na **Tabela 5**. Observa-se que a representatividade dos agrupamentos varia entre os reservatórios analisados, indicando diferenças na frequência de ocorrência dos padrões ambientais identificados em cada sistema. Em alguns casos, verifica-se predominância de um cluster específico, como observado em Taperoá, onde a maior parte das amostras está concentrada no Cluster 2. Situação semelhante ocorre em Eptácio Pessoa, onde o Cluster 3 apresenta maior representatividade. Por outro lado, reservatórios como Camalaú e Cordeiro apresentam uma distribuição mais equilibrada entre os agrupamentos, indicando maior variabilidade nas condições ambientais representadas no conjunto de dados.

89

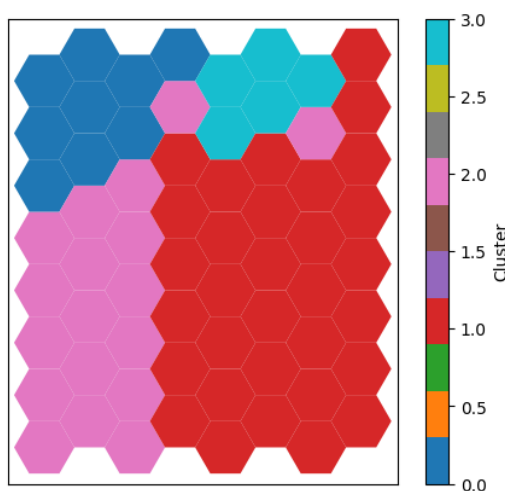
Tabela 5 – Distribuição das amostras por cluster identificado nos mapas individuais dos reservatórios analisados

	Cluster 1	Cluster 2	Cluster 3	Cluster 4
Acauã	21	39	76	12
Camalaú	42	18	33	25
Cordeiro	51	34	16	1
Epitácio Pessoa	42	7	76	1
Poçoões	26	53	43	1
Sumé	57	7	24	23
Taperoá	8	101	16	1

Fonte: Autora (2026)

Após a identificação da distribuição das amostras entre os clusters (**Tabela 5**), torna-se possível analisar a organização espacial desses agrupamentos nos mapas auto-organizáveis de cada reservatório. A seguir, são apresentados os mapas de clusters obtidos para cada reservatório, permitindo observar a organização espacial dos agrupamentos identificados e sua relação com os padrões previamente observados nos mapas de componentes.

A distribuição espacial dos clusters no SOM do reservatório de Acauã é apresentada na **Figura 38**. Observa-se a formação de quatro agrupamentos distintos no mapa, indicando a presença de diferentes condições ambientais entre as amostras analisadas.

Figura 38 – Mapa de clusters do SOM de Acauã

Fonte: Autora (2026)

O Cluster 2 (vermelho) apresenta maior extensão no mapa, ocupando grande parte da região central e da direita do SOM. O Cluster 1 (azul) concentra-se principalmente na porção superior esquerda, enquanto o Cluster 3 (rosa) aparece em uma região mais restrita do mapa, localizada na porção inferior esquerda. O Cluster 4 (verde) constitui um agrupamento menor

90

localizado na região superior central do SOM. Embora o Cluster 3 ocupe uma área relativamente pequena no mapa, concentra o maior número de amostras (**Tabela 5**). Isso ocorre porque diferentes observações podem ser mapeadas para o mesmo neurônio vencedor (BMU), indicando que determinados padrões ambientais são recorrentes no conjunto de dados.

De acordo com as médias das variáveis associadas a cada agrupamento (**Tabela 6**), os clusters apresentam diferenças marcantes nas condições limnológicas do reservatório.

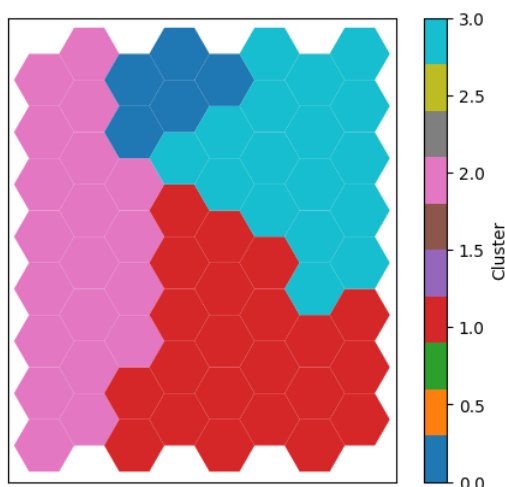
Tabela 6 – Valores médios das variáveis ambientais para os clusters identificados no SOM individual do reservatório de Acauã

	Cluster 1	Cluster 2	Cluster 3	Cluster 4
Temperatura da água (°C)	26.93	28.55	28.74	27.72
pH	7.77	9.38	8.32	7.52
Potencial redox	207.63	129.44	125.76	-210.58
Condutividade elétrica (mS cm ⁻¹)	1.91	1.36	1.73	1.95
Turbidez (NTU)	23.28	72.77	56.16	17.1
Oxigênio dissolvido (mg/L)	6.09	7.63	6.50	1.72
Sólidos totais dissolvidos (g L ⁻¹)	1.22	0.87	13.30	1.25
Salinidade (ppt)	0.10	0.65	0.09	0.10
Amônia (µg L ⁻¹)	31.07	31.96	32.44	469.55
Nitrito (µg L ⁻¹)	83.00	2.61	3.10	3.88
Nitrato (µg L ⁻¹)	847.27	26.97	29.49	26.14
Fósforo total (µg L ⁻¹)	281.37	300.8	270.38	395.61
Clorofila total (µg L ⁻¹)	73.72	150.49	115.58	39.82

Fonte: Autora (2026)

O Cluster 2 apresenta os maiores valores de pH, turbidez e clorofila total, indicando condições associadas a maior biomassa fitoplantônica. O Cluster 3 caracteriza-se pelos maiores valores de sólidos totais dissolvidos, sugerindo maior concentração de material dissolvido na água. Por sua vez, o Cluster 1 apresenta elevadas concentrações de nitrato e nitrito em comparação aos demais agrupamentos. Já o Cluster 4 destaca-se por apresentar baixos valores de oxigênio dissolvido, associados a potencial redox negativa e a elevada concentração de amônia, indicando condições mais redutoras da água e possíveis processos de decomposição de matéria orgânica.

A distribuição espacial dos clusters no SOM do reservatório de Camalaú é apresentada na **Figura 39**.

Figura 39 – Mapa de clusters do SOM de Camalaú

Fonte: Autora (2026)

No mapa, observa-se que o Cluster 3 (rosa) ocupa predominantemente a região esquerda do SOM, enquanto o Cluster 1 (azul) se concentra na porção superior central do mapa. O Cluster 2 (verde) distribui-se principalmente na região direita do SOM, enquanto o Cluster 4 (vermelho) ocupa grande parte da porção inferior e central do mapa, indicando a presença de diferentes condições ambientais entre os agrupamentos identificados. As diferenças entre esses agrupamentos tornam-se mais evidentes ao analisar os valores médios das variáveis ambientais associados a cada cluster (**Tabela 7**).

Tabela 7 – Valores médios das variáveis ambientais para os clusters identificados no SOM individual do reservatório de Camalaú

	Cluster 1	Cluster 2	Cluster 3	Cluster 4
Temperatura da água (°C)	25.80	26.46	26.13	27.03
pH	8.70	8.39	5.22	8.76
Potencial redox	151.19	217.5	270.61	183.04
Condutividade elétrica (mS cm⁻¹)	0.20	0.30	0.27	0.18
Turbidez (NTU)	5.75	13.18	5.10	22.49
Oxigênio dissolvido (mg/L)	2.96	6.41	3.91	5.30
Sólidos totais dissolvidos (g L⁻¹)	0.13	0.20	0.24	0.12
Salinidade (ppt)	0.02	0.13	0.01	0.06
Amônia (µg L⁻¹)	43.30	27.27	39.86	29.86
Nitrito (µg L⁻¹)	1.96	1.24	0.64	2.07
Nitrato (µg L⁻¹)	29.10	37.89	21.10	14.80
Fósforo total (µg L⁻¹)	23.23	8.02	39.45	27.33
Clorofila total (µg L⁻¹)	10.50	12.34	11.54	27.98

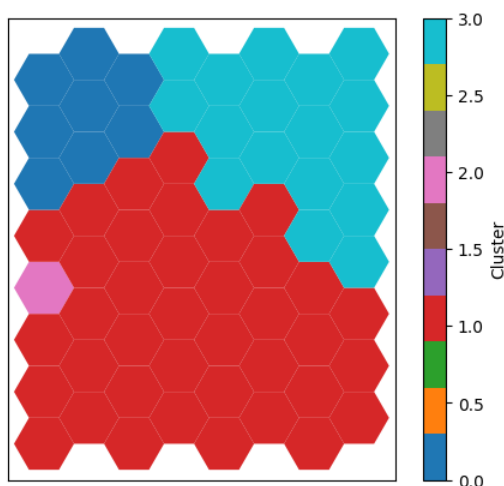
Fonte: Autora (2026)

92

De acordo com as médias das variáveis, os clusters apresentam diferenças nas condições limnológicas observadas no reservatório. O Cluster 4 apresenta os maiores valores de temperatura da água, pH, turbidez e clorofila total, indicando condições associadas à maior biomassa fitoplanctônica. O Cluster 3, por sua vez, apresenta os maiores valores de potencial redox e de fósforo total, sugerindo condições mais oxidantes e maior disponibilidade desse nutriente. O Cluster 2 caracteriza-se por maiores valores de oxigênio dissolvido, nitrato e salinidade, indicando condições mais oxigenadas da coluna d'água. Já o Cluster 1 apresenta maiores concentrações de amônia e nitrito em comparação aos demais agrupamentos.

A distribuição espacial dos clusters no SOM do reservatório de Cordeiro é apresentada na **Figura 40**.

Figura 40 – Mapa de clusters do SOM de Cordeiro



Fonte: Autora (2026)

No mapa, o Cluster 1 (azul) aparece concentrado na porção superior esquerda, enquanto o Cluster 2 (vermelho) ocupa grande parte da região central e inferior do SOM. Diferentemente do observado em Acuã e Camalaú, o Cluster 3 (rosa) ocorre em uma área bastante restrita no mapa, indicando menor representatividade desse padrão ambiental na estrutura do SOM. O cluster 4 (verde) distribui-se principalmente nas regiões superiores e direitas do mapa, apresentando padrão de distribuição semelhante ao observado para o reservatório de Camalaú.

As diferenças entre os agrupamentos podem ser melhor compreendidas a partir da análise dos valores médios das variáveis ambientais apresentados na **Tabela 8**.

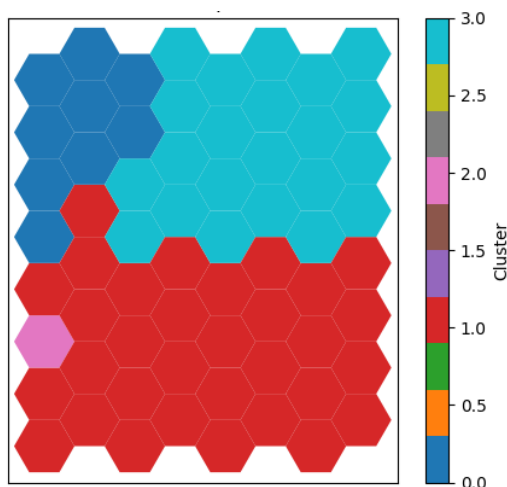
Tabela 8 – Valores médios das variáveis ambientais para os clusters identificados no SOM individual do reservatório de Cordeiro

	Cluster 1	Cluster 2	Cluster 3	Cluster 4
Temperatura da água (°C)	26.13	27.72	26.28	25.21
pH	8.11	8.98	5.65	5.83
Potencial redox	226.27	172.21	245.23	310
Condutividade elétrica (mS cm ⁻¹)	0.5	0.48	0.76	0.57
Turbidez (NTU)	2.87	15.66	3.61	1.4
Oxigênio dissolvido (mg/L)	7.1	10.58	7.2	6.3
Sólidos totais dissolvidos (g L ⁻¹)	0.33	0.3	0.49	0.38
Salinidade (ppt)	0.02	0.17	0.04	0.03
Amônia (µg L ⁻¹)	34.7	36.89	20.62	77.5
Nitrito (µg L ⁻¹)	1.38	1.46	0.67	4.44
Nitrato (µg L ⁻¹)	22.25	13	11.68	48.29
Fósforo total (µg L ⁻¹)	24.78	15.38	42.21	110.33
Clorofila total (µg L ⁻¹)	10.97	10.54	10.36	23.7

Fonte: Autora (2026)

A análise dos valores médios da **Tabela 8** evidencia diferenças importantes entre os agrupamentos identificados. O Cluster 4 apresenta os maiores valores para diversas variáveis, incluindo potencial redox, oxigênio dissolvido, amônia, nitrito, nitrato, fósforo total e clorofila total, indicando áreas mais eutrofizadas. As demais variáveis distribuem-se principalmente entre os Clusters 2 e 3, refletindo condições intermediárias entre os diferentes padrões ambientais representados no reservatório. A condutividade elétrica e os sólidos totais dissolvidos apresentam valores mais elevados no Cluster 3, sugerindo maior concentração de material dissolvido na água nesse agrupamento específico.

A organização espacial dos agrupamentos obtidos para o reservatório de Eptácio Pessoa pode ser visualizada na **Figura 41**.

Figura 41 – Mapa de clusters do SOM de Epitácio Pessoa

Fonte: Autora (2026)

O mapa de clusters do reservatório de Epitácio Pessoa apresenta uma distribuição espacial parcialmente semelhante à observada em Cordeiro, embora com diferenças na representatividade dos agrupamentos. Observa-se a predominância dos Clusters 2 (vermelho) e 4 (verde) na estrutura do mapa, sendo o Cluster 2 o mais extenso. O Cluster 3 (rosa) ocorre em uma área bastante restrita na porção esquerda do SOM, o que indica menor frequência desse padrão ambiental. Já o Cluster 1 (azul) aparece concentrado na região superior esquerda do mapa, em posição semelhante à observada em Cordeiro, porém com área ligeiramente maior.

A **tabela 9** apresenta os valores médios das variáveis para cada cluster identificado no reservatório. De modo geral, os clusters evidenciam diferenças claras entre os padrões limnológicos representados no conjunto de dados.

Tabela 9 – Valores médios das variáveis ambientais para os clusters identificados no SOM individual do reservatório de Epitácio Pessoa

	Cluster 1	Cluster 2	Cluster 3	Cluster 4
Temperatura da água (°C)	27.00	26.49	27.54	26.91
pH	6.42	8.12	8.93	6.03
Potencial redox	251.38	147.57	154.54	254.00
Condutividade elétrica (mS cm ⁻¹)	0.46	0.39	0.37	256.00
Turbidez (NTU)	4.72	297.06	18.21	5.10
Oxigênio dissolvido (mg/L)	6.19	5.74	7.90	7.30
Sólidos totais dissolvidos (g L ⁻¹)	0.29	0.25	0.24	0.28
Salinidade (ppt)	0.02	0.02	0.07	0.02
Amônia (µg L ⁻¹)	29.78	34.06	25.05	24.32
Nitrito (µg L ⁻¹)	0.60	18.73	2.58	0
Nitrato (µg L ⁻¹)	29.51	152.57	22.05	12.57

Fósforo total ($\mu\text{g L}^{-1}$)	12.56	18.90	12.90	15.33
Clorofila total ($\mu\text{g L}^{-1}$)	27.91	11.94	14.53	63.80

Fonte: Autora (2026)

O Cluster 2 apresenta valores bastante elevados de turbidez (297,06 NTU) e nitrato ($152,57 \mu\text{g L}^{-1}$), além de concentrações relativamente altas de nitrito e fósforo total, indicando condições associadas à maior presença de material particulado e nutrientes dissolvidos na água. Esse conjunto de variáveis sugere episódios de maior aporte de nutrientes ou de ressuspensão de sedimentos no reservatório, registrados durante as campanhas de monitoramento. O Cluster 3 destaca-se pelos maiores valores de oxigênio dissolvido ($7,90 \text{ mg L}^{-1}$) e de salinidade (0,07 ppt), indicando condições de água relativamente mais oxigenadas em comparação aos demais agrupamentos. Por sua vez, o Cluster 4 apresenta valores elevados de clorofila total ($63,80 \mu\text{g L}^{-1}$) e de potencial redox (254), sugerindo ambientes associados a maior atividade biológica e a possíveis episódios de aumento da biomassa fitoplanctônica.

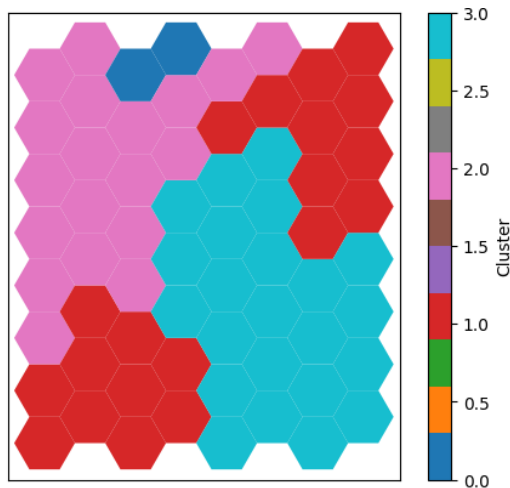
Observa-se, ainda, que alguns agrupamentos apresentam valores médios bastante elevados para determinadas variáveis, como ocorre com a condutividade elétrica no Cluster 4. A verificação do banco de dados original confirmou a presença de algumas observações com valores extremos, que foram mantidas na análise, uma vez que não foi realizado tratamento ou remoção de outliers no conjunto de dados. Dessa forma, tais valores refletem condições específicas registradas durante as campanhas de monitoramento e podem influenciar as médias de determinados agrupamentos.

De forma geral, os resultados indicam que o reservatório de Epitácio Pessoa apresenta variabilidade significativa nas condições físico-químicas e biológicas, refletida na diferenciação dos clusters identificados pelo SOM.

A organização espacial dos agrupamentos identificados nos mapas auto-organizáveis dos reservatórios de Poções e Sumé é apresentada nas **Figuras 42 e 43**, permitindo visualizar como os diferentes padrões ambientais foram distribuídos na grade do SOM de cada sistema.

Figura 42 – Mapa de clusters do SOM de

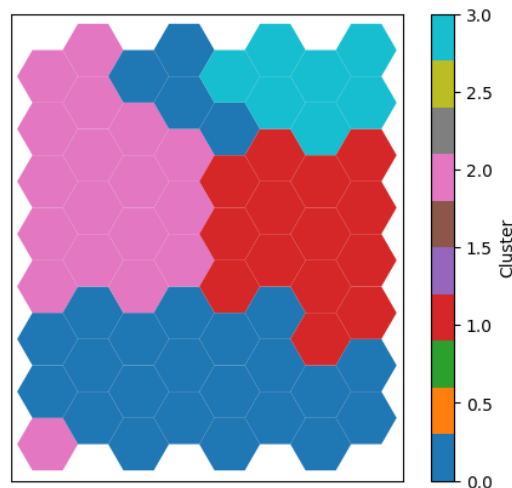
Poções



Fonte: Autora (2026)

Figura 43 – Mapa de clusters do SOM de

Sumé



Fonte: Autora (2026)

No reservatório de Poções, observa-se uma divisão relativamente clara entre os agrupamentos no mapa. O Cluster 4 (verde) ocupa grande parte da região central e inferior direita do SOM, representando o agrupamento mais extenso no mapa. O Cluster 2 (vermelho) aparece principalmente na porção direita e em parte da região inferior esquerda da grade, formando áreas relativamente contínuas. O Cluster 3 (rosa) distribui-se predominantemente ao longo da porção esquerda do mapa, ocupando uma faixa relativamente extensa nessa região. Já o Cluster 1 (azul) aparece de forma mais restrita, concentrando-se principalmente na região superior central do SOM. Essa organização espacial evidencia a presença de diferentes padrões ambientais no reservatório.

No reservatório de Sumé, a distribuição espacial dos clusters apresenta uma configuração distinta em relação ao reservatório de Poções. Observa-se a predominância do Cluster 1 (azul) no mapa, que ocupa grande parte da grade do SOM. O Cluster 2 (vermelho) está concentrado principalmente na região central e direita do mapa, formando uma faixa relativamente contínua nessa área. O Cluster 4 (verde) ocorre em pequena proporção na porção superior direita do SOM, enquanto o Cluster 3 (rosa) distribui-se predominantemente na região superior esquerda da grade.

Após a identificação da organização espacial dos agrupamentos no SOM, torna-se possível analisar as características ambientais associadas a cada cluster. Os valores médios das variáveis físico-químicas e biológicas do reservatório de Poções são apresentados na **Tabela 10**.

Tabela 10 – Valores médios das variáveis ambientais para os clusters identificados no SOM individual do reservatório de Poções

	Cluster 1	Cluster 2	Cluster 3	Cluster 4
Temperatura da água (°C)	26.76	26.58	27.47	26.5
pH	9.09	6.77	8.93	7.20
Potencial redox	177.85	245.21	162.19	196.00
Condutividade elétrica (mS cm ⁻¹)	0.16	0.22	0.21	164
Turbidez (NTU)	57.77	14.09	16.46	20.5
Oxigênio dissolvido (mg/L)	8.27	6.72	5.90	5.92
Sólidos totais dissolvidos (g L ⁻¹)	0.10	0.14	0.12	0.11
Salinidade (ppt)	0.10	0.01	0.02	0.01
Amônia (µg L ⁻¹)	66.00	27.76	43.2	79.77
Nitrito (µg L ⁻¹)	2.62	1.60	0.81	0.62
Nitrato (µg L ⁻¹)	19.33	15.65	21.59	4.00
Fósforo total (µg L ⁻¹)	80.14	29.26	29.33	43.67
Clorofila total (µg L ⁻¹)	81.21	35.30	31.01	56.60

Fonte: Autora (2026)

De modo geral, observa-se que o Cluster 1 concentra os maiores valores para a maioria das variáveis analisadas, incluindo pH, turbidez, salinidade, oxigênio dissolvido, nitrito, fósforo total e clorofila total. O Cluster 2 destaca-se principalmente pelo potencial redox, indicando condições relativamente mais oxidantes em comparação aos demais agrupamentos. Já o Cluster 3 apresenta os maiores valores de temperatura da água e de nitrato, enquanto o Cluster 4 se destaca pelos maiores valores de condutividade elétrica e de amônia, sugerindo maior concentração de íons dissolvidos na água.

De forma geral, a predominância de valores elevados em diversas variáveis no Cluster 1 indica que esse agrupamento representa um conjunto de amostras associado a condições limnológicas mais intensas, possivelmente relacionadas à maior disponibilidade de nutrientes e maior produtividade biológica no reservatório.

Os dados da **Tabela 11** evidenciam as diferenças limnológicas no reservatório de Sumé.

Tabela 11 – Valores médios das variáveis ambientais para os clusters identificados no SOM individual do reservatório de Sumé

	Cluster 1	Cluster 2	Cluster 3	Cluster 4
Temperatura da água (°C)	26.42	29.57	24.96	26.98
pH	8.92	8.02	7.27	6.46
Potencial redox	178.25	94.77	231.71	230.22
Condutividade elétrica (mS cm ⁻¹)	0.74	1.97	0.68	1.24
Turbidez (NTU)	66.38	62.36	45.70	25.38

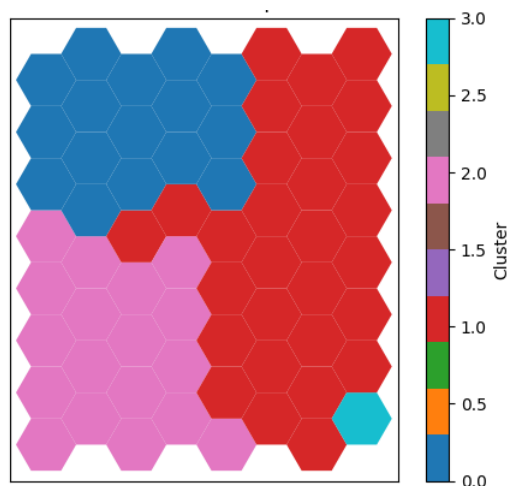
Oxigênio dissolvido (mg/L)	7.26	5.55	5.82	6.54
Sólidos totais dissolvidos (g L⁻¹)	0.47	21.77	0.43	0.79
Salinidade (ppt)	0.21	0.09	0.03	0.07
Amônia (µg L⁻¹)	28.09	44.38	30.85	23.75
Nitrito (µg L⁻¹)	1.67	0.45	3.96	0.69
Nitrato (µg L⁻¹)	7.45	20.94	22.15	9.73
Fósforo total (µg L⁻¹)	56.36	134.86	74.29	214.03
Clorofila total (µg L⁻¹)	16.91	17.94	14.15	15.40

Fonte: Autora (2026)

A análise da tabela evidencia diferenças importantes entre os agrupamentos identificados. O Cluster 1 apresenta valores elevados de pH, turbidez, oxigênio dissolvido e salinidade, além de concentrações relativamente altas de clorofila total, indicando condições que podem favorecer o desenvolvimento de organismos fitoplanctônicos. O Cluster 2 destaca-se pelos maiores valores de temperatura da água, condutividade elétrica, sólidos totais dissolvidos e amônia. Esses resultados sugerem condições associadas à maior presença de íons dissolvidos e de nutrientes no sistema aquático. Por sua vez, o Cluster 3 apresenta os maiores valores de potencial redox, de nitrito e de nitrato, indicando condições de água relativamente mais oxidantes e maior disponibilidade de formas oxidadas de nitrogênio no reservatório. Já o Cluster 4 destaca-se principalmente pelos maiores valores de fósforo total, sugerindo possíveis episódios de maior disponibilidade desse nutriente na água, o que pode estar associado a processos de aporte externo ou ressuspensão de sedimentos.

Por último, apresenta-se a organização espacial dos agrupamentos identificados no SOM do reservatório de Taperoá, conforme ilustrado na **Figura 44**.

Figura 44 – Mapa de clusters do SOM de Taperoá



Fonte: Autora (2026)

99

Observa-se que o Cluster 1 (azul) ocupa grande parte da porção superior esquerda do mapa, indicando predominância desse padrão ambiental no conjunto de dados do reservatório. O Cluster 2 (vermelho) distribui-se principalmente na região central e direita da grade do SOM, formando uma área contínua relativamente extensa. O Cluster 3 (rosa) aparece concentrado na região inferior esquerda do mapa, enquanto o Cluster 4 (verde) ocorre de forma bastante restrita, representado por poucas células na porção inferior direita da grade.

A caracterização das condições ambientais associadas aos agrupamentos identificados pode ser observada nos valores médios das variáveis apresentados na **Tabela 12**.

Tabela 12 – Valores médios das variáveis ambientais para os clusters identificados no SOM individual do reservatório de Taperoá

	Cluster 1	Cluster 2	Cluster 3	Cluster 4
Temperatura da água (°C)	24.77	26.85	24.98	31.21
pH	9.60	8.27	7.55	9.92
Potencial redox	105.75	179.66	208.52	31.00
Condutividade elétrica (mS cm ⁻¹)	3.98	0.91	4.16	152.00
Turbidez (NTU)	685.75	25.21	419.69	112.00
Oxigênio dissolvido (mg/L)	7.26	8.10	6.96	9.02
Sólidos totais dissolvidos (g L ⁻¹)	3.60	0.58	2.66	0.98
Salinidade (ppt)	0.19	0.11	0.22	0.08
Amônia (µg L ⁻¹)	419.32	35.80	957.70	37.05
Nitrito (µg L ⁻¹)	7.53	1.43	30.91	4.74
Nitrato (µg L ⁻¹)	29.00	16.13	79.27	29.71
Fósforo total (µg L ⁻¹)	400.33	39.76	248.98	108.67
Clorofila total (µg L ⁻¹)	234.56	21.04	128.22	79.40

Fonte: Autora (2026)

A análise da tabela evidencia diferenças marcantes entre os agrupamentos. O Cluster 1 apresenta valores extremamente elevados de turbidez (685,75 NTU), sólidos totais dissolvidos (3,60 g L⁻¹), fósforo total (400,33 µg L⁻¹) e clorofila total (234,56 µg L⁻¹), além de altas concentrações de amônia. Esses resultados sugerem episódios de forte enriquecimento nutricional e de alta produtividade biológica no reservatório. O Cluster 3 também apresenta valores elevados de diversas variáveis, destacando-se principalmente amônia (957,70 µg L⁻¹), nitrato (79,27 µg L⁻¹) e nitrito (30,91 µg L⁻¹), indicando grande disponibilidade de compostos nitrogenados dissolvidos na água. Por sua vez, o Cluster 4 apresenta os maiores valores de temperatura da água, pH, condutividade elétrica e oxigênio dissolvido, sugerindo condições ambientais distintas das observadas nos demais agrupamentos, possivelmente associadas a diferenças nas características físico-químicas da água.

100

De forma geral, observa-se que o reservatório de Taperoá apresenta valores extremamente elevados de diversas variáveis limnológicas, superiores aos observados nos demais reservatórios analisados. Esses resultados indicam elevada variabilidade nas condições ambientais do sistema,

A análise dos SOM individuais evidenciou diferenças nos padrões ambientais entre os reservatórios estudados, refletindo variações nas condições físico-químicas da água e na disponibilidade de nutrientes entre os sistemas. A identificação desses agrupamentos permitiu caracterizar as particularidades limnológicas de cada reservatório. Esses resultados reforçam a importância da análise individual de cada reservatório, ao mesmo tempo em que fornecem base para a comparação com o modelo SOM integrado, apresentada na seção seguinte.

4.6 COMPARAÇÃO ENTRE MODELAGEM INTEGRADA E A MODELAGEM INDIVIDUAL

A aplicação do SOM, considerando simultaneamente todas as amostras dos reservatórios, permitiu identificar padrões ambientais gerais no conjunto de dados. Essa abordagem permitiu sintetizar a variabilidade das variáveis limnológicas analisadas e evidenciar gradientes ambientais comuns entre os sistemas estudados. Dessa forma, o modelo integrado fornece uma visão mais ampla da organização dos dados, permitindo identificar tendências gerais na qualidade da água na região analisada.

Por outro lado, a modelagem realizada individualmente para cada reservatório revelou particularidades ambientais específicas de cada sistema, que não necessariamente se evidenciam quando os dados são analisados conjuntamente. Ao considerar cada reservatório separadamente, foi possível observar diferenças mais detalhadas na distribuição das variáveis limnológicas, evidenciando padrões associados às condições locais de cada sistema hídrico.

Entre as principais diferenças observadas entre as duas abordagens, destaca-se a forma como os valores extremos e as variações ambientais mais acentuadas são representados. Na modelagem integrada, a presença de amostras provenientes de diferentes reservatórios tende a suavizar parcialmente essas variações. Já nas modelagens individuais, essas variações tornam-se mais evidentes, permitindo identificar com maior clareza situações específicas de cada reservatório, como os valores elevados de turbidez, condutividade elétrica e de nutrientes.

Dessa forma, observa-se que as duas abordagens apresentam características complementares. Enquanto a modelagem integrada permite identificar padrões ambientais gerais e gradientes comuns entre os reservatórios, as modelagens individuais permitem

101

compreender com maior detalhe a variabilidade interna de cada sistema aquático.

Assim, a utilização conjunta dessas duas estratégias de análise contribui para uma interpretação mais abrangente dos dados, permitindo tanto a identificação de padrões ambientais em escala de bacia hidrográfica, quanto a evidência das características específicas de cada reservatório.

CAPÍTULO V

CONCLUSÕES

Este capítulo apresenta as principais conclusões obtidas a partir das análises realizadas ao longo da pesquisa sobre a qualidade da água nos reservatórios investigados. São sintetizados os resultados das análises estatísticas exploratórias e da modelagem por redes neurais do tipo Self-Organizing Map (SOM), considerando tanto a abordagem integrada quanto os modelos individuais aplicados a cada reservatório. A partir desses resultados, destacam-se as principais contribuições do estudo e são apresentadas recomendações para pesquisas futuras.

5 CONCLUSÕES

Este estudo desenvolveu e aplicou um modelo de rede neural artificial do tipo Self-Organizing Map (SOM) para a análise da qualidade da água em sete reservatórios da bacia hidrográfica do Rio Paraíba, localizados na região semiárida brasileira. A proposta buscou utilizar técnicas de aprendizado de máquina não supervisionado como ferramenta complementar às análises estatísticas convencionais, permitindo identificar estruturas e relações complexas em um conjunto de dados multivariado composto por variáveis físicas, químicas e biológicas de qualidade da água.

As análises estatísticas exploratórias constituíram uma etapa fundamental para a compreensão inicial do comportamento das variáveis de qualidade da água. A estatística descritiva evidenciou elevada variabilidade em diversos parâmetros limnológicos, especialmente os relacionados à concentração de nutrientes e ao material particulado. De forma complementar, a análise de correlação permitiu identificar relações relevantes entre as variáveis, destacando-se a associação entre nitrito e nitrato, relacionada aos processos do ciclo do nitrogênio, bem como a relação entre fósforo total e clorofila total, indicando a influência da disponibilidade de nutrientes sobre a biomassa fitoplanctônica. A Análise de Componentes Principais contribuiu para sintetizar essas relações, evidenciando gradientes ambientais associados principalmente ao enriquecimento nutricional e às transformações biogeoquímicas do nitrogênio.

Embora essas técnicas estatísticas permitam identificar tendências gerais e relações lineares entre as variáveis, apresentam limitações na identificação de estruturas mais complexas presentes em conjuntos de dados ambientais multivariados. Nesse contexto, a aplicação do Self-Organizing Map mostrou-se uma abordagem capaz de explorar relações não lineares entre as variáveis e revelar padrões ambientais presentes no conjunto de dados.

O modelo SOM integrado foi capaz de organizar as amostras em regiões do mapa com diferentes níveis de similaridade, evidenciando a existência de agrupamentos ambientais distintos entre as amostras analisadas. A etapa de clusterização permitiu identificar quatro grupos principais, associados a diferentes condições de qualidade da água. Esses agrupamentos evidenciaram um gradiente ambiental relacionado ao aumento da disponibilidade de nutrientes, da turbidez e da biomassa fitoplanctônica, permitindo distinguir desde condições de menor enriquecimento trófico até situações associadas à maior produtividade biológica e ao maior potencial de eutrofização.

Além da abordagem integrada, a modelagem individual, realizada para cada

reservatório, permitiu identificar as particularidades limnológicas de cada sistema hídrico. Os resultados indicaram que, embora os reservatórios pertençam à mesma bacia hidrográfica, cada sistema apresenta uma organização própria das variáveis ambientais no espaço do SOM, refletindo diferenças nas condições hidrológicas, nos aportes de nutrientes e nos processos biogeoquímicos que influenciam a qualidade da água.

A comparação entre o modelo integrado e os modelos individuais demonstrou que ambas as abordagens têm caráter complementar. Enquanto o modelo integrado permite identificar tendências gerais, os modelos individuais possibilitam uma análise mais detalhada das características ambientais específicas de cada sistema. Dessa forma, a utilização conjunta dessas duas estratégias amplia as possibilidades de interpretação dos dados e contribui para uma compreensão mais abrangente da dinâmica da qualidade da água na bacia hidrográfica.

Outro aspecto relevante refere-se à aplicação do SOM em estudos de qualidade da água. Embora essa técnica seja amplamente utilizada em pesquisas hidrológicas e ambientais, grande parte das aplicações está relacionada à modelagem de vazões ou à análise de séries temporais. Nesse contexto, ainda são relativamente limitados os estudos que utilizam Mapas Auto-Organizáveis especificamente para análise da qualidade da água em reservatórios. Os resultados desta pesquisa demonstram que a técnica apresenta desempenho satisfatório na organização e na interpretação de dados limnológicos, evidenciando seu potencial como ferramenta de apoio à análise multivariada desses dados, contribuindo para o avanço metodológico na área de monitoramento e gestão de sistemas aquáticos.

Dessa forma, conclui-se que o objetivo geral desta pesquisa foi alcançado, uma vez que o modelo SOM desenvolvido foi capaz de representar a estrutura multivariada dos dados de qualidade da água e identificar diferentes condições ambientais nos reservatórios analisados. Assim, a utilização dessa abordagem pode contribuir para ampliar as possibilidades de análise e interpretação de dados provenientes de programas de monitoramento ambiental, fornecendo subsídios importantes para o gerenciamento e a gestão dos recursos hídricos em regiões semiáridas.

5.1 RECOMENDAÇÕES

- Integrar dados hidrológicos e informações de uso e ocupação do solo na bacia hidrográfica para avaliar com maior precisão os fatores que influenciam a dinâmica da qualidade da água;
- Aplicar e comparar técnicas de aprendizado de máquina, como Random Forest ou outros tipos de redes neurais, para avaliar o desempenho de diferentes abordagens na modelagem da qualidade da água;
- Realizar análises temporais mais detalhadas para investigar a influência de eventos hidrológicos extremos, como secas prolongadas e eventos intensos de precipitação, sobre as condições limnológicas dos reservatórios;
- Avaliar a aplicabilidade do modelo desenvolvido em outros reservatórios da região semiárida ou em diferentes bacias hidrográficas, verificando a robustez e a capacidade de generalização da metodologia proposta.

REFERÊNCIAS BIBLIOGRÁFICAS

AGÊNCIA EXECUTIVA DE GESTÃO DAS ÁGUAS DO ESTADO DA PARAÍBA – AESA. Portal institucional. João Pessoa: AESA, 2025. Disponível em: <http://www.aesa.pb.gov.br/aesa-website/>. Acesso em: 9 fev. 2026.

AGÊNCIA EXECUTIVA DE GESTÃO DAS ÁGUAS DO ESTADO DA PARAÍBA – AESA. Rio Paraíba. Comitês de Bacias. João Pessoa: AESA, 2025. Disponível em: <https://www.aesa.pb.gov.br/conselho-comites/comites-de-bacias/rio-paraiba/>. Acesso em: 9 fev. 2026.

AGÊNCIA NACIONAL DE ÁGUAS – ANA. Nota Técnica nº 15/2021/COMAR/SRE: marco regulatório estabelecendo condições de uso dos recursos hídricos no sistema hídrico Sumé, no Estado da Paraíba. Brasília, 2021.

AGÊNCIA NACIONAL DE ÁGUAS E SANEAMENTO BÁSICO – ANA. Reservatórios do semiárido brasileiro: hidrologia, balanço hídrico e operação. Brasília: ANA, 2017.

AGÊNCIA NACIONAL DE ÁGUAS E SANEAMENTO BÁSICO – ANA. Sistema de acompanhamento de reservatórios (SAR): Nordeste e Semiárido – Paraíba. Brasília: ANA, 2026.

AHMADMOAZZAM, Mehdi et al. Assessment of the water quality of Karun River catchment using artificial neural networks, self-organizing maps and K-means algorithm. *Journal of Environmental Accounting and Management*, v. 9, n. 1, p. 43–58, 2021.

ALPAYDIN, Ethem. *Introduction to machine learning*. 4. ed. Cambridge: MIT Press, 2020.

ALVALÁ, Regina Célia dos Santos et al. Drought monitoring in the Brazilian semiarid region. *Anais da Academia Brasileira de Ciências*, v. 91, 2019.

ÁLVARO, Érica Luana Ferreira et al. Phytoplankton and macroinvertebrate diversity and eco-exergy responses to rainfall diverge in semiarid reservoirs. *Ecological Indicators*, v. 147, 2023.

ARAÚJO, Rosivaldo José de. Gestão dos recursos hídricos: uma análise sobre o município de Taperoá-PB. 2023. Trabalho de Conclusão de Curso (Graduação em Administração Pública) – Universidade Federal da Paraíba, João Pessoa, 2023.

AZEVEDO, E. et al. Application of a statistical model for the assessment of environmental quality in neotropical semi-arid reservoirs. *Environmental Monitoring and Assessment*, v. 189, 2017.

BADRUDEEN, A. T. et al. A critical review of statistical, signal processing and machine learning methods for continuous and high-frequency water quality data improvement. *Ecological Informatics*, 2026.

BARBOSA, Pedro Augusto Oliveira. Proposta para avaliação de processos de alocação negociada de água a partir do caso do reservatório Epitácio Pessoa (PB). 2025. Dissertação (Mestrado em Engenharia Civil e Ambiental) – Universidade Federal de Campina Grande, Campina Grande, 2025.

BARROS, Mário Ubirajara Gonçalves et al. Environmental factors associated with toxic cyanobacterial blooms across drinking water reservoirs in a semi-arid region of Brazil. *Harmful Algae*, v. 86, p. 128–137, 2019.

BOURHABI, Bouchra et al. Water quality assessment in Bin El Ouidane dam (Morocco): integrating machine learning and multivariate analysis. *Environmental Nanotechnology, Monitoring & Management*, v. 25, 2026.

BRITO, Yáscara Maia Araújo et al. Multiannual drought overview in the Brazilian semiarid region: hydrological and agricultural impacts and coexistence initiatives. *Desenvolvimento e Meio Ambiente*, v. 58, p. 911–931, 2021.

CAI, N. et al. Distribution and source apportionment of potentially toxic elements in rivers via self-organizing map and positive matrix factorization. *Water*, v. 17, 2025.

CANTALICE, Luciana Ramos. Gestão hídrica do reservatório Sumé-PB: potencialidades e fragilidades. 2010. Dissertação (Mestrado em Recursos Naturais) – Universidade Federal de

Campina Grande, Campina Grande, 2010.

CARNEIRO, Brenda Lara D. S. et al. Predicting anoxia in the wet and dry periods of tropical semiarid reservoirs. *Journal of Environmental Management*, v. 326, 2023.

CHEA, R.; GRENOUILLET, G.; LEK, S. Evidence of water quality degradation in Lower Mekong Basin revealed by self-organizing map. *PLOS ONE*, 2016.

CHEN, Q. et al. Analysis of algal bloom risk with uncertainties in lakes by integrating self-organizing map and fuzzy information theory. *Science of the Total Environment*, v. 482–483, p. 135–145, 2014.

CHOU, J. S.; HO, C. C.; HOANG, H. S. Determining quality of water in reservoir using machine learning. *Ecological Informatics*, v. 44, p. 57–75, 2018.

CONSELHO NACIONAL DE DESENVOLVIMENTO CIENTÍFICO E TECNOLÓGICO – CNPq. Chamada CNPq nº 21/2020 – Programa de Pesquisa Ecológica de Longa Duração (PELD). Brasília: CNPq, 2020.

CORTEZ, Fernanda et al. Effects of prolonged drought on water quality after drying of a semiarid tropical reservoir, Brazil. *Limnologia*, v. 93, 2022.

COSTA, Davi Freire da; BARBOSA, José Etham de Lucena; DANTAS, Ênio Wocylí. Productivity-diversity relationships in reservoir phytoplankton communities in the semi-arid region of northeastern Brazil. *Journal of Arid Environments*, v. 129, p. 64–70, 2016.

CRUZ, Patrícia Silva et al. Dinâmica do fitoplâncton e evolução trófica de reservatório no semiárido brasileiro. *Revista Verde de Agroecologia e Desenvolvimento Sustentável*, v. 13, n. 4, p. 495, 2018.

EID, Mohamed Hamdy et al. Integrating unsupervised machine learning, statistical analysis and Monte Carlo simulation to assess toxic metal contamination. *Results in Engineering*, v. 26, 2025.

FARIAS, André Aires de. Estudo da degradação ambiental no entorno da bacia hidráulica do açude Manoel Marcionilo, Taperoá-PB. 2012. Dissertação (Mestrado em Recursos Naturais) – Universidade Federal de Campina Grande, Campina Grande, 2012.

FARIAS, Camilo Allyson Simões; BEZERRA, Ulisses Alencar; SILVA FILHO, José Adalberto. Runoff-erosion modeling at micro-watershed scale: a comparison of self-organizing maps structures. *Geoenvironmental Disasters*, v. 2, 2015.

FERNANDES, Jussara Mara Lima Queiroz. Problemas socioambientais no entorno do açude Jeremias e do rio Taperoá no município de Desterro – PB. 2022. Monografia – Universidade Federal de Campina Grande, Patos, 2022.

FILHO, Erivaldo dos Santos Medeiros. Análise dos conflitos e da alocação da água do sistema hídrico Sumé (PB). 2024. Trabalho de Conclusão de Curso – Universidade Federal de Campina Grande, Campina Grande, 2024.

FREIRE, Leticia Lacerda et al. Quantitative and qualitative aspects of managed water releases from dryland reservoirs under negotiated allocation in the Brazilian semiarid. *Journal of Hydrology: Regional Studies*, v. 61, 2025.

GAD, Mohamed et al. Comprehensive evaluation and prediction of groundwater quality using multivariate analysis and machine learning models. *Heliyon*, v. 10, 2024.

GARCIA, Hilario López; GONZALEZ, Ivan Machón; NOVAL, Soledad Rocés. Waste treatment monitoring using self-organizing map and condition achievement maps. In: *IFAC Intelligent Components and Instruments for Control Applications*. Aveiro, Portugal: IFAC/Elsevier, 2003.

GOMES, Maurício de Gois; MAIA, Adelená Gonçalves; MEDEIROS, Joana D. arc Freire de. Reservoir operation rule in semiarid areas: the quantity-quality approach. *Journal of Hydrology*, v. 610, 2022.

GOOGLE. Google Colaboratory: ambiente de desenvolvimento em nuvem para Python. Mountain View, 2023. Disponível em: <https://colab.research.google.com>. Acesso em: 9 fev.

2026.

GU, Lei et al. The contribution of internal climate variability to climate change impacts on droughts. *Science of the Total Environment*, v. 684, p. 229–246, 2019.

GUTIÉRREZ, A. P. A. et al. Drought preparedness in Brazil. *Weather and Climate Extremes*, v. 3, p. 95–106, 2014.

HADJISOLOMOU, Ekaterini et al. Assessment of eutrophication-related environmental parameters by integrating statistical techniques and self-organizing maps. *International Journal of Environmental Research and Public Health*, v. 15, 2018.

INTERGOVERNMENTAL PANEL ON CLIMATE CHANGE – IPCC. *Climate Change 2023: Synthesis Report*. Geneva: IPCC, 2023.

JEONG, K. S. et al. Patterning limnological characteristics using a self-organizing map. *Limnology*, v. 9, n. 3, p. 181–191, 2008.

JIN, X. et al. Fluorescence analysis of river DOM spectra using PARAFAC in combination with a self-organizing map to distinguish organic matter sources. *International Journal of Environmental Research*, v. 18, 2024.

KLOMPENBURG, Thomas Van; KASSAHUN, Ayalew; CATAL, Cagatay. Crop yield prediction using machine learning: a systematic literature review. *Computers and Electronics in Agriculture*, v. 177, 2020.

KOHONEN, Teuvo. Essentials of the self-organizing map. *Neural Networks*, v. 37, p. 52–65, 2013.

LACERDA, Mateus Clemente de. *Aprendizado de máquina aplicado à predição da ocorrência de coliformes totais e Escherichia coli em reservatórios de abastecimento de água*. 2024. Dissertação – Universidade Federal de Campina Grande, 2024.

LEMOS, Filipe Carvalho. *Monitoramento remoto e automático de nível no reservatório*

Argemiro de Figueiredo (Acauã). 2019. Trabalho de Conclusão de Curso – Universidade Federal da Paraíba, João Pessoa, 2019.

LIMA NETO, Iran Eduardo et al. Assessment of phosphorus loading dynamics in a tropical reservoir with high seasonal water level changes. *Science of the Total Environment*, v. 815, 2022.

LU, R.; LO, S. L. Diagnosing reservoir water quality using self-organizing maps and fuzzy theory. *Water Research*, v. 36, n. 9, p. 2265–2272, 2002.

MEDEIROS, C. R. et al. How does the storage volume of semi-arid reservoirs change water quality and modulate the diversity of benthic macroinvertebrates? *Science of the Total Environment*, v. 958, 2025.

MELO, Darllene S. et al. Self-organizing maps for evaluation of biogeochemical processes in water quality of subtropical reservoirs. *Water Resources Research*, v. 55, 2019.

MENGISTU, Tarekegn Dejen; CHUNG, Il Moon; CHANG, Sun Woo. Machine learning for water quality prediction and uncertainty assessment. *Physics and Chemistry of the Earth*, 2026.

MIA, Md. Yousuf et al. Analysis of self-organizing maps and explainable artificial intelligence to identify hydrochemical factors that drive drinking water quality in Haor region. *Science of the Total Environment*, v. 904, p. 166927, 2023.

NAUMANN, G. et al. Global changes in drought conditions under different levels of warming. *Geophysical Research Letters*, v. 45, p. 3285–3296, 2018.

PAIVA, Franciely. Macroinvertebrados bentônicos como indicadores de estressores ambientais em reservatórios no semiárido. 2023. Tese – Universidade Estadual da Paraíba, 2023.

PARAÍBA (Estado). Secretaria de Infraestrutura e dos Recursos Hídricos. Memorial descritivo – Barragem de Camalaú (v. 1). Disponível em: <https://paraiba.pb.gov.br/diretas/secretaria-de-infraestrutura-e-dos-recursos-hidricos/arquivos/obras-barragem-camalau/02-20-bar-camalau-tx-001-r0-1-volume-01-memorial-descritivo-1.pdf>. Acesso em: 26 fev. 2026.

PARK, Y. S. et al. Characterizing effects of landscape factors on water quality of reservoirs using a self-organizing map. *Environmental Modelling & Software*, v. 55, p. 37–45, 2014.

PEREIRA NETO, Manoel Cirício. Prospects of the construction of dam in semiarid Brazilian and its implications. *Sociedade & Natureza*, v. 29, 2017.

RAULINO, João B. S.; SILVEIRA, Cleiton S.; LIMA NETO, Iran Eduardo. Assessment of climate change impacts on hydrology and water quality of large semi-arid reservoirs in Brazil. *Hydrological Sciences Journal*, v. 66, n. 8, p. 1321–1336, 2021.

RODRIGUES, Ítalo Sampaio et al. Effects of floating macrophytes on open-water evaporation in a Brazilian tropical reservoir. *Journal of South American Earth Sciences*, v. 164, 2025.

ROCHA JUNIOR, Carlos et al. Influence of land use on spatial distribution of phosphorus in sediment of a tropical semi-arid reservoir. *Science of the Total Environment*, v. 914, 2024.

ROCHA, Maria Aparecida Melo; BARROS, Mário Ubirajara Gonçalves; SOUZA FILHO, Francisco de Assis de; LIMA NETO, Iran Eduardo. Total phosphorus modeling and carrying capacity of the largest reservoir in the Brazilian semiarid region. *Water, Air, & Soil Pollution*, v. 236, n. 525, 2025.

SALAH, Ayari; HAYETTE, Gatfaoui. A meta-analysis of supervised and unsupervised machine learning algorithms. *Expert Systems with Applications*, v. 271, 2025.

SALEM, H. S.; PUDZA, M. Y.; YIHDEGO, Y. Water strategies and water–food nexus. *Sustainable Water Resources Management*, v. 8, 2022.

SANTOS, Ester Milena et al. Modeling water quality in the Brazilian semiarid region using remote sensing. *Environmental Monitoring and Assessment*, v. 197, 2025.

SILVA, Maycon Breno Macena da; RIBEIRO, Márcia Maria Rios. Analysis of water governance through the robustness framework: the case of the Epitácio Pessoa Reservoir, PB. *Engenharia Sanitária e Ambiental*, v. 28, 2023.

SILVA, Maycon Breno Macena da et al. Espacialização do nível de clorofila-a em reservatórios utilizando imagens Sentinel-2: aplicação no reservatório Poções, PB. In: SIMPÓSIO BRASILEIRO DE RECURSOS HÍDRICOS, 25., 2023.

SILVA, M. T. et al. Análise das séries temporais de uso e ocupação da terra em Áreas de Preservação Permanente (APP) da Bacia do Rio Paraíba. RA'EGA – O Espaço Geográfico em Análise, v. 58, p. 20–38, 2023.

SONG, Han et al. The use of the self-organizing map methodology to study dissolved organic matter in water bodies. *Water Science and Technology*, v. 90, 2024.

TUNDISI, J. G.; MATSUMURA-TUNDISI, T.; TUNDISI, J. E. M. Reservoirs and human wellbeing. *Brazilian Journal of Biology*, v. 68, p. 1133–1135, 2008.

VASCONCELOS, Janiele França de. Ocorrência de estados estáveis alternativos em reservatórios do semiárido e sua influência na dinâmica das comunidades aquáticas. 2013. Tese (Doutorado em Ciências Ambientais) – Universidade Estadual de Maringá, Maringá, 2013.

VESANTO, Juha; ALHONIEMI, Esa. Clustering of the self-organizing map. *IEEE Transactions on Neural Networks*, 2000.

WIEGAND, Mário César et al. Trophic state changes of semi-arid reservoirs as a function of hydro-climatic variability. *Journal of Arid Environments*, v. 184, 2021.

ZHAO, X. et al. Analysis of water quality using self-organizing map and fuzzy set-pair analysis. *Water*, v. 17, 2025.