



UNIVERSIDADE FEDERAL DE PERNAMBUCO
CENTRO DE TECNOLOGIA E GEOCIÊNCIAS
DEPARTAMENTO DE ENGENHARIA CIVIL E AMBIENTAL
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA CIVIL

HIDAIANE FAYGA MATIAS CALDAS

**ANÁLISE INTEGRADA DA QUALIDADE DA ÁGUA EM BACIA HIDROGRÁFICA
APLICANDO FERRAMENTAS DA CIÊNCIA DE DADOS**

Recife

2025

HIDAIANE FAYGA MATIAS CALDAS

**ANÁLISE INTEGRADA DA QUALIDADE DA ÁGUA EM BACIA HIDROGRÁFICA
APLICANDO FERRAMENTAS DA CIÊNCIA DE DADOS**

Tese apresentada ao Programa de Pós-Graduação em Engenharia Civil da Universidade Federal de Pernambuco, Centro de Tecnologia e Geociências, como requisito parcial para obtenção do título de Doutora em Engenharia Civil, Área de Tecnologia Ambiental.

Orientadora: Prof^ª. Dr^ª. Maria do Carmo Martins Sobral

Recife

2025

.Catalogação de Publicação na Fonte. UFPE - Biblioteca Central

Caldas, Hidaiane Fayga Matias.

Análise integrada da qualidade da água em Bacia Hidrográfica aplicando ferramentas da Ciência de Dados / Hidaiane Fayga Matias Caldas. - Recife, 2025.
150f.: il.

Tese (Doutorado) - Universidade Federal de Pernambuco, Centro de Tecnologia e Geociências, Programa de Pós-Graduação em Engenharia Civil, 2025.

Orientação: Maria do Carmo Martins Sobral.

Inclui referências e apêndice.

1. Ciência de dados ambientais; 2. Análise de dados; 3. Recursos hídricos; 4. Poluição da água. I. Sobral, Maria do Carmo Martins. II. Título.

UFPE-Biblioteca Central

HIDAIANE FAYGA MATIAS CALDAS

**ANÁLISE INTEGRADA DA QUALIDADE DA ÁGUA EM BACIA
HIDROGRÁFICA APLICANDO FERRAMENTAS DA CIÊNCIA DE DADOS**

Tese apresentada ao Programa de Pós-Graduação em Engenharia Civil da Universidade Federal de Pernambuco, Centro de Tecnologia e Geociências, como requisito parcial para obtenção do título de Doutora em Engenharia Civil, Área de Tecnologia Ambiental.

Aprovada em 08/05/2025

Orientadora: Prof.^a Dr.^a Maria do Carmo Martins Sobral, UFPE

BANCA EXAMINADORA

participação por videoconferência
Prof. Dr. Wanderli Rogério Moreira Leite (examinador interno)
Universidade Federal de Pernambuco

participação por videoconferência
Prof.^a Dr.^a Maristela Casé Costa Cunha (examinadora externa)
Universidade do Estado da Bahia

participação por videoconferência
Dr. José Luís Said Cometti (examinador externo)
Agência Estadual de Meio Ambiente e Recursos Hídricos

participação por videoconferência
Dr.^a Éricka Alves Tavares Marques (examinadora externa)
Universidade Federal de Pernambuco

participação por videoconferência
Prof.^a Dr.^a Simone Machado Santos (examinadora externa)
Universidade Federal de Pernambuco

À minha família e meus amigos, que foram meu apoio durante essa jornada, e contribuem para o meu crescimento pessoal a cada dia.

AGRADECIMENTOS

Agradeço, em primeiro lugar, aos meus pais, Carlos (*in memoriam*) e Hidelvânia, por toda a força, apoio e incentivo na busca pelos meus objetivos. À minha irmã Hithalla e ao meu sobrinho Carlos Neto, que me ensinam, a cada dia, o valor de compartilhar momentos e sentimentos. À minha avó Emília e aos meus tios Marcílio e Matsuzuk, pela referência de família e afeto que sempre me transmitiram. Aos demais familiares, por todo o estímulo nos momentos de conquista e nas adversidades. Ao meu namorado Wladimir, pelo apoio constante e pela paciência inestimável.

Minha gratidão especial à minha orientadora, Professora Doutora Maria do Carmo Sobral, por sua confiança, dedicação e pela oportunidade de aprender e crescer junto à sua equipe. Sua orientação foi essencial para a realização deste trabalho, e é para mim motivo de profunda admiração.

Agradeço a todos os colegas do Grupo de Gestão Ambiental que estiveram ao meu lado ao longo desta jornada: Ariane Cardoso, Anthony Epifânio, Janaina Assis, Gérsica Nogueira, Maiara Melo, Sophia Telino, Érika Marques, Athos Farias e Tatyane Guimarães. Cada um contribuiu de maneira única e significativa para esta trajetória.

À Universidade de Évora, em Portugal, e a todos que tive o prazer de conhecer durante as duas oportunidades em que estive nesse país, meu muito obrigada pela calorosa recepção e apoio. Em especial, à Professora Doutora Manuela Morais, pelos ensinamentos compartilhados e pela inspiração acadêmica. Às colegas Ana Pedro, Ângela Guerra e Amália Oliveira, do Laboratório da Água; e ao Domingos Romão, da Divisão de Internacionalização e Mobilidade, meu sincero agradecimento.

Aos bons amigos que sempre estiveram presentes com carinho, apoio e palavras de encorajamento: Agnes Pinheiro, Vanine Farias, Ellen Marie, Samara Fernanda, Rochelle Vasconcelos, Joelma Ferreira, Daliana Teles, Sandra Farinha, Roberto Chillihua, Joseph Greene e Peder Gedda.

Aos membros da banca examinadora: Professor Doutor Wanderli Rogério, Doutor Luís Cometti, Professora Doutora Maristela Casé, Doutora Érika Marques e Professora Doutora Simone Machado, pelas contribuições valiosas para o aprimoramento desta pesquisa.

À Universidade Federal de Pernambuco (UFPE), aos docentes e às secretárias do Programa de Pós-Graduação em Engenharia Civil (PPGEC), pela colaboração contínua ao longo do desenvolvimento deste trabalho.

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001. Agradeço, ainda, à CAPES pela concessão da bolsa de intercâmbio CAPES-PrInt (Processo 88887.935329/2024-00), sendo ambas as bolsas essenciais para o desenvolvimento desta pesquisa.

Por fim, expresso minha mais profunda gratidão a todas as pessoas e instituições que, direta ou indiretamente, contribuíram para o desenvolvimento desta pesquisa.

RESUMO

À medida que os conjuntos de dados relacionados aos recursos hídricos crescem em tamanho e complexidade, a necessidade de respostas rápidas e personalizadas também aumenta. Em um contexto atual, pautado pelos desafios da Agenda 2030 e do Objetivo de Desenvolvimento Sustentável 6 – Água Potável e Saneamento, a gestão da água exige abordagens holísticas para lidar com questões cada vez mais diversas. Este estudo tem como objetivo principal propor uma metodologia para análise integrada da qualidade da água aplicando ferramentas da Ciência de Dados, tendo como estudo de caso a bacia hidrográfica do rio Ipojuca, em Pernambuco. A proposta metodológica inicia com a avaliação espaço-temporal da qualidade da água, seguida pela análise da sua relação com o uso do solo, o desenvolvimento de um modelo conceitual e a elaboração de um painel de visualização dos dados. Como resultado, observou-se que os parâmetros com maior número de violações à legislação foram oxigênio dissolvido, DBO, fósforo total e coliformes termotolerantes, indicando forte influência da poluição. As estações de amostragem localizadas em Caruaru, Gravatá e Ipojuca apresentaram as piores condições de qualidade da água, e as variáveis associadas à poluição antrópica e mineral exerceram maior influência. Foram identificadas mudanças significativas nas séries temporais e tendências de piora em parâmetros como DBO, oxigênio dissolvido e coliformes termotolerantes, refletindo possíveis impactos do crescimento da área urbanizada e de fatores ambientais. A análise do uso e ocupação do solo ao redor das estações de amostragem revelou transformações importantes, com destaque para a expansão de áreas urbanas e formação savânica. Verificou-se que maiores áreas com cobertura florestal estão associadas a melhor qualidade da água, enquanto áreas agropecuárias correlacionaram-se positivamente com o fósforo total, indicador de poluição. A análise de redundância evidenciou maior influência do uso do solo no período seco, explicando 65,4% da variância total da qualidade da água. O modelo conceitual proposto é formado por tabelas relacionais que estruturam informações de diversas fontes, facilitando e simplificando o desenvolvimento do painel interativo. Este foi elaborado com visualizações dinâmicas e georreferenciadas, permitindo a exploração detalhada dos dados de monitoramento, infraestrutura e outorgas na bacia. A abordagem integrada proposta demonstrou-se eficaz e replicável, servindo como referência para o aprimoramento da gestão da qualidade da água em bacias hidrográficas.

Palavras-chave: Ciência de dados ambientais. Análise de dados. Recursos hídricos. Poluição da água.

ABSTRACT

As datasets related to water resources grow in size and complexity, the need for fast and customized responses also increases. In the current global scenario, framed by the challenges of the 2030 Agenda and Sustainable Development Goal (SDG) 6 – Clean Water and Sanitation, effective water management demands holistic and integrated approaches to address increasingly complex and multifaceted issues. This study primarily aims to propose a methodology for the integrated analysis of water quality by applying Data Science tools, using the Ipojuca River watershed in Pernambuco as a case study. The proposed methodology begins with the spatiotemporal assessment of water quality, followed by the analysis of its relationship with land use, the development of a conceptual model, and the creation of a data visualization dashboard. As a result, it was observed that the parameters with the highest number of violations of legal standards were dissolved oxygen, BOD (Biochemical Oxygen Demand), total phosphorus, and thermotolerant coliforms, indicating strong pollution influence. The sampling stations located in Caruaru, Gravatá, and Ipojuca presented poor water quality conditions, with variables associated with anthropogenic and mineral pollution exerting the greatest influence. Significant changes were identified in the time series, along with worsening trends in parameters such as BOD, dissolved oxygen, and thermotolerant coliforms, reflecting potential impacts of urban expansion and environmental factors. The analysis of land use and land cover around the sampling stations revealed important transformations, particularly the expansion of urban areas and savanna formation. It was found that larger forested areas are associated with better water quality, while agricultural and livestock areas correlated positively with total phosphorus, a pollution indicator. Redundancy analysis showed greater influence of land use during the dry season, explaining 65.4% of the total variance in water quality. The proposed conceptual model consists of relational tables that structure data from multiple sources, facilitating and simplifying the development of an interactive dashboard. This dashboard was designed with dynamic and georeferenced visualizations, enabling detailed exploration of monitoring data, infrastructure, and water use permits in the watershed. The proposed integrated approach was proved effective and replicable, serving as a reference for improving water quality management in river basins.

Keywords: Environmental data science. Data analysis. Water resources. Water pollution.

LISTA DE FIGURAS

Figura 1 -	Objetivos do Desenvolvimento Sustentável firmados na Agenda 2030	25
Figura 2 -	Os cinco instrumentos de gestão dos recursos hídricos previstos na PNRH....	26
Figura 3 -	Matriz institucional dos integrantes do SINGREH.....	28
Figura 4 -	Unidades de Planejamento Hídrico definidas no PERH/PE - 2022.....	30
Figura 5 -	Comportamento da curva correspondente às metas progressivas de qualidade da água ao longo do tempo.....	34
Figura 6 -	Diagrama de Venn da Ciência de Dados.....	37
Figura 7 -	Elementos que contribuem para a Ciência de Dados.....	38
Figura 8 -	Arquitetura do processo de Ciência de Dados.....	40
Figura 9 -	Esquema Estrela (esquerda) vs. Esquema Floco de Neve (direita).....	42
Figura 10 -	Fluxograma de execução da pesquisa.....	46
Figura 11 -	Localização da Bacia Hidrográfica do Rio Ipojuca.....	47
Figura 12 -	Distribuição espacial das estações de amostragem da bacia do rio Ipojuca selecionadas para avaliação.....	59
Figura 13 -	Porcentagem de dados faltantes por parâmetro no banco de dados bruto.....	60
Figura 14 -	Etapas da elaboração do painel de visualização da qualidade da água.....	66
Figura 15 -	Variação dos parâmetros de qualidade da água nas estações de monitoramento da BHRI entre 2005 e 2023.....	71
Figura 16 -	Percentuais de violação da Resolução CONAMA nº 357/2005 para águas doces Classe 2 nas estações de amostragem analisadas.....	77
Figura 17 -	Precipitação média mensal observada nas unidades de análise da bacia do rio Ipojuca no período de 2005 a 2023.....	79
Figura 18 -	Precipitação média mensal observada para cada ano entre 2005 e 2023, nas unidades de análise da bacia do rio Ipojuca.....	80
Figura 19 -	Vazão média mensal observada nas unidades de análise da bacia do rio Ipojuca no período de 2005 a 2023.....	81
Figura 20 -	Vazão média mensal observada para cada ano entre 2005 e 2023, nas unidades de análise da bacia do rio Ipojuca.....	83
Figura 21 -	Diferenças entre a precipitação (a) e a vazão (b) médias nas unidades de análise da BHRI.....	84

Figura 22 - Variação dos parâmetros de qualidade da água nas estações de monitoramento da bacia do rio Ipojuca entre 2005 e 2023, nos períodos seco e chuvoso.....	85
Figura 23 - Matriz de gráficos de dispersão dos dados do período chuvoso, com gráficos de dispersão bivariados abaixo da diagonal, histogramas na diagonal e correlação de Spearman acima da diagonal.....	89
Figura 24 - Matriz de gráficos de dispersão dos dados do período seco, com gráficos de dispersão bivariados abaixo da diagonal, histogramas na diagonal e correlação de Spearman acima da diagonal.....	90
Figura 25 - Biplot da Análise de Componentes Principais para os dois primeiros eixos nos períodos de 2005-2019 e 2020-2023.....	93
Figura 26 - Parâmetros com pontos de mudança significativos nas estações de amostragem na BHRI.....	98
Figura 27 - Parâmetros que apresentaram tendências significativas na BHRI.....	101
Figura 28 - Evolução do uso do solo na BHRI nos anos de 2005 e 2019.....	103
Figura 29 - Composição do uso do solo em raio de 500 metros ao redor das estações de amostragem de qualidade da água da BHRI.....	105
Figura 30 - Agrupamentos de acordo com (a) uso do solo e (b) indicadores de qualidade da água no período chuvoso e (c) no período seco.....	107
Figura 31 - Gráficos de ordenação (triplots) mostrando a relação entre as classes de uso do solo (setas pretas), os parâmetros de qualidade da água (setas cinzas) e as amostras individuais nos períodos chuvoso e seco, para o modelo completo.....	110
Figura 32 - Modelo conceitual de dados usado para elaboração do painel de visualização da qualidade da água da bacia do rio Ipojuca.....	113
Figura 33 - Aba de apresentação do painel de visualização da qualidade da água desenvolvido para a bacia do rio Ipojuca.....	116
Figura 34 - Aba de visão geral do monitoramento da qualidade da água da bacia do rio Ipojuca.....	117
Figura 35 - Interatividade na aba sobre a visão geral, mostrando informações populacionais e de saneamento.....	118
Figura 36 - Linha do tempo da Gestão de Recursos Hídricos na Bacia do Ipojuca.....	119

Figura 37 -	Aba de análise espacial da qualidade da água ao longo da bacia do rio Ipojuca.....	121
Figura 38 -	Aba de análise espacial da qualidade da água ao longo da bacia do rio Ipojuca com detalhamento da porcentagem de violação aos limites recomendados por estação de amostragem.....	121
Figura 39 -	Aba de análise espacial da qualidade da água ao longo da bacia do rio Ipojuca com detalhamento sazonal.....	122
Figura 40 -	Aba de análise temporal da qualidade da água da bacia do rio Ipojuca com detalhamento em forma de gráfico.....	123
Figura 41 -	Aba de análise temporal da qualidade da água da bacia do rio Ipojuca com detalhamento em forma de tabela.....	123
Figura 42 -	Aba de análise do uso do solo com detalhamento para a bacia do rio Ipojuca	124
Figura 43 -	Aba de análise do uso do solo com detalhamento para os municípios inseridos na bacia.....	125
Figura 44 -	Aba de análise do uso do solo com detalhamento para o raio de 500 metros ao redor das estações de amostragem.....	126
Figura 45 -	Aba de análise das outorgas concedidas na bacia do rio Ipojuca.....	127
Figura 46 -	Aba de análise das outorgas concedidas na bacia do rio Ipojuca, com destaque para a topologia lançamento de efluentes.....	128
Figura 47 -	Variação dos parâmetros de qualidade da água nas estações de amostragem da BHRI ao longo do tempo.....	147
Figura 48 -	Continuação.....	148
Figura 49 -	Continuação.....	149
Figura 50 -	Continuação.....	150

LISTA DE QUADROS

Quadro 1 -	Descrição dos cinco instrumentos de gestão dos recursos hídricos previstos na PNRH.....	28
Quadro 2 -	Caracterização trófica dos corpos hídricos.....	36
Quadro 3 -	Elementos que contribuem para a Ciência de Dados e suas descrições.....	40
Quadro 4 -	Municípios que integram a bacia hidrográfica do rio Ipojuca.....	48
Quadro 5 -	Principais reservatórios da BHRI.....	50
Quadro 6 -	Estações de amostragem da rede de monitoramento da bacia do rio Ipojuca.....	52
Quadro 7 -	Empreendimentos concluídos do PSA Ipojuca e suas respectivas fontes de recursos.....	55
Quadro 8 -	Frequência das coletas para monitoramento da qualidade da água da BHRI entre os anos de 2005 e 2023.....	57
Quadro 9 -	Postos pluviométricos levantados na BHRI.....	58
Quadro 10 -	Postos fluviométricos levantados na BHRI.....	58
Quadro 11 -	Frequência de monitoramento das estações de amostragem da bacia do rio Ipojuca.....	59
Quadro 12 -	Estações de amostragem da rede de monitoramento da BHRI selecionadas para avaliação.....	59

LISTA DE TABELAS

Tabela 1 -	Indicadores de saneamento por município para o ano de referência 2022 (NI - não informado)	51
Tabela 2 -	Parâmetros de qualidade da água utilizados na pesquisa que possuem valores máximos permitidos de acordo com a Resolução CONAMA nº 357/2005.....	61
Tabela 3 -	Resultados da estatística W e valores de p do teste de Shapiro-Wilk.....	70
Tabela 4 -	Resumo dos parâmetros de qualidade da água analisados na BHRI de 2005 a 2023.....	71
Tabela 5 -	Matriz de componentes principais derivada da ACP com dados sobre parâmetros de qualidade da água na bacia do rio Ipojuca, entre 2005-2019...	95
Tabela 6 -	Matriz de componentes principais derivada da ACP com dados sobre parâmetros de qualidade da água na bacia do rio Ipojuca, entre 2020-2023...	96
Tabela 7 -	Resultados do teste de Pettitt para os parâmetros de qualidade da água na BHRI.....	98
Tabela 8 -	Resultados do teste de Mann-Kendall para séries temporais segmentadas em antes e depois dos pontos de mudança identificados.....	100
Tabela 9 -	Resultados do teste de Mann-Kendall para séries temporais de parâmetros médios anuais nas estações de amostragem da BHRI.....	101
Tabela 10 -	Área e proporção de área das classes de uso do solo na BHRI em 2005 e 2019.....	103
Tabela 11 -	Resultados da análise de correlação entre os padrões de uso do solo e os parâmetros de qualidade da água na BHRI.....	106
Tabela 12 -	Resultados da análise de redundância para a porcentagem da variância geral da qualidade da água explicada pelo uso do solo na BHRI.....	111

LISTA DE SIGLAS E ABREVIATURAS

ACP	Análise de Componentes Principais
AESA	Agência Executiva de Gestão das Águas da Paraíba
AGNU	Assembleia Geral das Nações Unidas
ANA	Agência Nacional de Águas e Saneamento Básico
APAC	Agência Pernambucana de Águas e Clima
APP	Área de Preservação Permanente
BA	<i>Business Analytics</i>
BHRI	Bacia Hidrográfica do Rio Ipojuca
BI	<i>Business Intelligence</i>
BID	Banco Interamericano de Desenvolvimento
CBH	Comitê de Bacia Hidrográfica
CERH	Conselho Estadual de Recursos Hídricos
CNRH	Conselho Nacional de Recursos Hídricos
COMPESA	Companhia Pernambucana de Saneamento
CONAMA	Conselho Nacional de Meio Ambiente
CPRH	Agência Estadual de Meio Ambiente de Pernambuco
DBO	Demanda Bioquímica de Oxigênio
EEE	Estação Elevatória de Esgoto
ETA	Estação de Tratamento de Água
ETE	Estação de Tratamento de Esgoto
ETL	<i>Extraction, Transformation and Load</i>
GIRH	Gestão Integrada de Recursos Hídricos
IA	Inteligência Artificial
IBGE	Instituto Brasileiro de Geografia e Estatística
IET	Índice de Estado Trófico
IQA	Índice de Qualidade da Água
MIDR	Ministério da Integração e do Desenvolvimento Regional
OD	Oxigênio Dissolvido
ODM	Objetivos de Desenvolvimento do Milênio
ODS	Objetivo de Desenvolvimento Sustentável
OLAP	<i>Online Analytical Processing</i>

ONU	Organização das Nações Unidas
PDRH	Plano Diretor de Recursos Hídricos
PERH	Política Estadual de Recursos Hídricos
PERH/PE	Plano Estadual de Recursos Hídricos de Pernambuco
PHA	Plano Hidroambiental
PNRH	Política Nacional de Recursos Hídricos
PSA IPOJUCA	Programa de Saneamento Ambiental da bacia hidrográfica do rio Ipojuca
RDA	Análise de Redundância
RHN	Rede Hidrometeorológica Nacional
RMR	Região Metropolitana do Recife
RNQA	Rede Nacional de Monitoramento da Qualidade de Água
SAD	Sistema de Apoio à Decisão
SEINFRA	Secretaria de Infraestrutura e Recursos Hídricos
SEMOBI	Secretaria de Mobilidade e Infraestrutura
SES	Sistema de Esgotamento Sanitário
SGBD	Sistemas de Gerenciamento de Banco de Dados
SIGRH/PE	Sistema Integrado de Gerenciamento de Recursos Hídricos de Pernambuco
SINGREH	Sistema Nacional de Gerenciamento dos Recursos Hídricos
SNIRH	Sistema Nacional de Informações sobre Recursos Hídricos
SNIS	Sistema Nacional de Informações sobre Saneamento
SNSH	Secretaria Nacional de Segurança Hídrica
UA	Unidade de Análise
ULAB	Unidade de Análises Laboratoriais

SUMÁRIO

1	INTRODUÇÃO	19
1.1	APRESENTAÇÃO DA AUTORA	19
1.2	JUSTIFICATIVA	20
1.3	HIPÓTESE	21
1.4	OBJETIVOS	22
1.4.1	Objetivo geral	22
1.4.2	Objetivos específicos	22
1.5	ESTRUTURA DA TESE	22
2	REFERENCIAL TEÓRICO	24
2.1	GESTÃO DOS RECURSOS HÍDRICOS.....	24
2.1.1	Gestão de Recursos Hídricos no Brasil	25
2.1.2	Gestão de Recursos Hídricos em Pernambuco	28
2.1.3	Avaliação da qualidade da água	31
2.2	CIÊNCIA DE DADOS	36
2.2.1	Visualização de dados usando ferramentas de Business Intelligence	40
2.2.2	Ciência de Dados na área das Ciências Ambientais e de Recursos Hídricos	43
3	METODOLOGIA	46
3.1	CARACTERIZAÇÃO DA ÁREA DE ESTUDO	47
3.2	PROCEDIMENTOS METODOLÓGICOS	55
3.2.1	Avaliação da evolução espaço-temporal da qualidade da água	55
3.2.2	Relação entre o uso e cobertura do solo e a qualidade da água	64
3.2.3	Desenvolvimento de painel de visualização da qualidade da água	66
4	RESULTADOS E DISCUSSÃO	69
4.1	EVOLUÇÃO ESPAÇO-TEMPORAL DA QUALIDADE DA ÁGUA NA BACIA DO RIO IPOJUCA.....	69
4.1.1	Padrões espaciais de qualidade da água	69
4.1.2	Influência dos períodos seco e chuvoso	78
4.1.3	Parâmetros que exercem maior influência na qualidade da água	92
4.1.4	Análise de tendências	96
4.2	ANÁLISE DA RELAÇÃO ENTRE O USO E COBERTURA DO SOLO E A QUALIDADE DA ÁGUA NA BACIA DO RIO IPOJUCA.....	102
4.3	PAINEL DE VISUALIZAÇÃO DA QUALIDADE DA ÁGUA	112
5	CONCLUSÕES E RECOMENDAÇÕES	130
5.1	CONCLUSÕES	130
5.2	RECOMENDAÇÕES	131

REFERÊNCIAS	134
APÊNDICE A - GRÁFICOS DA SÉRIE TEMPORAL DOS PARÂMETROS DE QUALIDADE DA ÁGUA	147

1 INTRODUÇÃO

Nesta seção, é feita uma breve apresentação da autora, seguida pela justificativa da temática escolhida para o desenvolvimento desta pesquisa, juntamente com a apresentação da hipótese e dos objetivos do trabalho.

1.1 APRESENTAÇÃO DA AUTORA

A autora possui curso técnico em Desenvolvimento de *Software* pelo Instituto Federal do Piauí (IFPI), finalizado em 2011. Concluiu a graduação em Engenharia Civil na Universidade Federal de Campina Grande (UFCG) em 2016, e obteve Especialização em Gerenciamento de Obras pelo UNIFIP Centro Universitário em 2020. Em 2021, concluiu o mestrado em Engenharia Civil pela Universidade Federal de Pernambuco (UFPE).

Seu interesse pela área de Recursos Hídricos e Saneamento surgiu durante a graduação, quando conduziu um estudo sobre assoreamento em açudes urbanos, tendo como estudo de caso o Açude Velho em Campina Grande/PB.

Durante o mestrado, foi bolsista da Fundação de Amparo à Ciência e Tecnologia de Pernambuco (FACEPE) e começou a contribuir para o Programa de Monitoramento de Qualidade de Água e Limnologia (PBA-22) do Projeto de Integração do Rio São Francisco com Bacias Hidrográficas do Nordeste Setentrional (PISF), financiado pelo Ministério da Integração e do Desenvolvimento Regional. Sua pesquisa de mestrado envolveu uma análise da evolução espaço-temporal da qualidade da água no Eixo Leste do PISF.

Iniciou o doutorado em Engenharia Civil na UFPE em 2021, participando do Projeto MonitoraQuali: Avaliação do processo de monitoramento da qualidade da água de bacias hidrográficas de Pernambuco, realizado em parceria entre a FACEPE, UFPE e a Agência Estadual de Meio Ambiente de Pernambuco (CPRH), finalizado em 2023. No mesmo ano, foi contemplada com uma bolsa do programa de mobilidade de estudantes do projeto ERASMUS AMIGO, passando três meses desenvolvendo pesquisa no Laboratório da Água da Universidade de Évora – Portugal.

Ao longo de sua trajetória acadêmica, buscou conhecimento interdisciplinar, participando de cursos de curta duração em Sensoriamento Remoto usando o *software* QGIS, Linguagem R, Linguagem *Python*, *Machine Learning* e *Microsoft Power BI*, a fim de agregar valor e promover uma interação mais ampla entre o mercado de trabalho e a pesquisa científica.

1.2 JUSTIFICATIVA

Os recursos hídricos desempenham um papel fundamental nas esferas econômica, social e biológica. Diante dessa relevância, é imperativo que sua exploração seja cuidadosamente planejada, gerenciada e monitorada, com o objetivo de atenuar os conflitos decorrentes dos diversos usos da água. Esses conflitos são agravados, principalmente, por eventos climáticos extremos (SOBRAL et al., 2007). A qualidade da água assume papel central em escala global, impactando diretamente a saúde humana, a biodiversidade e a sustentabilidade ambiental. Em um cenário contemporâneo caracterizado por desafios ambientais cada vez mais complexos, a adoção de abordagens inovadoras torna-se fundamental para enfrentar tais questões.

O sexto Objetivo de Desenvolvimento Sustentável (ODS 6) da Agenda 2030 das Nações Unidas para o Desenvolvimento Sustentável tem como meta assegurar a disponibilidade e o manejo sustentável da água e do saneamento para todos (ONU, 2015). Nesse contexto, a Gestão Integrada de Recursos Hídricos (GIRH) emerge como uma importante abordagem que busca equilibrar as exigências hídricas da sociedade, preservando, ao mesmo tempo, a sustentabilidade dos ecossistemas (UNEP, 2021).

Para que seja possível a avaliação do estado de poluição, degradação ou conservação de rios, lagos, represas, estuários ou águas costeiras, é necessário o monitoramento das suas propriedades físicas, químicas e biológicas. O monitoramento da qualidade da água trata da coleta de informações regulares e a formação de um banco de dados fundamental para futuras ações, proporcionando o direcionamento dos recursos para áreas prioritárias no controle da poluição hídrica (TUNDISI J.; TUNDISI T., 2008).

Programas de monitoramento de qualidade da água são um importante pilar da GIRH, gerando uma enorme quantidade de dados que muitas vezes são difíceis de serem interpretados (CHAPMAN; SULLIVAN, 2022). Com isso, uma série de abordagens vêm sendo amplamente utilizadas para um melhor entendimento da qualidade da água, retratando variações temporais, espaciais e de fatores antropogênicos (VEGA et al., 1998; PALMA et al., 2010; JARDIM et al., 2014; CALAZANS et al., 2018).

Ciência de Dados é um campo multidisciplinar que combina análise de dados com métodos de processamento e experiência no domínio, transformando dados em conhecimento relevante para uma tomada de decisão informada (GILBERT et al., 2018). Para o caso da qualidade da água, a aplicação da Ciência de Dados na forma de técnicas estatísticas avançadas, técnicas de visualização e Aprendizado de Máquina, oferece uma abordagem inovadora para lidar com a gestão desse recurso. Ao integrar essas ferramentas, é possível analisar grandes

volumes de dados provenientes de diversas fontes, como sensores remotos, estações de monitoramento e dados meteorológicos. Essa perspectiva permite identificar padrões complexos, tendências e correlações que podem passar despercebidos por métodos tradicionais.

A motivação para a pesquisa proposta é estimulada pelo potencial transformador da Ciência de Dados na gestão da qualidade da água. A capacidade de prever eventos futuros, identificar fontes de contaminação de maneira rápida e precisa, otimizar estratégias de monitoramento, além da crescente disponibilidade de dados e evolução das tecnologias são aspectos que instigam o uso dessa abordagem, sempre com o objetivo de melhorar a capacidade de resposta a eventos críticos e aprimorar a gestão sustentável dos recursos hídricos.

A bacia hidrográfica do rio Ipojuca foi escolhida como estudo de caso neste trabalho, dada a sua importância no contexto pernambucano e as diversas intervenções em curso pelo governo do Estado. Essas intervenções visam aprimorar a qualidade da água através da implementação de sistemas de abastecimento e esgotamento sanitário, além da criação de parques horizontais ao longo do rio, entre outras iniciativas. Vale ressaltar que a bacia possui um plano hidroambiental que define as medidas a serem implementadas para a proteção dos ecossistemas locais.

Outro ponto relevante é a recente regulamentação do novo marco regulatório de saneamento sancionado pela Lei Federal 14.026, de 15 de julho de 2020, que define metas de universalização a serem alcançadas até o final de 2033, ao mesmo tempo em que introduz mecanismos para incentivar a implementação de novos sistemas de saneamento ambiental (BRASIL, 2020). Dessa forma, a perspectiva renovada da sustentabilidade regional demanda uma abordagem inovadora na gestão sustentável dos recursos hídricos, com a aplicação de metodologias e instrumentos tecnológicos específicos para cada contexto.

Consciente dos desafios apresentados na gestão dos recursos hídricos, representados no estudo de caso da bacia hidrográfica do rio Ipojuca, essa pesquisa propõe uma metodologia para análise integrada da qualidade da água na bacia aplicando ferramentas da Ciência de Dados. Dessa forma, busca-se contribuir para uma atuação eficaz do Estado, por meio da identificação de ações prioritárias e políticas públicas a serem implementadas a curto, médio e longo prazo.

1.3 HIPÓTESE

A aplicação de ferramentas da Ciência de Dados contribui para o desenvolvimento de uma metodologia para análise integrada da qualidade da água em bacias hidrográficas.

1.4 OBJETIVOS

1.4.1 Objetivo geral

Propor uma metodologia para análise integrada da qualidade da água, aplicando ferramentas da Ciência de Dados, tendo como estudo de caso a bacia hidrográfica do rio Ipojuca, em Pernambuco.

1.4.2 Objetivos específicos

- Avaliar a evolução espaço-temporal e tendências de alteração da qualidade da água;
- Analisar as mudanças de uso e ocupação do solo e suas relações com a qualidade da água na bacia;
- Elaborar um modelo conceitual de dados representativo com informações de suporte à gestão da qualidade da água;
- Desenvolver um painel de integração e visualização dos dados para auxiliar as análises e o gerenciamento da qualidade da água;
- Propor recomendações para a gestão da qualidade de água do rio Ipojuca.

1.5 ESTRUTURA DA TESE

Este documento está organizado em cinco capítulos. O primeiro capítulo serve como introdução ao tema, apresentando a autora, justificando a problemática abordada, formulando uma hipótese, delineando os objetivos geral e específicos, e esboçando a estrutura do documento.

No segundo capítulo, é abordado o referencial teórico que fundamenta a pesquisa. Isso inclui uma explanação geral da Gestão dos Recursos Hídricos, com um foco específico no Brasil e no estado de Pernambuco. Além disso, são explorados conceitos e considerações sobre a qualidade da água. Também é contextualizada a temática da Ciência de Dados, discutindo definições relacionadas à visualização de dados por meio de ferramentas de *Business Intelligence* e a aplicação da Ciência de Dados no campo das Ciências Ambientais e dos Recursos Hídricos.

No terceiro capítulo, é apresentada a metodologia adotada, descrevendo a área de estudo e os procedimentos metodológicos empregados para alcançar cada objetivo específico proposto.

Além disso, são explicitados detalhes sobre o levantamento do banco de dados utilizado na pesquisa.

No capítulo quatro, são apresentados os resultados e a discussão, os quais abordam a avaliação espaço-temporal da qualidade da água na bacia, a análise da relação entre o uso do solo e a qualidade da água, além do painel de visualização da qualidade da água elaborado na pesquisa. O quinto capítulo delinea as conclusões e recomendações para trabalhos futuros.

2 REFERENCIAL TEÓRICO

Este tópico apresenta um levantamento teórico sobre os temas relacionados a esta pesquisa, obtido por meio de estudo bibliográfico e documental. Inicia-se com uma abordagem sobre a gestão dos recursos hídricos, incluindo aspectos sobre a qualidade da água em pontos relevantes para o entendimento deste estudo, e finaliza com uma apresentação sobre o tema Ciência de Dados.

2.1 GESTÃO DOS RECURSOS HÍDRICOS

A água é um recurso finito, e uma maior demanda resulta em estresse hídrico, que é uma função da disponibilidade, demanda e qualidade da água. Esses estresses são causados principalmente pelo aumento da população humana e do uso *per capita* de água, bem como a concentração crescente de pessoas em áreas urbanas. Somado a isso, ainda temos os efeitos das mudanças climáticas, demanda por mais água para irrigação visando o aumento da produção de alimentos e requisitos ambientais (fluxos ambientais) para a biodiversidade e proteção ambiental (MILAN, 2017).

Dessa forma, enquanto as mudanças climáticas afetam os ecossistemas aquáticos e a disponibilidade de recursos hídricos, fatores socioeconômicos aumentam a demanda por água e deterioram os corpos d'água. Para lidar com a problemática, se faz necessário desenvolver métodos eficientes de planejamento, coordenação e fiscalização dos usos da água, visando otimizar esse recurso. Isso remete à ideia de gestão de recursos hídricos, cujo objetivo é preservar e recuperar a quantidade e qualidade das águas, garantindo a segurança hídrica para as gerações futuras (ANA, 2017).

A Assembleia Geral das Nações Unidas (AGNU) adotou formalmente os Objetivos de Desenvolvimento Sustentável (ODS) em 5 de setembro de 2015, para estabelecer o caminho para um esforço contínuo e uniforme de desenvolvimento em nível global, substituindo os Objetivos de Desenvolvimento do Milênio (ODM). Os ODS consistem em um conjunto de 17 objetivos (Figura 1) e 169 metas destinados a moldar a política social, econômica e ambiental global até 2030 (ONU, 2015).

O ODS 6 (Água potável e saneamento) refere-se à água limpa e acessível para todos. O direito à segurança hídrica garante a todos o acesso a água suficiente, segura, aceitável, fisicamente acessível e acessível financeiramente para usos domésticos (UN, 2015). Entretanto, segundo Milan (2017), as alterações climáticas, o aumento nas projeções populacionais e a

consequente poluição resultante desafiarão esse objetivo e a maneira como ele interage com outros. A autora afirma que o ODS 6 interage de perto com muitos dos ODS propostos, especialmente em contextos urbanos, e ainda sugere que futuras pesquisas devem se concentrar em como esse objetivo seria impactado e executado em um clima em mudança, e quais medidas apropriadas são necessárias para reduzir ameaças e riscos decorrentes das mudanças climáticas à segurança hídrica e ao alcance de suas metas.

Figura 1 - Objetivos do Desenvolvimento Sustentável firmados na Agenda 2030



Fonte: ONU (2015)

A gestão sustentável dos recursos hídricos é essencial para alcançar os ODS e atender às metas estabelecidas pela Agenda 2030. Enfrentar os desafios associados às mudanças climáticas, ao crescimento populacional e à poluição hídrica requer ação coordenada em níveis global, nacional e local. Ao integrar práticas sustentáveis, tecnologias inovadoras e políticas eficazes, podemos garantir a preservação dos recursos hídricos para as gerações presentes e futuras.

2.1.1 Gestão de Recursos Hídricos no Brasil

O Brasil é reconhecido por possuir uma das maiores reservas de água doce do mundo, porém essa abundância não se traduz uniformemente em todo o país, tanto em termos geográficos quanto temporais. A disparidade na distribuição dos recursos hídricos, aliada à intensiva utilização da água em diversas atividades econômicas nas bacias hidrográficas brasileiras e aos desafios de qualidade associados à poluição hídrica, destacam a necessidade de implementação de medidas de gestão dos recursos hídricos cada vez mais eficazes (ANA, 2019).

A gestão dos recursos hídricos no Brasil é estabelecida pela Lei nº 9.433, de 08 de janeiro de 1997, que institui a Política Nacional de Recursos Hídricos (PNRH). Essa legislação abrange conceitos, fundamentos, diretrizes, instrumentos de gestão e a estrutura de governança. Os objetivos dessa política incluem garantir a disponibilidade necessária de água, em padrões

de qualidade adequados para diversos usos, promover a utilização racional e integrada dos recursos hídricos para o desenvolvimento sustentável, além de prevenir e defender contra eventos hidrológicos críticos de origem natural ou decorrentes do uso inadequado dos recursos naturais (BRASIL, 1997).

A lei indica que a gestão dos recursos hídricos deve seguir princípios como descentralização e participação social, considerando a bacia hidrográfica como unidade territorial de planejamento. Essa abordagem integrada não separa aspectos de qualidade e quantidade da água, reconhece a água como um bem público com valor econômico e busca integração com outras políticas setoriais, especialmente a ambiental. Além disso, a gestão leva em conta as particularidades regionais do país, considerando seus aspectos naturais, socioeconômicos, culturais e as políticas públicas de desenvolvimento regionais vigentes (ANA, 2022).

Com o intuito de alcançar seus objetivos, a legislação prevê a implementação de cinco instrumentos de gestão (Figura 2 e Quadro 1): os Planos de Recursos Hídricos; o enquadramento dos corpos de água em classes, segundo os usos preponderantes da água; a outorga dos direitos de uso de recursos hídricos; cobrança pelo uso de recursos hídricos; e o Sistema de Informações sobre Recursos Hídricos (BRASIL, 1997). Esses instrumentos têm como objetivo proporcionar meios e estruturar a administração por meio de iniciativas que abrangem planejamento, regulação, fiscalização e divulgação de informações.

Figura 2 - Os cinco instrumentos de gestão dos recursos hídricos previstos na PNRH



Fonte: Adaptado de ANA (2022)

Quadro 1 - Descrição dos cinco instrumentos de gestão dos recursos hídricos previstos na PNRH

Instrumento	Descrição
Planos de Recursos Hídricos	Fornecem diretrizes para a gestão e ações de regulação, enquadramento, cobrança e fiscalização. São elaborados por bacia, por estado ou para o país.
Enquadramento dos corpos de água	Estabelece metas de qualidade da água (classes), que podem variar ao longo do rio em função dos tipos de uso. Alguns usos são mais restritivos do que outros.
Outorga dos direitos de uso de recursos hídricos	É uma autorização de direito de uso da água obtida pelos usuários que causam algum impacto na qualidade e na quantidade de água na bacia hidrográfica.
Cobrança pelo uso de recursos hídricos	Ocorre para incentivar o uso racional da água pelos diversos usuários e os recursos arrecadados são utilizados para ações em prol dos recursos hídricos presentes na própria bacia.
Sistema de Informações sobre Recursos Hídricos	O SNIRH é um amplo sistema de coleta, tratamento, armazenamento e divulgação de informações sobre recursos hídricos

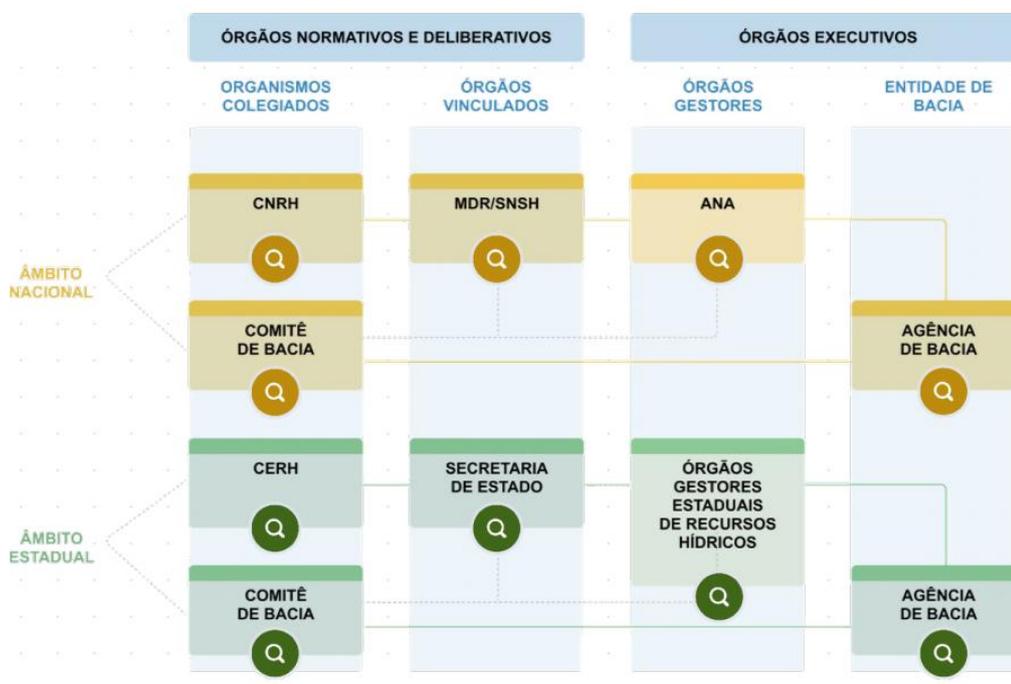
Fonte: Adaptado de ANA (2019)

Existe uma interligação entre os instrumentos da PNRH, e em alguns casos, uma interdependência entre eles. Um exemplo disso é a relação com a cobrança pelo uso dos recursos hídricos, que demanda a existência do plano de recursos hídricos da bacia para sua subsequente implementação. Da mesma forma, a concessão de outorga de direito de uso dos recursos hídricos requer a observância da classe de enquadramento do corpo d'água antes de ser emitida. O cadastro de usuários e a fiscalização são ações destinadas à regularização dos usos, impactando principalmente nas áreas de outorga e cobrança. É essencial que todas essas atividades estejam acessíveis por meio de um sistema de informações bem estruturado e constantemente atualizado (ANA, 2022).

A Lei da Águas também instituiu o Sistema Nacional de Gerenciamento dos Recursos Hídricos (SINGREH), criado para implementar a PNRH por meio das instituições que o integram (Figura 3) (BRASIL, 1997). O SINGREH representa a estrutura de governança estabelecida no Brasil para a administração dos recursos hídricos. Ele é composto por organismos colegiados, responsáveis por debates e deliberações acerca da gestão dos recursos hídricos (atuando como órgãos consultivos e deliberativos), além de órgãos administrativos encarregados de implementar a PNRH. Estes entes operam em âmbito federal ou estadual, dependendo do domínio dos corpos hídricos em questão. O SINGREH abrange o Conselho Nacional de Recursos Hídricos (CNRH), a Secretaria Nacional de Segurança Hídrica (SNSH) vinculada ao Ministério da Integração e do Desenvolvimento Regional (MIDR), a Agência Nacional de Águas e Saneamento Básico (ANA), os Conselhos Estaduais de Recursos Hídricos

(CERHs), os órgãos gestores estaduais de recursos hídricos, os comitês de bacia hidrográfica (intermunicipais e estaduais) e as agências de água (ligadas aos comitês) (ANA, 2019).

Figura 3 - Matriz institucional dos integrantes do SINGREH



Fonte: Adaptado de ANA (2022)

2.1.2 Gestão de Recursos Hídricos em Pernambuco

O estado de Pernambuco situa-se na região Nordeste do Brasil, possui uma área de aproximadamente 98.068 km², sendo a maior parte na região semiárida. Possui uma população estimada de 9.058.931 habitantes distribuída em 185 municípios (IBGE, 2022). Segundo a Agência Pernambucana de Águas e Clima (APAC), as grandes bacias hidrográficas de Pernambuco possuem duas vertentes: o rio São Francisco e o Oceano Atlântico. As bacias que escoam para o rio São Francisco formam os chamados rios interiores, tendo como principais: Pontal, Garças, Brígida, Terra Nova, Pajeú, Moxotó, Ipanema, além de grupos de pequenos rios interiores. As bacias que escoam para o Oceano Atlântico constituem os chamados rios litorâneos, sendo os principais: Goiana, Capibaribe, Ipojuca, Sirinhaém, Una e Mundaú e grupos de pequenos rios litorâneos (APAC, 2023).

A gestão dos recursos hídricos no estado é estabelecida pela Lei nº 12.984, de 30 de dezembro de 2005, que instituiu a Política Estadual de Recursos Hídricos (PERH). Baseada na PNRH, a lei estadual introduziu os objetivos, fundamentos, diretrizes e instrumentos para a

gestão dos recursos hídricos em Pernambuco, substituindo a Lei nº 11.426, de 17 de janeiro de 1997. Na PERH, o estado dispõe de mais dois instrumentos além dos citados na PNRH, que são: a fiscalização do uso de recursos hídricos e o monitoramento dos recursos hídricos. A Lei nº 12.984/2005 também instituiu o Sistema Integrado de Gerenciamento de Recursos Hídricos (SIGRH/PE), que tem por finalidade formular, atualizar, aplicar, coordenar e executar a PERH. No tocante à sua estrutura organizacional, o SIGRH/PE tem a seguinte composição básica: Conselho Estadual de Recursos Hídricos (CERH); Comitês de Bacia Hidrográfica (CBHs); Órgão gestor de recursos hídricos do estado; Órgãos executores do SIGRH/PE; Organizações civis de recursos hídricos e Agências de Bacia (PERNAMBUCO, 2005).

Nesse contexto, a Lei Ordinária nº 14.028, de 26 de março de 2010 criou a APAC, vinculada à SEINFRA (antiga Secretaria de Infraestrutura e Recursos Hídricos, atual Secretaria de Mobilidade e Infraestrutura - SEMOBI) e atuante na gestão de recursos hídricos com o objetivo de fortalecer o planejamento e a regulação dos seus usos múltiplos no estado e consolidar o SIGRH/PE. A APAC ainda gerencia o monitoramento hidrometeorológico dos rios e reservatórios no estado, além de realizar previsões de tempo e temperatura (APAC, 2017).

A Agência Estadual de Meio Ambiente (CPRH), desde 1984, realiza o monitoramento sistemático da qualidade da água nas bacias hidrográficas do Estado, disponibilizando para o usuário os Mapas da Qualidade das Águas Superficiais de Pernambuco com os resultados do monitoramento dos últimos anos. Desde 2017 existe uma parceria entre a CPRH e a APAC, para atender ao programa QUALIÁGUA da ANA. A CPRH atua no monitoramento dos rios com 84 estações de amostragem, enquanto a APAC realiza o mesmo em 54 estações nos reservatórios (CPRH, 2023).

O primeiro Plano Estadual de Recursos Hídricos de Pernambuco (PERH/PE) foi concluído em 1998. O objetivo principal do PERH/PE é realizar o planejamento do uso dos recursos hídricos com o intuito de assegurar a qualidade, disponibilidade, conservação e utilização racional desses recursos. Tais esforços visam beneficiar não apenas as gerações atuais, mas também as futuras, promovendo, assim, o desenvolvimento sustentável da região. Em 2022 o plano passou por uma atualização, onde uma nova divisão do território em Unidades de Planejamento Hídrico foi introduzida (Figura 4). Essa reorganização teve como base critérios técnicos, como hidrogeologia, geologia, uso do solo, rede de adutoras e a presença de perímetros irrigados, entre outros. Como resultado, o número de Unidades de Planejamento foi reduzido de 29 para 16. Essa redução ocorreu principalmente pela agregação de Unidades de Planejamento que originalmente correspondiam a grupos de bacias de pequenos rios às bacias de grandes rios adjacentes (PERNAMBUCO, 2022).

Figura 4 - Unidades de Planejamento Hídrico definidas no PERH/PE - 2022



Fonte: PERNAMBUCO (2022)

O PERH/PE se trata de um Plano Diretor de Recursos Hídricos (PDRH), que é um dos instrumentos da política estadual que pode ser elaborado para uma bacia hidrográfica (ou grupos de bacias hidrográficas), Unidade de Planejamento Hídrico ou para todo o estado (caso do PERH/PE). Os PDRH constituem estudos que envolvem a elaboração de diagnósticos abrangendo os aspectos físicos, bióticos, socioeconômicos e relacionados aos recursos hídricos. A partir desses diagnósticos, são desenvolvidos prognósticos, nos quais se projeta a evolução de variáveis não controláveis em cenários futuros baseados em premissas conhecidas. Com base nessas projeções, são propostas linhas de ação, delineadas na forma de um Plano de Investimentos (APAC, 2023).

Alguns Planos Diretores foram desenvolvidos para determinadas bacias hidrográficas, focando principalmente no diagnóstico dos recursos hídricos. Recentemente, foram elaborados Planos Hidroambientais (PHA), que representam a nomenclatura atualizada dos antigos PDRH. Essa mudança ocorreu após a incorporação e integração desses planos com a abordagem ambiental. Atualmente, Pernambuco conta com o PHA Capibaribe (PERNAMBUCO, 2010a), o PHA Ipojuca (PERNAMBUCO, 2010b) e o PHA Una, GL4 e GL5 (APAC, 2019).

Segundo a PERH, os PDRH devem ser elaborados com o envolvimento e aprovação dos respectivos CBHs, bem como assegurada a efetiva participação dos municípios e da sociedade civil organizada. Os comitês são órgãos colegiados consultivos e deliberativos, com área de atuação abrangendo a totalidade de uma bacia hidrográfica, uma sub-bacia tributária ou um grupo contíguo de bacias ou sub-bacias hidrográficas. O estado de Pernambuco possui 8

comitês de bacia hidrográfica formados: Metropolitano Sul, Metropolitano Norte, Capibaribe, Goiana, Ipojuca, Una, Pajeú e Sirinhaém (APAC, 2023).

2.1.3 Avaliação da qualidade da água

O declínio da qualidade da água se revelou como uma preocupação global à medida que as atividades antropogênicas se expandem e as alterações climáticas ameaçam provocar mudanças significativas no ciclo da água (ABBOTT et al., 2019). A preocupação global é centrada na garantia de uma quantidade suficiente de água para a saúde pública, segurança alimentar e atendimento à crescente demanda de acesso aos recursos hídricos. Com uma estimativa de uma população global de 9,7 bilhões de pessoas até 2050, segundo a UNESCO, espera-se um aumento na demanda mundial de água entre 20% e 30% (WWAP, 2019). No entanto, já enfrentamos grandes desafios para suprir a demanda atual por água de boa qualidade, o que torna mais difícil assegurar esse recurso a longo prazo (MELLO et al., 2020).

Diante da projeção de um futuro mais quente e incerto, também é previsto que a deterioração da qualidade da água será agravada (MOLINA-NAVARRO et al., 2014). A análise dos impactos das alterações climáticas nos recursos hídricos desempenha um papel fundamental no desenvolvimento de políticas e estratégias de planejamento e gestão da água, tornando-as mais adaptáveis a essas mudanças do clima. Esta informação pode ser utilizada para minimizar os impactos negativos diante de um futuro imprevisível, contribuindo assim para o desenvolvimento sustentável (LI et al., 2020).

Nesse contexto, a qualidade da água varia de acordo com a localização, o período, o clima e a presença de fontes de poluição pontuais e difusas (GIRI; QIU, 2016). As fontes pontuais são confinadas e facilmente identificáveis, porém ainda existem regiões onde o esgotamento sanitário é insuficiente, tornando esse tipo de fonte de poluição um grande problema. Já no caso das fontes difusas, elas têm uma natureza mais complexa pois interagem com o uso da terra, o solo, o manejo e o clima em toda a paisagem, sendo assim mais difíceis de controlar (GIRI, 2021). São medidas variáveis para caracterização das fontes e análise da saúde do corpo hídrico, que também fornecem fundamentos que facilitam a implementação adequada de medidas corretivas para um sistema ambiental saudável.

A eutrofização é uma consequência da poluição da água, caracterizada pelo crescimento acelerado de organismos fotossintetizantes devido ao enriquecimento de nutrientes, especialmente nitrogênio e fósforo, em um sistema aquático. Esse aumento nos nutrientes resulta na depleção de oxigênio no corpo d'água, ocasionada pela intensa mineralização

bacteriana. Embora a eutrofização seja um fenômeno natural que ocorre ao longo do tempo geológico durante a evolução de ecossistemas aquáticos, a ação humana pode acelerá-la, sendo então denominada eutrofização cultural. Nas últimas décadas, a eutrofização emergiu como um problema significativo para as massas de água superficiais em todo o mundo (BRAGA et al., 2005; TUNDISI.; TUNDISI, 2008; LE MOAL et al.; 2019).

A modificação nos padrões de uso da terra devido ao aumento da população e sua demanda por recursos está substituindo florestas por agricultura, pastagem e urbanização. A conversão dessas áreas florestais em outros usos tem o potencial de degradar fortemente a qualidade da água, perturbando os ecossistemas aquáticos e prejudicando o abastecimento humano de água (MELLO et al., 2020). Os indicadores de qualidade da água têm respostas diferentes aos padrões de uso da terra, e diferentes escalas espaciais e temporais podem contribuir para esses efeitos. A definição de uma escala adequada (dimensão espacial e temporal) e a consideração da influência do uso da terra na avaliação da qualidade da água pode ajudar a desenvolver a gestão das bacias hidrográficas e as projeções futuras (MELLO et al., 2018).

Os impactos do uso da terra na qualidade da água são mensurados na maioria dos estudos em três escalas: ponto de amostragem, zona ribeirinha e bacia hidrográfica (MELLO et al., 2020). Utilizar uma abordagem multiescala ajuda na identificação de áreas críticas na bacia, dando suporte à implementação de melhores práticas de gestão para controlar a poluição difusa (GIRI, 2021), visto que não é claro se efeitos locais são mais fortes do que efeitos regionais e vice-versa (ZORZAL-ALMEIDA et al., 2018).

Usualmente, a análise da qualidade da água de bacias hidrográficas é realizada de três maneiras: (i) Modelagem hidrológica/qualidade da água; (ii) Modelagem estatística e (iii) Monitoramento (GIRI; QIU, 2016). O uso de algoritmos de Aprendizado de Máquina (*Machine Learning*) e Inteligência Artificial para melhorar a modelagem da qualidade da água também é uma técnica que está em ascensão nos últimos anos (GIRI et al., 2019; MOHAMMED; LONGVA; SEIDU, 2018; MOHAN et al., 2025; PÉREZ-BELTRÁN et al., 2024). Essas técnicas são comumente aplicadas para fins regulatórios/de formulação de políticas, planejamento e exploração.

Os programas de monitoramento da qualidade da água produzem uma grande matriz de dados com muitos parâmetros físico-químicos. Para a interpretação dessa massa de dados por meio de um método simples e que fornece resultados aceitáveis, é aplicada a Modelagem Estatística com técnicas multivariadas como a Análise de *Cluster*, a Análise Discriminante e a Análise de Componentes Principais/Análise Fatorial. Essas metodologias têm se mostrado

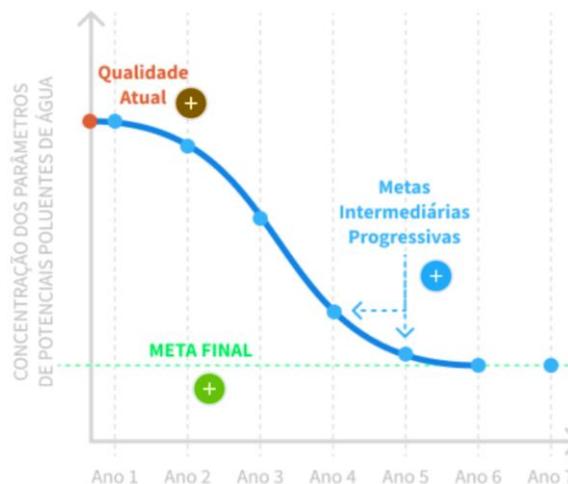
eficientes na redução de conjunto de dados complexos para uma dimensão consideravelmente baixa, bem como na obtenção dos padrões invisíveis dentro dos dados originais (LI et al., 2017). Com base nessa avaliação, gestores de bacias hidrográficas e formuladores de políticas são informados e é facilitada a tomada de ações correspondentes para proteger a qualidade da água.

Em relação à análise dos efeitos do uso da terra na qualidade da água, muitos estudos utilizam a técnica estatística Análise de Redundância (DING et al., 2016; MELLO et al., 2018; HAN et al., 2023; PAK et al., 2021; SHEN et al., 2015; WEI et al., 2020; XU et al., 2021; ZHANG et al., 2019; ZORZAL-ALMEIDA et al., 2018) e modelos de regressão (BU et al., 2014, 2016; MELLO et al., 2018; PAK et al., 2021; PRATT; CHANG, 2012; SINGH; BANSAL, 2025; ZHOU et al., 2016). A abordagem de métricas do uso da terra e da paisagem também são frequentemente usadas para prever a qualidade da água (DING et al., 2016; HAN et al., 2023; PAK et al., 2021; SHEN et al., 2015; ZHANG et al., 2019).

A avaliação da qualidade das águas superficiais no Brasil enfrenta desafios devido à ausência de redes estaduais de monitoramento em algumas unidades da federação e à heterogeneidade nas redes existentes, incluindo variações no número de parâmetros analisados e na frequência de coleta. A Rede Hidrometeorológica Nacional (RHN) e a Rede Nacional de Monitoramento da Qualidade de Água (RNQA), sob gestão da ANA, representam as principais fontes de dados no Brasil. Enquanto a RHN concentra-se principalmente no monitoramento quantitativo, abrangendo parâmetros básicos como pH, temperatura, oxigênio dissolvido (OD), turbidez e condutividade elétrica, a RNQA, operacionalizada pelo Programa QUALIÁGUA, amplia o escopo ao incluir parâmetros adicionais que requerem coleta e análises laboratoriais. Nos últimos anos, a RNQA experimentou um aumento significativo no número de pontos de monitoramento, quase dobrando de 1.486 em 2016 para 2.938 em 2020 (ANA, 2022).

A importância do monitoramento para a gestão eficaz dos recursos hídricos é destacada pela ANA (2022), pois contribui para a elaboração de planos de recursos hídricos e o enquadramento dos corpos hídricos em classes de uso. As referências para estabelecer as classes de qualidade são as resoluções CONAMA nº 357, de 17 de março de 2005, e nº 396, de 03 de abril de 2008, que consideram a bacia hidrográfica como unidade de gestão e os usos preponderantes mais restritivos, conforme estabelecido pela Resolução CNRH nº 91 de 2008. O enquadramento define metas progressivas de qualidade da água (Figura 5), alinhadas aos usos preponderantes, sujeitas à aprovação do Comitê de Bacia Hidrográfica e do respectivo Conselho de Recursos Hídricos, seja a nível estadual ou nacional, dependendo do domínio do corpo d'água.

Figura 5 - Comportamento da curva correspondente às metas progressivas de qualidade da água ao longo do tempo



Fonte: Adaptado de ANA (2022)

A Resolução CONAMA nº 357/2005 estipula limites permitidos para parâmetros físicos, químicos e biológicos em cinco classes de enquadramento para águas doces, e quatro classes para águas salobras e salinas. A classe especial, a mais restritiva, preserva as condições naturais do corpo d'água, proibindo lançamentos de efluentes, mesmo tratados. A numeração da classe indica a exigência de qualidade da água, sendo que classes de número mais alto permitem usos menos restritivos. Navegação e harmonia paisagística são usos menos exigentes, enquanto a preservação do equilíbrio natural das comunidades aquáticas e abastecimento humano após simples tratamento com desinfecção são mais exigentes. Até que os enquadramentos sejam aprovados, águas doces são consideradas classe 2 e salinas ou salobras, classe 1, a menos que as condições atuais justifiquem a aplicação da classe mais rigorosa.

A avaliação da qualidade da água envolve a análise de suas características físicas, bem como das substâncias e organismos presentes. Para simplificar a comunicação sobre a qualidade da água, são utilizados índices que classificam os recursos hídricos em categorias, funcionando como alertas para situações de impacto ambiental significativo. Um dos métodos mais comuns para essa avaliação é o Índice de Qualidade da Água (IQA), que abrange nove parâmetros (OD, DBO, temperatura, pH, coliformes termotolerantes, nitrogênio, fósforo, sólidos e turbidez), cada um ponderado de acordo com sua relevância para a qualidade global da água. O IQA foi desenvolvido para avaliar a qualidade da água bruta destinada ao abastecimento público, após tratamento, por meio de parâmetros que são indicadores de contaminação por lançamento de esgotos domésticos (ANA, 2022).

É preciso destacar que, embora o IQA seja vantajoso para a comunicação com o público leigo, ele apresenta a desvantagem de não contemplar variáveis importantes para o

abastecimento público, como metais pesados, compostos orgânicos com potencial mutagênico, substâncias que afetam as propriedades organolépticas da água, quantidade de células de cianobactérias e o potencial de formação de trihalometanos (CETESB, 2019).

Além do IQA, outro indicador bastante utilizado na avaliação da qualidade da água é o Índice de Estado Trófico (IET). O IET faz referência aos diferentes graus de trofia de um corpo hídrico, quantificados por meio de variáveis que se relacionam diretamente com o processo de eutrofização. Geralmente são empregadas a “clorofila-a” (admitida como uma medida da biomassa de algas), as espécies de algas presentes, a transparência das águas e as concentrações de nutrientes e oxigênio dissolvido (ADAMOVICH et al, 2019).

Segundo Braga et al. (2005), a classificação trófica dos ecossistemas lacustres de acordo com a produtividade biológica distingue classicamente três tipos de lagos:

- Oligotróficos: lagos com baixa produtividade e baixa concentração de nutrientes;
- Mesotróficos: lagos com características intermediárias entre oligotrófico e eutrófico;
- Eutróficos: lagos com elevada produtividade e alta concentração de nutrientes.

O Quadro 2 a seguir apresenta uma adaptação de Von Sperling (2005) para caracterização qualitativa entre os principais graus de trofia conjuntamente com mais dois níveis extremos (ultraoligotrófico e hipereutrófico), de forma a se caracterizar com uma particularidade ainda mais elevada os corpos d’água.

Quadro 2 - Caracterização trófica dos corpos hídricos

Item	Classe de trofia				
	Ultraoligotrófico	Oligotrófico	Mesotrófico	Eutrófico	Hipereutrófico
Biomassa	Bastante baixa	Reduzida	Média	Alta	Bastante alta
Fração de algas verdes e/ou cianobactérias	Baixa	Baixa	Variável	Alta	Bastante alta
Macrófitas	Baixa ou ausente	Baixa	Variável	Alta ou baixa	Baixa
Dinâmica de produção	Bastante baixa	Baixa	Média	Alta	Alta, instável
Dinâmica de oxigênio na camada superior	Normalmente saturado	Normalmente saturado	Variável em torno da supersaturação	Frequentemente supersaturado	Bastante instável, de supersaturação à ausência
Dinâmica de oxigênio na camada inferior	Normalmente saturado	Normalmente saturado	Variável abaixo da saturação	Abaixo da saturação à completa ausência	Bastante instável, de supersaturação à ausência
Prejuízo aos usos múltiplos	Baixo	Baixo	Variável	Alto	Bastante alto

Fonte: Adaptado de VON SPERLING (2005)

Assim como no caso do IQA, as limitações do IET também devem ser consideradas, já que para o seu cálculo, geralmente utiliza-se apenas dois parâmetros: fósforo total e clorofila-a. Esteves (2011) destaca que a classificação trófica de um ecossistema aquático deve se basear em um número mais amplo possível de características, especialmente ao aplicar em regiões tropicais, onde os corpos hídricos são classificados com base em indicadores geralmente obtidos por meio de amostragens esporádicas.

2.2 CIÊNCIA DE DADOS

A sociedade passou de predominantemente analógica para digital em apenas alguns anos, gerando um grande impacto na forma como fazemos negócios e nos comunicamos. A sociedade (organizações e pessoas) está "sempre conectada", produzindo dados sobre qualquer coisa, que são coletados a qualquer momento e em qualquer lugar. Hoje em dia, o termo *Big Data* (ou Dados Massivos) se refere a conjuntos de dados enormes e complexos que são tão grandes que o *software* tradicional não consegue processar. No entanto, eles oferecem oportunidades para resolver problemas de negócios que anteriormente não eram considerados (VAN DER AALST, 2016).

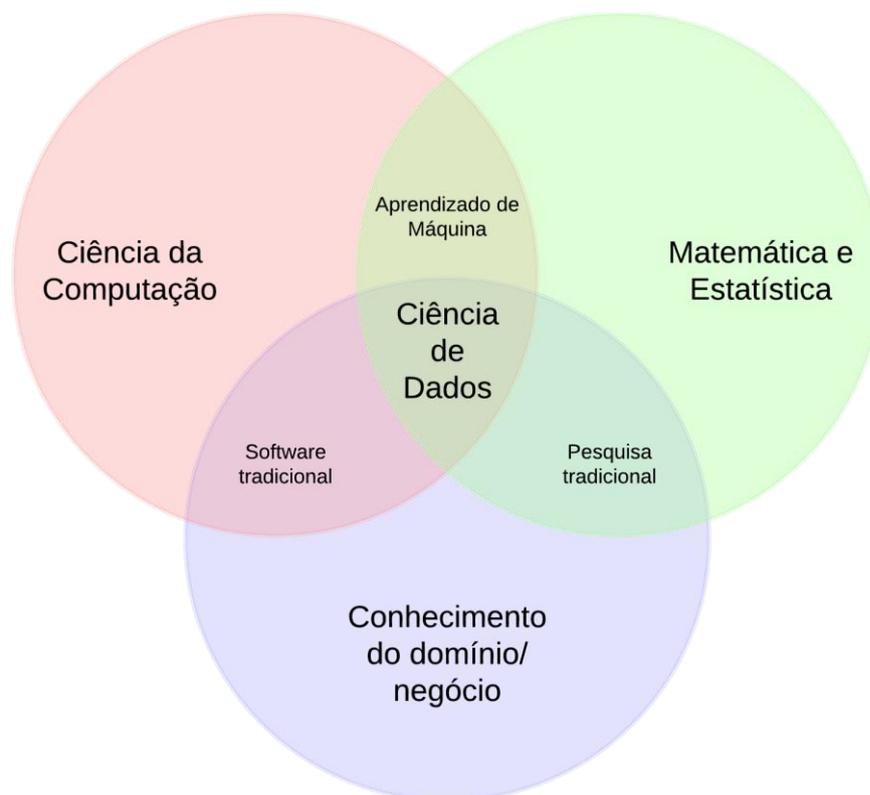
Os atributos que definem *Big Data* são volume, variedade, velocidade e variabilidade, normalmente chamados de quatro V's. O primeiro "V" (Volume) destaca a escala massiva dos dados, como exemplificado pelos milhares de usuários ativos de redes sociais e seus respectivos dados. O segundo "V" (Velocidade) refere-se à rapidez com que os dados são gerados e precisam ser processados, muitas vezes superando as abordagens tradicionais de processamento em lote. O terceiro "V" (Variedade) aborda a diversidade de tipos de dados, incluindo estruturados e não estruturados, que precisam ser combinados para obter *insights* completos. O quarto "V" (Veracidade) ressalta a importância da confiabilidade dos dados, considerando incertezas como dados de sensores e conteúdos gerados automaticamente (REIS et al., 2015; TSAI et al., 2015; VAN DER AALST, 2016). Posteriormente, outros V's foram propostos como relevantes: Valor – apontando a necessidade de filtrar dados para obter informações valiosas; Visualização – significando a necessidade de apresentar estruturas de dados complexas e informações ricas de forma eficaz (NATIVI et al., 2015).

No atual contexto do universo digital, a predominância de dados não estruturados apresenta um desafio significativo, principalmente para lidar com sua imensa quantidade. A extração de informações e valor desses dados se tornou um requisito fundamental para as organizações. Nesse cenário, a Ciência de Dados surge como uma resposta, buscando preencher

a lacuna entre dados e decisões. Para Gibert et al. (2018), a Ciência de Dados é entendida como um campo multidisciplinar que integra análise e processamento de dados com conhecimento especializado em um determinado domínio, resultando na transformação de dados em informações relevantes e acionáveis para decisões informadas.

O Diagrama de Venn da Ciência de Dados (Figura 6) de Conway (CONWAY, 2013) oferece uma representação de como habilidades de programação (chamadas por Conway de habilidades de *hacking*), conhecimentos em Matemática e Estatística, e *expertise* em um domínio específico (chamado por Conway de *expertise* substantiva) se combinam para possibilitar a Ciência de Dados. A maioria da pesquisa tradicional ocorre na intersecção entre *expertise* no domínio, conhecimentos em Matemática e Estatística. Programar com esses conhecimentos pode gerar *insights* através do Aprendizado de Máquina, mas sem considerar questões científicas específicas do domínio, a compreensão dos processos pode ser limitada. Habilidades de programação combinadas apenas com *expertise* no domínio podem levar a interpretações incorretas dos resultados sem conhecimento em Matemática e Estatística. A intersecção desses elementos é onde a Ciência de Dados é mais eficaz, evidenciando a importância da multidisciplinaridade nesse campo.

Figura 6 - Diagrama de Venn da Ciência de Dados



Fonte: Adaptado de CONWAY (2013)

Para van der Aalst (2016), a Ciência de Dados é uma combinação de diferentes subdisciplinas parcialmente sobrepostas, e a Figura 7 e o Quadro 3 mostram os seus principais elementos. O autor aponta que, no entanto, o diagrama deve ser interpretado com cautela, pois as subdisciplinas se sobrepõem e variam em tamanho. Além disso, os limites não são claros e parecem mudar ao longo do tempo. Considerando, por exemplo, a diferença entre Mineração de Dados e Aprendizado de Máquina ou Estatística. Suas origens são muito diferentes: a Mineração de Dados surgiu da comunidade de Bancos de Dados e o Aprendizado de Máquina surgiu da comunidade de Inteligência Artificial (IA), ambas bastante desconectadas da comunidade de Estatística. Apesar das origens diferentes, as três subdisciplinas definitivamente se sobrepõem no dia a dia da Ciência de Dados.

Figura 7 - Elementos que contribuem para a Ciência de Dados



Fonte: VAN DER AALST (2016)

Quadro 3 - Elementos que contribuem para a Ciência de Dados e suas descrições

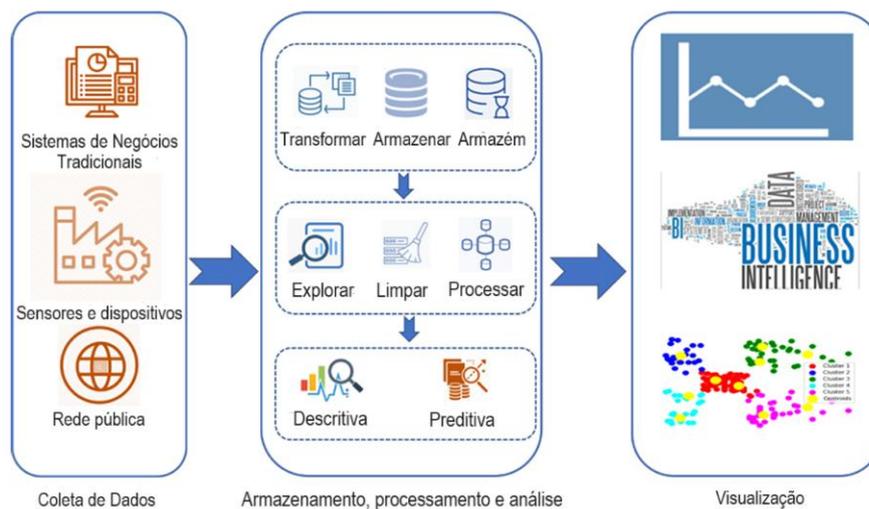
Elemento	Descrição
Estatística	É considerada a origem da Ciência de Dados, dividida em Estatística Descritiva (para resumir dados de amostra usando conceitos como média, desvio padrão e frequência) e Estatística Inferencial (usando dados de amostra para estimar características de todos os dados ou testar uma hipótese).
Algoritmos	São cruciais em qualquer abordagem de análise de dados. À medida que os conjuntos de dados aumentam, a complexidade dos algoritmos se torna uma preocupação primária.
Mineração de dados	É a análise de conjuntos de dados (geralmente grandes) para encontrar relacionamentos inesperados e resumir os dados de maneiras novas e úteis para o proprietário. Baseia-se em estatísticas, bancos de dados e algoritmos, e concentra-se na escalabilidade e em aplicações práticas.
Aprendizado de máquina	Busca construir programas de computador que melhorem automaticamente com a experiência. Esses algoritmos permitem que os computadores aprendam com os dados e façam previsões ou tomem decisões com base nesse aprendizado.
Mineração de processos	Adiciona a perspectiva de processo ao Aprendizado de Máquina e à Mineração de Dados, confrontando dados de eventos com modelos de processo para analisar conformidade e desempenho.
Análise preditiva	É a prática de extrair informações de conjuntos de dados existentes para determinar padrões e prever resultados e tendências futuras.
Bancos de dados	Armazenam dados e formam um dos pilares da Ciência de Dados. Sistemas de Gerenciamento de Banco de Dados (SGBD) estruturam dados para gerenciamento fácil e fornecem escalabilidade e desempenho confiável.
Sistemas distribuídos	Fornecem a infraestrutura para análise de dados em grande escala, dividindo tarefas em muitos computadores para alcançar escalabilidade.
Visualização e análise visual	São elementos essenciais da Ciência de Dados, permitindo que as pessoas interpretem os resultados e orientem a análise.
Modelos de negócio e <i>marketing</i>	São importantes na Ciência de Dados pois transformam dados em valor, incluindo valor comercial.
Ciências comportamentais e sociais	São relevantes para entender o comportamento humano e o contexto social em que indivíduos e organizações operam, porque a maioria dos dados é (indiretamente) gerada por pessoas e os resultados da análise são frequentemente usados para influenciar pessoas.
Privacidade, segurança, lei e ética	São fundamentais para proteger indivíduos e organizações de práticas inadequadas de Ciência de Dados. Os dados devem ser precisos e armazenados com segurança, não permitindo acesso não autorizado, e os indivíduos precisam confiar na forma como os dados são armazenados e transmitidos.

Fonte: Adaptado de VAN DER AALST (2016)

O avanço da Ciência de Dados trouxe consigo uma nova abordagem na tomada de decisões, especialmente no âmbito empresarial. Agora, as decisões são fundamentadas em dados, e o valor acrescentado às organizações não está mais ligado apenas à tecnologia ou ao capital, mas principalmente à informação. Os dados são considerados a principal fonte de conhecimento nesse contexto. Essa mudança não se limita ao mundo dos negócios, sendo amplamente reconhecido o valor que os processos de Ciência de Dados podem trazer para a compreensão de fenômenos complexos em diversas áreas (GIBERT et al., 2018).

Com o avanço das tecnologias *web*, uma grande quantidade de dados está sendo gerada em alta velocidade e variedade por fontes importantes como Internet das Coisas/sensores, redes sociais e sistemas de negócios tradicionais. Esses dados podem ter estruturas variadas, podendo ser estruturados, não estruturados ou uma combinação de ambos. A Ciência de Dados utiliza diversas abordagens para limpar, explorar e transformar esses dados antes de aplicar técnicas como Aprendizado de Máquina, visualização de dados e análise de texto para extrair informações. Além disso, a Ciência de Dados está profundamente integrada à Inteligência Artificial para viabilizar a tomada de decisões automatizada (KUMAR et al., 2021). A Figura 8 ilustra a arquitetura detalhada do processo de Ciência de Dados.

Figura 8 - Arquitetura do processo de Ciência de Dados



Fonte: KUMAR et al. (2021)

2.2.1 Visualização de dados usando ferramentas de *Business Intelligence*

Business Intelligence (BI) é um conjunto de tecnologias, processos e práticas para coletar, integrar, analisar e apresentar informações de negócios, que tem como objetivo principal ajudar as organizações a tomar decisões estratégicas com base em dados. As

ferramentas de BI acessam e analisam conjuntos de dados e apresentam descobertas analíticas em relatórios, resumos, painéis (*dashboards*), gráficos, tabelas e mapas para fornecer aos usuários uma visão detalhada sobre o estado do negócio (IBM, 2023). Embora a Ciência de Dados compartilhe alguns conceitos e técnicas com o BI, existem notáveis diferenças entre eles.

Enquanto o BI tem um foco mais voltado para o negócio e tende a usar dados estruturados e históricos para análise e relatórios (OLAVSRUD; FRUHLINGER, 2023), a Ciência de Dados é mais ampla em seu escopo e pode lidar com uma variedade de dados, estruturados e não estruturados, em tempo real ou em lote, incorporando técnicas avançadas de análise, como Aprendizado de Máquina e Mineração de Dados (VAN DER AALST, 2016). No entanto, tanto BI quanto Ciência de Dados compartilham o objetivo comum de extrair valor dos dados para apoiar a tomada de decisões por meio do uso da informação. Eles também podem se complementar, com as visualizações e análises geradas pelo BI alimentando os modelos e algoritmos da Ciência de Dados, e vice-versa.

Tanto as plataformas de BI quanto a Ciência de Dados tradicionalmente dependem de *Data Warehouses* (Armazém de dados) para obter os dados necessários. Um *Data Warehouse* é responsável por reunir informações de diversas fontes, como aplicativos, arquivos e bancos de dados. Esses dados são processados utilizando diferentes ferramentas para que estejam prontos para análise. Por exemplo, o *Data Warehouse* pode coletar dados de um banco de dados relacional que organiza as informações em tabelas de linhas e colunas (AWS, 2024).

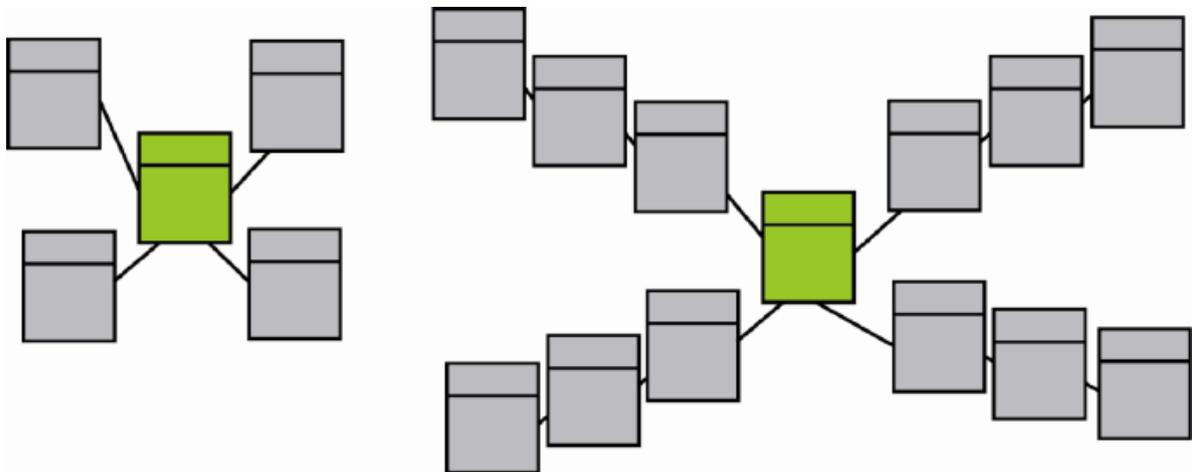
Os *Data Warehouses* podem incluir um mecanismo de processamento analítico *online* (OLAP, do inglês *Online Analytical Processing*) para dar suporte a consultas multidimensionais. OLAP é uma abordagem para análise e gestão de dados que se baseia em visualizações dimensionais. Essas visualizações são criadas utilizando esquemas estrela e, quando preenchidas com dados, são chamadas de cubos. Cubos ou *arrays* multidimensionais são estruturas utilizadas para armazenar valores pré-calculados em uma estrutura dimensional específica. Embora a construção dessas estruturas possa demandar tempo, elas são altamente eficientes em termos de leitura e recuperação de dados (OSIANLIS; ARNOTT, 2002).

O Esquema Estrela (em inglês, *Star Schema*) é uma técnica de modelagem de dados multidimensionais utilizados em *Data Warehouses* ou bancos de dados OLAP. Ele consiste em uma tabela principal, chamada de tabela de fatos, e várias tabelas secundárias, chamadas tabelas de dimensões. A tabela de fatos armazena valores numéricos relacionados a um processo de negócios, enquanto as tabelas de dimensões contêm informações descritivas sobre os atributos da tabela de fatos. A tabela de fatos faz referência às tabelas de dimensões usando chaves estrangeiras, que são identificadores únicos que relacionam os dados na tabela de fatos às

informações correspondentes nas tabelas de dimensões. No Esquema Estrela, a tabela de fatos se conecta a várias tabelas de dimensões, criando uma estrutura de dados que se assemelha a uma estrela.

Outra técnica de modelagem de dados é o Esquema Floco de Neve, que é uma extensão do Esquema Estrela. Nesse esquema, algumas tabelas de dimensões podem se desdobrar em uma ou mais tabelas de dimensões secundárias, criando assim uma estrutura mais complexa. Essa subdivisão das tabelas de dimensões resulta em uma forma semelhante a um floco de neve quando o modelo de dados é visualizado (AWS, 2024). A Figura 9 mostra um modelo simples de visualização de Esquema Estrela e Esquema Floco de Neve em um *Data Warehouse*.

Figura 9 - Esquema Estrela (esquerda) vs. Esquema Floco de Neve (direita)



Fonte: BERNIER et al. (2008)

A aplicação de sistemas de *Business Intelligence* em cenários reais oferece uma vantagem significativa ao fornecer soluções inovadoras para resolver problemas do mundo real e contribuir para o avanço teórico da área. Essa abordagem tem sido amplamente adotada, como destacado por Muntean (2018), que ressalta a importância da gestão da sustentabilidade empresarial e a utilização de métodos e ferramentas de *Business Intelligence* para analisar as dimensões financeira, ambiental e social dos negócios. A visualização eficaz desempenha um papel importante em aplicações de suporte à decisão, seja para apresentar dados brutos ou resultados de análises e simulações. As tecnologias atuais oferecem ferramentas tão avançadas e interativas quanto os aplicativos de *desktop* tradicionais. Por exemplo, gráficos e mapas *online* permitem a interação do usuário, como a leitura dos valores dos pontos de dados ao passar o cursor sobre um gráfico, a capacidade de ampliar/reduzir partes específicas de um mapa ou gráfico e a sobreposição de diferentes informações e cenários, conforme necessário (VITOLLO et al., 2015).

2.2.2 Ciência de Dados na área das Ciências Ambientais e de Recursos Hídricos

A Ciência de Dados abrange a análise extensiva de dados disponíveis, não se limitando apenas ao *Big Data*, mas também considerando a complexidade dos dados. Devido ao valor estratégico que pode fornecer, a área está ganhando cada vez mais interesse em diversos setores, incluindo as Ciências Ambientais. À medida que os conjuntos de dados ambientais se tornam maiores e mais complexos, a necessidade de Ciência de Dados em aplicações ambientais está em ascensão. A Ciência de Dados tem o potencial de trazer benefícios significativos para as Ciências Ambientais em diversos aspectos (GIBERT et al., 2018). Tomando como exemplo o contexto dos recursos hídricos, a relação entre água e economia é tão evidente que as decisões sobre o tema raramente são postergadas. Dessa forma, decisões que são tomadas sem informações adequadas geralmente resultam em consequências que afetam o meio ambiente, saúde e sociedade (SWERSKY et al., 2022).

A Ciência de Dados desempenha seu papel na análise, interpretação e geração de informações úteis a partir de uma ampla gama de dados ambientais provenientes de diversos tipos de sensores, tanto *in situ* quanto remotamente (ATHANASIADIS; MITKAS, 2007; REIS et al., 2015). Esses conjuntos de dados abrangem diversas áreas das Ciências Ambientais, desde observações climáticas e meteorológicas até a qualidade do ar, hidrologia, qualidade da água, e rastreamento de movimento de organismos biológicos, entre outros (HILL et al., 2011; WIEMANN et al., 2016; WONG; KERKEZ, 2016; KRANSTAUBER et al., 2011). Conforme os custos dos sensores e sistemas relacionados diminuem, o tamanho dessas redes de monitoramento e dos conjuntos de dados resultantes está aumentando, destacando a crescente necessidade de analistas capazes de gerenciar e integrar esses dados com modelos de simulação e outras aplicações (GIBERT et al., 2018).

Outra área em rápido crescimento é a interpretação de dados provenientes de drones aéreos, drones aquáticos e sensoriamento remoto via satélite (ROELOFSEN et al., 2014; ELARAB et al., 2015; GAUCI et al., 2018). Esses dados de sensoriamento remoto estão se tornando essenciais para um monitoramento ambiental preciso e abrangente. Dada a amplitude dessas aplicações, é impraticável depender apenas de sensores individuais ou estações terrestres eficazes. O volume de dados produzidos por drones e satélites de sensoriamento remoto pode ser significativo, exigindo habilidades especializadas em análise de dados e o uso de ferramentas avançadas para extrair *insights* importantes (GIBERT et al., 2018).

Tomando mais uma vez como exemplo os recursos hídricos, a medição e registro detalhados do uso da água em termos de tempo e espaço podem auxiliar os gestores a identificar

os diferentes usos finais, compreender a quantidade e o momento específico do consumo de água, e descobrir oportunidades para economizar. Como resultado, tem havido um aumento no uso de medições em alta frequência (da ordem de segundos ou ainda mais frequentes), que geram uma grande quantidade de dados capazes de oferecer informações importantes sobre os padrões de consumo de água. Esses dados podem revelar quais programas ou medidas de conservação estão sendo eficazes, quais não estão surtindo efeito, e quais soluções de eficiência energética são mais vantajosas. Para trabalhar com essas aplicações, a implementação de algoritmos de pós-processamento sofisticados e técnicas de Ciência de Dados são recomendadas para extrair informações significativas dos dados coletados. Esse conhecimento é especialmente relevante para grandes instituições, onde a implementação de práticas mais eficientes de uso de água pode resultar em economias consideráveis no orçamento (HORSBURGH et al., 2017).

Nos últimos anos, técnicas de Aprendizado de Máquina (em inglês, *Machine Learning*), têm sido intensamente exploradas na área de recursos hídricos (ZHU et al., 2022). Muitos pesquisadores têm usado essa área que faz parte do escopo da Ciência de Dados para resolver problemas em vários aspectos dos sistemas de tratamento e gerenciamento de água, incluindo monitoramento em tempo real (SAGAN et al., 2020), previsão e estimativa de concentração de poluentes (LIU; LU, 2014; PARSIAIE et al., 2018; ASHERI ARNON et al., 2019; MA et al., 2020), rastreamento de fontes poluentes (WU et al., 2020; BURAS et al., 2021), avaliação do risco de contaminação de águas subterrâneas (SAJEDI HOSSEINI et al., 2018), otimização da tecnologia de tratamento de água (ZHANG et al., 2019), entre outros.

A preparação de conjuntos de dados destinados a apoiar modelagens e análises de alto desempenho em larga escala (por exemplo, escala continental) é outra área em que a Ciência de Dados está sendo usada nas Ciências Ambientais. Esses conjuntos de dados, geralmente derivados de dados geoespaciais existentes, são organizados em grandes escalas espaciais usando modelagem de dados sofisticada para fornecer a base para modelos ambientais. Além disso, a análise de saídas de modelos climáticos de alta resolução e longo período temporal representa outro tipo de dados ambientais que requer habilidades especializadas. Analistas e profissionais que trabalham com modelagem lidam com grandes volumes de dados que precisam ser reduzidos para análises posteriores, enfrentando desafios não apenas no uso dos dados, mas também na gestão de *hardware* e *software* para esses conjuntos massivos (GIBERT et al., 2018).

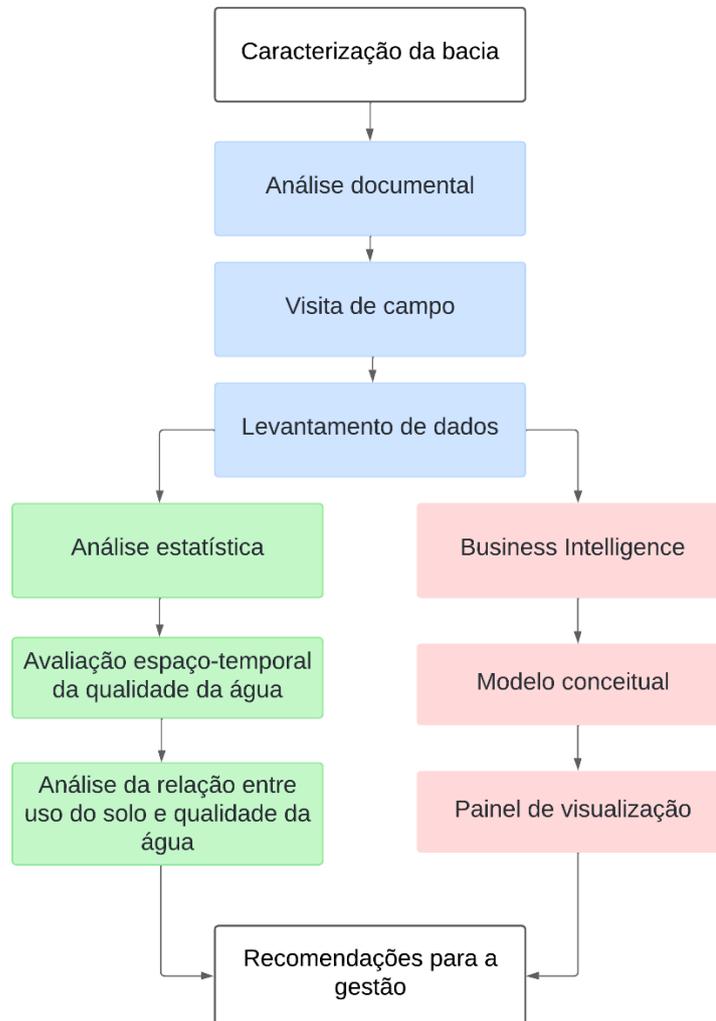
Outro ponto que merece ser destacado é que os sistemas ambientais do mundo real apresentam uma grande diversidade de dados heterogêneos provenientes de diversas fontes,

com diferentes formatos, resoluções e qualidades. Além disso, a coleta de dados ambientais geralmente é cara e difícil, com erros de medição significativos e amostragem espacial e temporal que podem não capturar completamente o comportamento do sistema (GIBERT et al., 2016). Os registros também podem conter valores discrepantes, ausentes e informações altamente incertas. Nesse contexto, as técnicas de Ciência de Dados têm demonstrado ser ferramentas eficazes para lidar com essa complexidade.

3 METODOLOGIA

Para aplicar a metodologia proposta para o desenvolvimento da pesquisa, optou-se por utilizar o estudo de caso da Bacia Hidrográfica do Rio Ipojuca (BHRI), situada no estado de Pernambuco, Brasil. Na Figura 10, apresenta-se um fluxograma que ilustra as etapas metodológicas adotadas para alcançar os objetivos do trabalho. O processo metodológico teve início com a caracterização da BHRI, seguido por etapas subsequentes delineadas para cumprir cada um dos objetivos específicos apresentados. Essas etapas foram detalhadas nos próximos tópicos deste capítulo.

Figura 10 - Fluxograma de execução da pesquisa

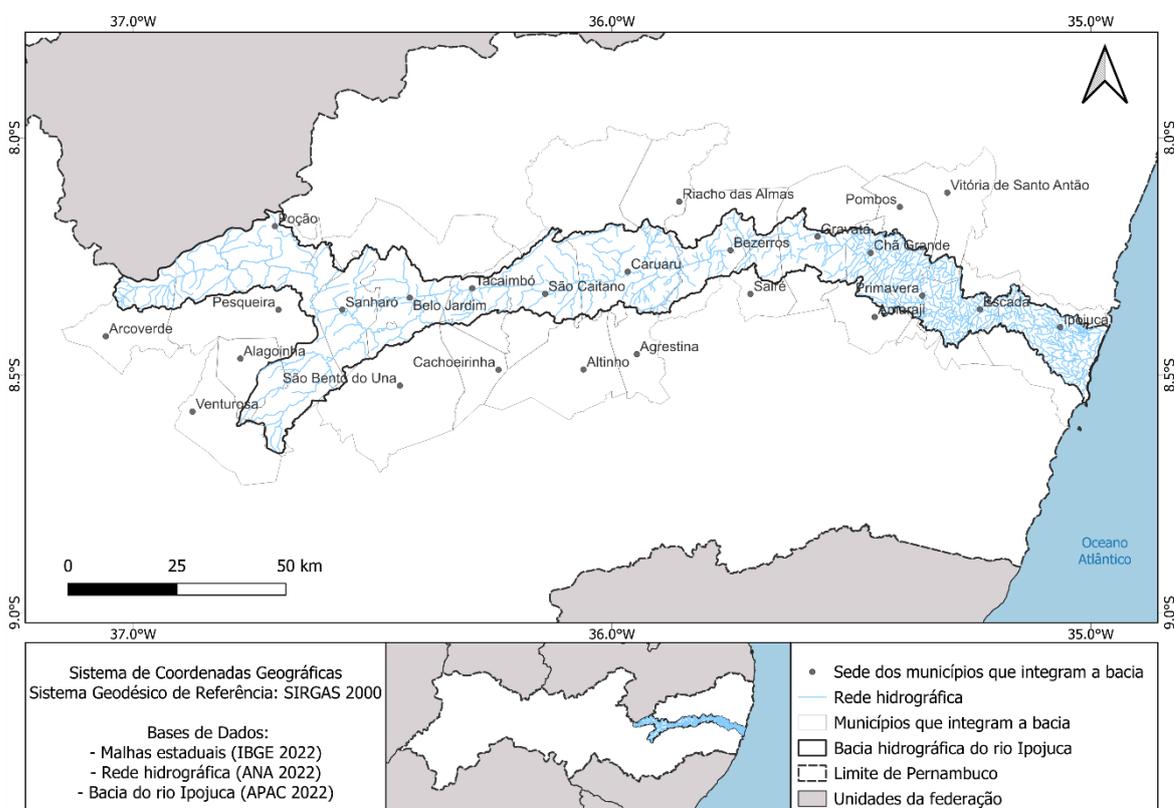


Fonte: Elaborado pela autora (2025)

3.1 CARACTERIZAÇÃO DA ÁREA DE ESTUDO

Conforme indicado no Plano Hidroambiental da Bacia Hidrográfica do Rio Ipojuca (PERNAMBUCO, 2010b), toda a área da bacia está localizada dentro do estado de Pernambuco (Figura 11). Ela está inserida na região hidrográfica do Atlântico Nordeste Oriental e abrange as Regiões de Desenvolvimento do Sertão do Moxotó, Agreste Meridional, Agreste Central, Mata Sul e Metropolitana. A bacia engloba áreas parciais de 25 municípios, sendo que apenas 12 deles têm suas sedes situadas dentro da própria bacia (Quadro 4).

Figura 11 - Localização da Bacia Hidrográfica do Rio Ipojuca



Fonte: Elaborado pela autora (2025)

Quadro 4 - Municípios que integram a Bacia Hidrográfica do Rio Ipojuca

Município	Área na bacia (%)	Município	Área na bacia (%)
Agrestina	0,04	Pesqueira	17,42
Alagoíinha	1,77	Poção*	5,34
Altinho	0,08	Pombos	1,95
Amaraji	1,76	Primavera*	2,60
Arcoverde	2,80	Riacho das Almas	0,24
Belo Jardim*	6,83	Sairé	2,25
Bezerros*	6,02	Sanharó*	7,12

Município	Área na bacia (%)	Município	Área na bacia (%)
Cachoeirinha	0,05	São Bento do Una	2,06
Caruaru*	11,31	São Caetano*	7,49
Chã Grande*	1,79	Tacaimbó*	4,10
Escada*	5,68	Venturosa	0,05
Gravatá*	5,55	Vitória de Santo Antão	1,14
Ipojuca*	4,45		

*Municípios com sede dentro da bacia

Fonte: PERNAMBUCO (2010b)

Com uma área de superfície de 3.433,58 km², a BHRI representa aproximadamente 3,49% do território total de Pernambuco. Sua configuração alongada no sentido Oeste-Leste confere-lhe uma posição estratégica dentro do estado, atuando como importante via hídrica de conexão entre a Região Metropolitana do Recife (RMR) e a região do Sertão, conforme destacado pelo CONDEPE/FIDEM em 2005.

O rio Ipojuca estende-se por cerca de 320 km, tendo suas nascentes na Serra do Pau d'arco, situada no município de Arcoverde. A bacia apresenta um regime fluvial intermitente até aproximadamente sua parte média, nas proximidades de Caruaru. A partir dos municípios de Gravatá e Chã Grande, o rio Ipojuca torna-se perene (APAC, 2013).

Os trechos médio e submédio da bacia estão situados na região Agreste do Estado de Pernambuco. Do ponto de vista climático, essa região é considerada intermediária entre as áreas de clima úmido (Zona da Mata) e de clima seco (Sertão), exibindo características ora de uma, ora de outra. Nas áreas mais próximas ao Sertão (trechos superior e parte do médio), o clima é quente e seco, com o período mais chuvoso ocorrendo de fevereiro a junho (chuvas de verão/outono). No trecho submédio (mais próximo da Zona da Mata), a estação chuvosa estende-se de março a julho (chuvas de outono/inverno). Por outro lado, o trecho inferior da bacia, em grande parte localizado na Zona da Mata, incluindo a faixa litorânea, apresenta características de clima quente e úmido, com médias pluviométricas superiores a 1.000 mm anuais, podendo ultrapassar 2.000 mm nas áreas litorâneas. O período chuvoso nessa região dura seis meses, indo de março a agosto (outono/inverno) (CONDEPE/FIDEM, 2005).

O Plano Hidroambiental da BHRI adotou a sua divisão em quatro unidades de análise (UA), que foram estabelecidas pelo Plano Diretor elaborado em 2002. A UA1 contempla o trecho mais a montante da bacia e tem sua seção de controle a jusante da sede do município de Belo Jardim; a UA2, a jusante da cidade de Caruaru; a UA3, a jusante da cidade de Gravatá; e a UA4, que corresponde ao trecho mais a jusante, tem sua seção de controle na foz do rio Ipojuca, no oceano Atlântico (PERNAMBUCO, 2010b).

Os principais reservatórios da bacia foram principalmente construídos para o abastecimento de água da população local (Quadro 5). Algumas dessas represas estão eutrofizadas devido à elevada poluição, enquanto suas áreas adjacentes (áreas de preservação permanente) exibem diversos graus de degradação ambiental devido ao uso inadequado do solo, necessitando de reabilitação para proteger as fontes de água, conforme indicado no Plano Hidroambiental da bacia (PERNAMBUCO, 2010b).

Quadro 5 - Principais reservatórios da BHRI

Reservatórios	Capacidade máxima (m ³)	Município	Finalidade
Pão de Açúcar	54.696.500	Pesqueira	Abastecimento urbano e irrigação
Pedro Moura Júnior (Belo Jardim)	30.740.000	Belo Jardim	Abastecimento urbano, piscicultura
Eng. Severino Guerra (Bitury)	17.776.470	Belo Jardim	Abastecimento urbano, irrigação horticultura
Manuíno	2.021.000	Bezerras	Abastecimento urbano
Brejão	1.625.000	Sairé	Abastecimento urbano
Taquara	1.100.000	Caruaru	Abastecimento rural
Duas Serras	2.032.289	Poção	Abastecimento urbano

Fonte: Adaptado de PERNAMBUCO (2010b) e APAC (2023)

A população urbana na bacia tem crescido rapidamente, com uma taxa de 1,2% ao ano entre 2015 e 2019, resultando em uma taxa de urbanização de 88% em 2019. Por outro lado, no mesmo período, a população rural diminuiu a uma taxa de -1,2% ao ano. Essa diminuição mais acentuada da população rural nos últimos anos pode ter sido impulsionada pela seca prolongada ocorrida na região, que provocou um aumento no êxodo rural (PERNAMBUCO, 2022).

No ano de 2013, o Instituto Brasileiro de Geografia e Estatística (IBGE) classificou o rio Ipojuca como o terceiro rio com a situação mais crítica no país, atrás apenas do rio Tietê, em São Paulo, e do rio Iguaçu, no Paraná. O modelo de desenvolvimento econômico dos municípios localizados na BHRI está diretamente relacionado à qualidade da água e à degradação ambiental observada. O estado enfrenta desafios como legislação e fiscalização deficientes, o que permite que setores como o sucroalcooleiro, indústrias e lavanderias de *jeans* despejem seus resíduos diretamente no rio. Esse cenário é particularmente evidente nos municípios de Ipojuca, Caruaru, Belo Jardim e Gravatá (COMPESA, 2016).

Além dos despejos do setor industrial, o lançamento de esgotos domésticos e resíduos sólidos ao longo da bacia é uma preocupação significativa que contribui para a eutrofização em alguns trechos do ambiente. A ocupação desordenada das margens do rio leva ao descarte indiscriminado de lixo e esgotos diretamente nas águas, sem qualquer tipo de tratamento prévio (PERNAMBUCO, 2010b; COMPESA, 2016).

A Tabela 1 apresenta os indicadores de saneamento dos 25 municípios que integram a bacia, extraídos do Sistema Nacional de Informações sobre Saneamento (SNIS, 2024) e referentes ao ano de 2022. Esses dados revelam uma cobertura insuficiente de saneamento na região, especialmente em relação à coleta e ao tratamento de esgoto nos municípios. As diretrizes do Novo Marco Legal do Saneamento estabelecem grandes metas, visando alcançar 99% da população com acesso à água potável e 90% com coleta e tratamento de esgotos até 2033 em todo o país (BRASIL, 2020). Portanto, fica evidente que muito ainda precisa ser investido na BHRI para atingir esses objetivos.

Tabela 1 - Indicadores de saneamento por município para o ano de referência 2022 (NI - não informado)

Município	Consumo médio <i>per capita</i> (L/hab/dia)	Índice de atendimento total de água (%)	Índice de coleta de esgoto (%)	Índice de tratamento de esgoto (%)
Agrestina	116,87	100	NI	0
Alagoinha	83,3	56,42	NI	0
Altinho	101,43	100	NI	0
Amaraji	NI	NI	NI	NI
Arcoverde	117,77	99,8	3,81	100
Belo Jardim	162,96	100	NI	50
Bezerros	103,54	94,78	NI	NI
Cachoeirinha	101,6	100	NI	0
Caruaru	149,02	100	40,87	100
Chã Grande	81,2	71,71	NI	33,33
Escada	88,81	92,73	NI	NI
Gravatá	168,57	100	12,89	100
Ipojuca	726,51	68,89	3,55	100
Pesqueira	140,89	74,54	NI	0
Poção	1,96	63,61	NI	0
Pombos	102,76	83,07	NI	NI
Primavera	91,17	50,54	NI	NI
Riacho das Almas	150,16	61,45	NI	0
Sairé	133,5	100	NI	NI
Sanharó	121,71	100	NI	NI
São Bento do Una	107,97	65,42	NI	NI
São Caetano	80,49	83,34	NI	NI
Tacaimbó	105,73	79,57	51,69	100
Venturosa	91,24	67,73	41,65	100
Vitória de Santo Antão	132,27	81,85	27,5	100

Fonte: SNIS (2024)

No que diz respeito ao Índice de Atendimento Total de Água, apenas 8 municípios relataram que 100% de sua população tem acesso a esse serviço. Apenas Arcoverde, Bezerros e Escada têm um índice de atendimento superior a 90%. Quanto ao Índice de Coleta de Esgoto, apenas 7 municípios afirmaram fornecer esse serviço, com uma cobertura inferior a 55%. Esses municípios também declararam tratar 100% do esgoto coletado. É importante destacar que essas informações são fornecidas pelas próprias prefeituras ou pelas concessionárias de água e esgoto que operam em cada localidade, o que pode resultar em disparidades com a realidade. Isso é notório no caso do consumo médio *per capita* de água nos municípios de Ipojuca e Poção, que apresentam valores muito acima e abaixo da média na bacia, respectivamente.

Os baixos índices de cobertura de saneamento básico na bacia têm impacto direto na qualidade da água do rio Ipojuca. A Agência Estadual de Meio Ambiente (CPRH) é responsável pelo monitoramento da qualidade da água do rio, iniciado em 1986 por meio de diversas estações de amostragem distribuídas em diferentes zonas homogêneas. A rede sofreu alterações ao longo do tempo, e as estações das quais se possuem dados estão detalhadas no Quadro 6. Ressalta-se que as estações IP-16, IP-21 e IP-24 foram incluídas na rede no final de 2021 e que o monitoramento das estações IP-01, IP-38, IP-85, IP-95 e IP-99 foi suspenso temporariamente. Atualmente, as estações ativas são monitoradas trimestralmente. Durante o monitoramento, são avaliados diversos parâmetros, incluindo temperatura da água, pH, condutividade elétrica, cloreto, oxigênio dissolvido (OD), demanda bioquímica de oxigênio (DBO), amônia, fósforo total, cromo, coliformes termotolerantes, salinidade, turbidez, nitrato, nitrito, teste de toxicidade aguda com *Daphnia magna*, sólidos totais, clorofila, potássio, alcalinidade, sólidos suspensos, sólidos dissolvidos, fósforo solúvel e nitrogênio total (CPRH, 2020). Este monitoramento é fundamental para avaliar a qualidade da água e identificar possíveis impactos ambientais decorrentes da falta de saneamento básico na região.

Quadro 6 - Estações de amostragem da rede de monitoramento da bacia do rio Ipojuca

Zona Homogênea	Estação	Corpo d'água	Local	Latitude	Longitude
Interesse Ambiental - trecho intermitente	IP-01	Rio Ipojuca	Na nascente do Rio Ipojuca, no Sítio Pedreira, em Arcoverde.	-8,3323	-37,0397
Pecuária Leiteira	IP-12	Rio Ipojuca	Na ponte da PE-180, que liga Belo Jardim a São Bento do Una.	-8,3865	-36,4234
Interesse Ambiental - brejo de altitude	IP-13	Rio Bitury	Nascente no Sítio de Seu Joaquim, na comunidade de Jussara, em Belo Jardim.	-8,2432	-36,4418
Interesse Ambiental - brejo de altitude	IP-14	Rio Bitury	Na ponte sobre o Rio Bitury, à montante do reservatório, em Belo Jardim.	-8,2824	-36,4350

Zona Homogênea	Estação	Corpo d'água	Local	Latitude	Longitude
Interesse Ambiental - brejo de altitude	IP-16	Riacho afluente do Rio Bitury	No Riacho afluente do Bitury, propriedade Dra. Elizabeth, em Belo Jardim.	-8,2729	-36,4474
Interesse Ambiental - brejo de altitude	IP-21	Riacho Taboquinha	Antes da Capela São Sebastião, em Belo Jardim.	-8,2977	-36,4512
Interesse Ambiental - brejo de altitude	IP-24	Riacho Taboquinha	Na passagem molhada a montante da Barragem do Bitury, em Belo Jardim.	-8,3026	-36,4561
Pecuária de Corte	IP-38	Rio Ipojuca	Na passagem molhada, próximo à Fazenda Pato Branco, à montante de São Caetano.	-8,3232	-36,1514
Urbana/Industrial - trecho intermitente	IP-49	Rio Ipojuca	A jusante da cidade de Caruaru, na Vila do Cedro (COAHB III), na cidade de Caruaru.	-8,2872	-35,9356
Pecuária de Corte	IP-55	Rio Ipojuca	Na ponte da BR-232, a montante de Gravatá, no município de Gravatá.	-8,2131	-35,6077
Policultura	IP-64	Rio Ipojuca	Na ponte à jusante da cidade de Chã Grande, no município de Chã Grande.	-8,2533	-35,4501
Agroindústria	IP-70	Rio Ipojuca	Na ponte à jusante da Usina União Indústria, no município de Primavera.	-8,3568	-35,3431
Urbana - trecho perene	IP-85	Rio Ipojuca	Na ponte BR-101 à jusante da cidade de Escada, no município de Escada.	-8,3640	-35,2190
Agroindústria	IP-90	Rio Ipojuca	Na ponte PE-60 à jusante da Usina Ipojuca, no município de Ipojuca.	-8,4046	-35,0668
Agroindústria	IP-95	Rio Ipojuca	À jusante da Usina Salgado, no município de Ipojuca.	-8,4159	-35,0113
Interesse Ambiental - área estuarina	IP-97	Estuário do Rio Ipojuca	No estuário dos rios Ipojuca e Merepe, em Suape, no município de Ipojuca.	-8,4068	-34,9784
Interesse Ambiental - mar	IP-99	Mar sob a influência do Rio Ipojuca	Próximo à desembocadura dos Rios Ipojuca e Merepe, no município de Ipojuca.	-8,4118	-34,9501

Fonte: CPRH (2020)

Com o objetivo de melhorar a qualidade da água do rio Ipojuca e seguindo as diretrizes do Plano Hidroambiental da bacia (PERNAMBUCO, 2010b), o governo do estado lançou o Programa de Saneamento Ambiental da Bacia Hidrográfica do Rio Ipojuca (PSA Ipojuca) em 2013, em parceria com o Banco Interamericano de Desenvolvimento (BID). Foram investidos US\$ 330 milhões no projeto, sendo US\$ 200 milhões provenientes do BID e US\$ 130 milhões de contrapartida estadual (COMPESA, 2016).

Conforme descrito pelo BID (2012), o projeto tem como finalidade desenvolver o saneamento ambiental da BHRI, por meio da ampliação da cobertura de esgotamento sanitário e os índices de tratamento de esgoto, principalmente nas 12 cidades-sede localizadas em suas

margens. Além disso, o projeto visa promover a recuperação das margens urbanas do rio e dos principais reservatórios de abastecimento de água da bacia. Para atingir este objetivo, o projeto conta com três conjuntos de ações específicas: i) a implementação de sistemas de coleta e tratamento de águas residuais nos municípios da bacia; ii) a recuperação de trechos das margens do rio em estágios avançados de degradação e das áreas de entorno dos reservatórios; e iii) ações de fortalecimento ambiental, técnico, operacional e financeiro da Companhia Pernambucana de Saneamento (COMPESA). O Quadro 7 mostra os empreendimentos propostos pelo projeto que já foram concluídos até junho de 2023.

Quadro 7 - Empreendimentos concluídos do PSA Ipojuca e suas respectivas fontes de recursos

Empreendimento	Recursos	Empreendimento	Recursos
Automação das Unidades Operacionais	BID	Implantação do SES de Venturosa (1ª Etapa)	Contrapartida
Implantação do SES Tacaimbó	BID	Implantação do SES de Itapetim	Contrapartida
Implantação do SES Gravatá 1ª etapa	BID	Implantação do SES de Garanhuns	Contrapartida
Ampliação do SES Caruaru - Limpeza, desobstrução, cadastramento e filmagem da rede existente	BID	Implantação da Adutora de Moxotó	Contrapartida
Ampliação do SES Caruaru - Núcleo de manutenção, almoxarifado e laboratório regional de esgoto e outros	BID	Implantação da Adutora do Agreste – Lote 1	Contrapartida
Elaboração de Projetos dos SES:		Obras de Implantação do parque urbano ambiental de Gravatá	BID
Implantação do SES de Poção, Chã Grande e Primavera	BID	Obras do Laboratório do CPRH	BID
Implantação do SES de Belo Jardim	BID	Obras de proteção da estação elevatória de esgoto da bacia A – EEEA (contra enchentes) - Sanharó	BID
Implantação do SES Gravatá 2ª Etapa	BID	Obras de Implantação do parque urbano ambiental de São Caetano	BID
Implantação do SES Caruaru	BID	Restauração Florestal - Pré-Plantio e Plantio	BID
Implantação do SES Bezerros	BID	Implantação do SES de Sanharó	BID
Implantação do SES Escada 2ª Etapa	BID	Implantação do SES Belo Jardim - SES 1ª etapa	BID
Modernização do Viveiro Florestal de Bonito	BID	Obras de Implantação da ETE Belo Jardim	BID
Obras complementares no SES Tacaimbó	BID	Obras de execução da complementação do SES Gravatá 1ª etapa	BID
Ampliação do SES Caruaru – Obras de recuperação de 6 EEES	BID	Obras de Implantação do parque urbano ambiental de Bezerros	BID
Substituição de adutora entre os municípios de Belo Jardim e São Bento do Una	BID	Implantação do SES Caraibeiras no município de Tacaratu	Contrapartida
Obras de adaptações das ETAs Petrópolis e Salgado	BID	Obras de Implantação do parque urbano ambiental de Belo Jardim	BID
Serviços Complementares do SES Tacaimbó	BID	Implantação do SES Escada 1ª etapa	BID
Implantação da Adutora de Porto de Galinhas	Contrapartida	Obras de Implantação do parque urbano ambiental de Escada	BID
Ampliação da ETA de Bezerros	Contrapartida		

Fonte: COMPESA (2023)

Legenda: Banco Interamericano de Desenvolvimento (BID), Estação Elevatória de Esgoto (EEE), Estação de Tratamento de Água (ETA), Estação de Tratamento de Esgoto (ETE), Sistema de Esgotamento Sanitário (SES).

3.2 PROCEDIMENTOS METODOLÓGICOS

Como mostrado por Conway (2013) e van der Aalst (2016), a Ciência de dados é uma combinação de diferentes subdisciplinas parcialmente sobrepostas, associando Ciência da Computação, Matemática/Estatística e Conhecimento do Domínio/Negócio.

O desenvolvimento desta pesquisa compreende etapas de avaliação da evolução espaço-temporal da qualidade da água, análise da relação entre o uso do solo e a qualidade da água e desenvolvimento do painel de visualização dos dados. Essas etapas têm como base a aplicação de algoritmos, bancos de dados, técnicas estatísticas, visualização e análise visual, que fazem parte dos campos da Ciência da Computação e da Matemática/Estatística. Os códigos em R desenvolvidos e utilizados neste trabalho estão disponíveis publicamente no repositório: <https://github.com/hidaiane/Tese>.

A discussão dos resultados, apresentada no próximo capítulo, corresponde ao conhecimento do domínio da BHRI. Dessa forma, para propor recomendações para a gestão da qualidade de água do rio Ipojuca, serão integradas ferramentas da Ciência de Dados adaptadas ao contexto ambiental.

3.2.1 Avaliação da evolução espaço-temporal da qualidade da água

Esta etapa da pesquisa visa fornecer uma compreensão das tendências e variações observadas na qualidade da água da BHRI ao longo do tempo e do espaço. Para isso, são discutidos o levantamento e seleção do banco de dados utilizado na pesquisa e detalhados os métodos de análise estatística aplicados.

3.2.1.1 Levantamento do banco de dados

Os dados de qualidade da água utilizados na pesquisa são provenientes do monitoramento da bacia hidrográfica do rio Ipojuca, executado pela CPRH de 2005 a 2016, e das campanhas do convênio CPRH-APAC de 2017 a 2023, atendendo ao programa QUALIÁGUA da Agência Nacional de Águas e Saneamento Básico. Ao todo, foram reunidos dados físico-químicos e biológicos ao longo de 86 coletas realizadas nos períodos chuvoso e seco.

Até 2016, a maioria das estações da rede de monitoramento operada pela CPRH geralmente realizava amostragens bimestrais. No entanto, houve uma mudança para

amostragens trimestrais devido a ajustes no laboratório da CPRH para implementação do sistema de qualidade e do programa QUALIÁGUA (Quadro 8). A Unidade de Análises Laboratoriais (ULAB) da CPRH foi responsável pelas atividades de coleta, análises físico-químicas e biológicas, bem como pela emissão de certificados. O método de ensaio adotado pelo laboratório segue o descrito no *Standard Methods for the Examination of Water and Wastewater – 23rd ed. 2017* (CPRH, 2020).

Quadro 8 - Frequência das coletas para monitoramento da qualidade da água da BHRI entre os anos de 2005 e 2023

Nº da Coleta	Mês	Ano	Nº da Coleta	Mês	Ano	Nº da Coleta	Mês	Ano
1	Fevereiro	2005	30	Fevereiro	2010	59	Outubro	2016
2	Abril		31	Abril		60	Dezembro	
3	Junho		32	Junho		61	Fevereiro	2017
4	Agosto		33	Agosto		62	Maio	
5	Outubro		34	Outubro		63	Agosto	
6	Dezembro		35	Dezembro		64	Novembro	
7	Fevereiro	2006	36	Fevereiro	2011	65	Fevereiro	2018
8	Abril		37	Abril		66	Maio	
9	Junho		38	Junho		67	Agosto	
10	Agosto		39	Agosto		68	Novembro	
11	Outubro		40	Outubro		69	Fevereiro	2019
12	Dezembro	41	Dezembro	70	Maio			
13	Fevereiro	2007	42	Fevereiro	2012	71	Agosto	2020
14	Abril		43	Abril		72	Novembro	
15	Junho		44	Junho		73	Fevereiro	
16	Agosto		45	Agosto		74	Novembro	
17	Outubro		46	Outubro		75	Fevereiro	2021
18	Dezembro		47	Dezembro		76	Maio	
19	Abril	2008	48	Fevereiro	2013	77	Agosto	2022
20	Junho		49	Abril		78	Novembro	
21	Agosto		50	Agosto		79	Fevereiro	
22	Outubro		51	Dezembro	80	Maio		
23	Novembro		52	Abril	2014	81	Agosto	
24	Dezembro	53	Junho	82		Novembro		
25	Fevereiro	2009	54	Abril	2015	83	Fevereiro	2023
26	Junho		55	Setembro		84	Maio	
27	Agosto		56	Abril	2016	85	Agosto	
28	Outubro		57	Junho		86	Novembro	
29	Dezembro		58	Agosto				

Fonte: Elaborado pela autora a partir dos dados de monitoramento disponibilizados pela CPRH (2025)

Além dos dados das estações de amostragem de qualidade da água, foram levantados dados complementares de postos pluviométricos e fluviométricos localizados nos municípios

da bacia. Os dados pluviométricos são provenientes de postos contidos no banco de dados da APAC e disponibilizados em seu portal, enquanto a série histórica dos dados fluviométricos está disponível no portal Hidroweb da ANA. Todos os cinco postos fluviométricos levantados estão localizados no leito do rio Ipojuca. Os postos pluviométricos (Quadro 9) e fluviométricos (Quadro 10) selecionados são apresentados a seguir.

Quadro 9 - Postos pluviométricos levantados na BHRI

Código	Nome	Município	Latitude	Longitude
20	Belo Jardim	Belo Jardim	-8,3367	-36,425
24	Caruaru (IPA)	Caruaru	-8,2383	-35,9158
117	Chã Grande	Chã Grande	-8,2422	-35,4592
452	Escada	Escada	-8,3667	-35,2333
58	Gravatá	Gravatá	-8,2006	-35,5431
387	Ipojuca	Ipojuca	-8,4011	-35,0622
108	Primavera	Primavera	-8,3483	-35,3475
114	Sanharó	Sanharó	-8,3639	-36,5664
122	São Caetano	São Caetano	-8,3283	-36,1375

Fonte: Elaborado pela autora a partir das informações disponibilizadas pela APAC (2025)

Quadro 10 - Postos fluviométricos levantados na BHRI

Código	Nome	Município	Latitude	Longitude	Área de drenagem (km³)
39310000	Sanharó	Sanharó	-8,3656	-36,5603	672
39340000	Caruaru	Caruaru	-8,3028	-36,0108	2000
39360000	Engenho Tabocas	Pombos	-8,2819	-35,3647	2960
39370100	Engenho Maranhão	Ipojuca	-8,3697	-35,1422	3310
39345000	Gravatá	Gravatá	-8,2128	-35,6042	2650

Fonte: Elaborado pela autora a partir das informações disponibilizadas pela ANA (2025)

3.2.1.2 Seleção e análise exploratória dos dados

Com relação ao critério de triagem das estações amostrais, optou-se por selecionar aquelas com maior frequência de dados e que não estivessem localizadas em ambiente de estuário ou mar. Dessa forma, estações que apresentavam mais de três anos sem nenhum dado coletado foram desconsideradas (Quadro 11). Como resultado, foram selecionados os pontos apresentados no Quadro 12, cuja distribuição espacial é ilustrada no mapa da Figura 12.

Quadro 11 - Frequência de monitoramento das estações de amostragem da bacia do rio Ipojuca

Estação	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023	Total de anos sem coleta	
IP-01																					12
IP-12																					3
IP-13																					1
IP-14																					1
IP-16																					16
IP-21																					16
IP-24																					16
IP-38																					8
IP-49																					1
IP-55																					1
IP-64																					1
IP-70																					0
IP-85																					4
IP-90																					0
IP-95																					4
IP-97																					0
IP-99																					7

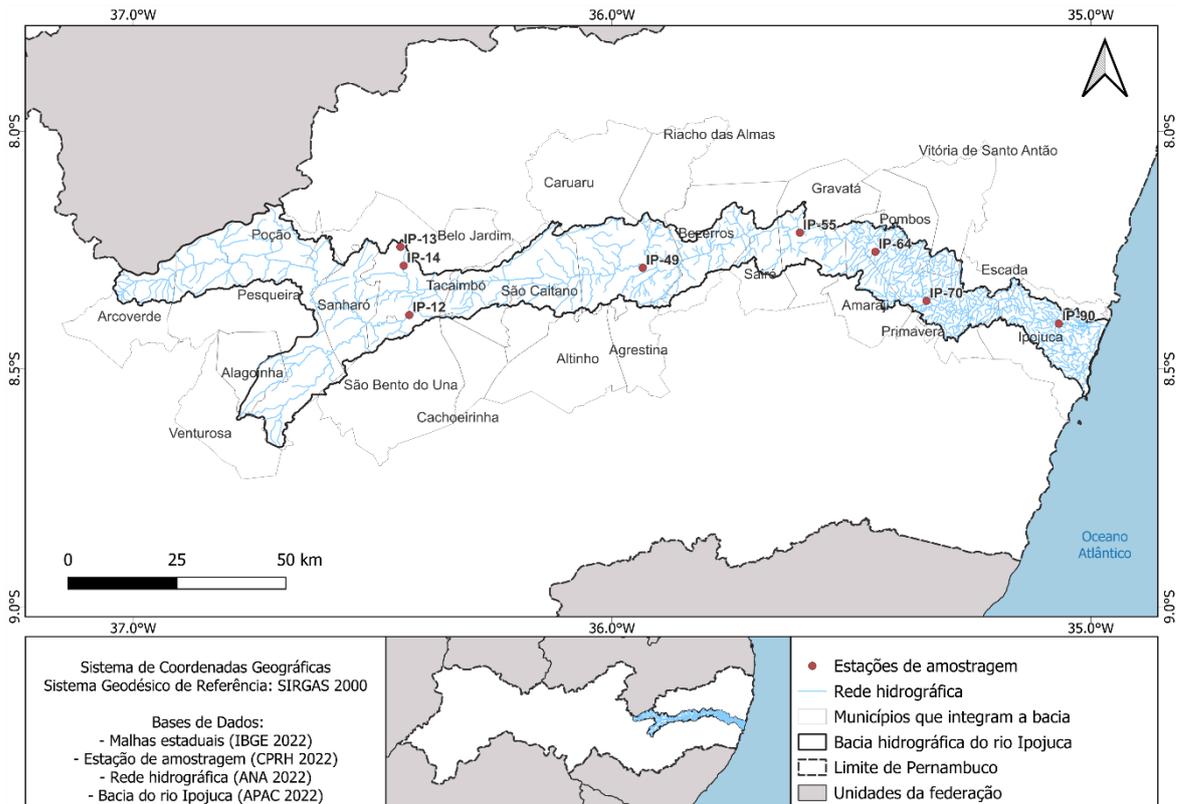
Fonte: Elaborado pela autora a partir dos dados de monitoramento disponibilizados pela CPRH (2025)

Quadro 12 - Estações de amostragem da rede de monitoramento da BHRI selecionadas para avaliação

Zona Homogênea	Estação	Corpo d'água	Local	Latitude	Longitude
Pecuária Leiteira	IP-12	Rio Ipojuca	Na ponte da PE-180, que liga Belo Jardim a São Bento do Una.	-8,3865	-36,4234
Interesse Ambiental - brejo de altitude	IP-13	Rio Bitury	Nascente no Sítio de Seu Joaquim, na comunidade de Jussara, em Belo Jardim.	-8,2432	-36,4418
Interesse Ambiental - brejo de altitude	IP-14	Rio Bitury	Na ponte sobre o Rio Bitury, à montante do reservatório, em Belo Jardim.	-8,2824	-36,4350
Urbana/Industrial - trecho intermitente	IP-49	Rio Ipojuca	A jusante da cidade de Caruaru, na Vila do Cedro (COAHB III), na cidade de Caruaru.	-8,2872	-35,9356
Pecuária de Corte	IP-55	Rio Ipojuca	Na ponte da BR-232, a montante de Gravatá, no município de Gravatá.	-8,2131	-35,6077
Policultura	IP-64	Rio Ipojuca	Na ponte à jusante da cidade de Chã Grande, no município de Chã Grande.	-8,2533	-35,4501
Agroindústria	IP-70	Rio Ipojuca	Na ponte à jusante da Usina União Indústria, no município de Primavera.	-8,3568	-35,3431
Agroindústria	IP-90	Rio Ipojuca	Na ponte PE-60 à jusante da Usina Ipojuca, no município de Ipojuca.	-8,4046	-35,0668

Fonte: CPRH (2020)

Figura 12 - Distribuição espacial das estações de amostragem da bacia do rio Ipojuca selecionadas para avaliação



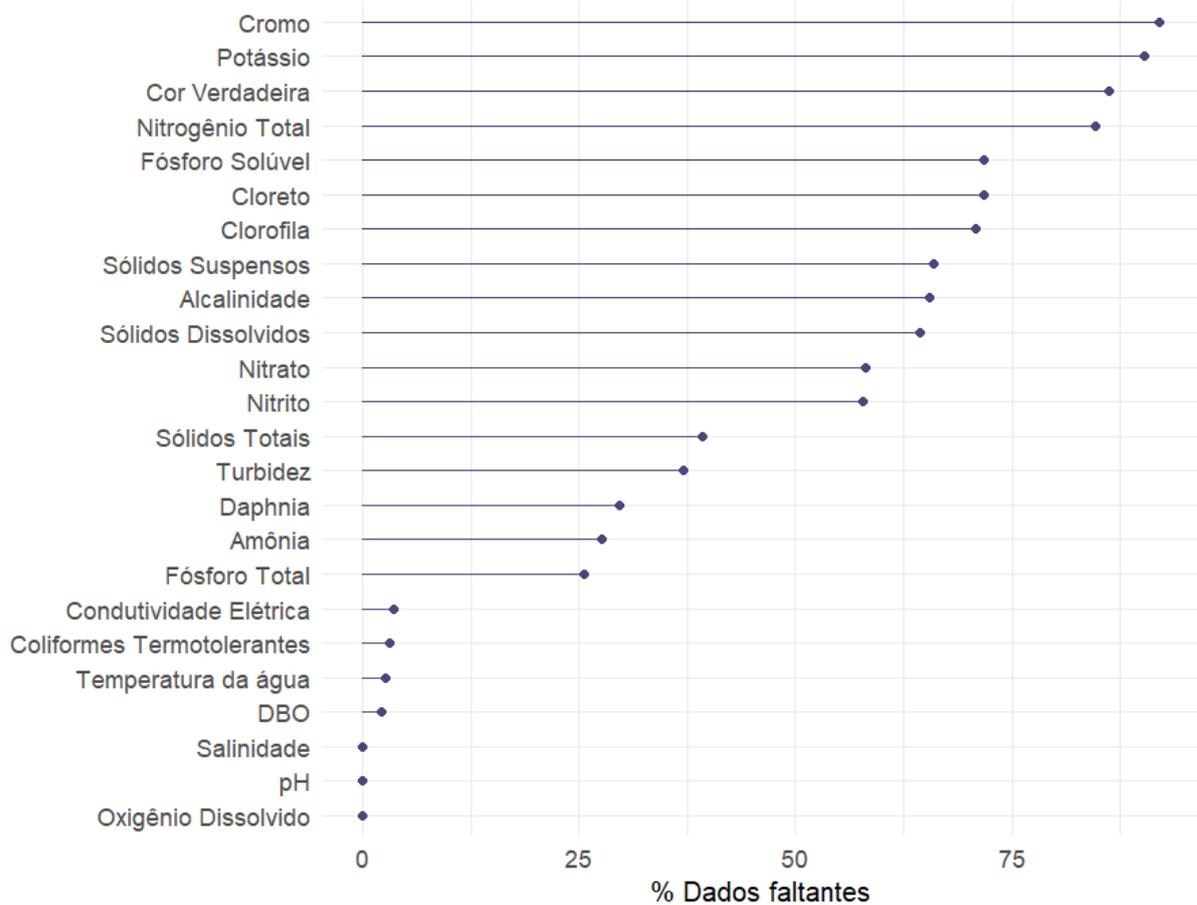
Fonte: Elaborado pela autora (2025)

Os dados brutos passaram por procedimentos iniciais de controle de qualidade para minimizar a ocorrência de valores faltantes. Parâmetros com mais de 50% de dados ausentes foram excluídos, pois não representavam adequadamente as variações na qualidade da água ao longo do período de monitoramento (Figura 13). Esse critério foi adotado para garantir que a análise estatística pudesse ser aplicada também a parâmetros que representam nutrientes responsáveis pela eutrofização, além de indicadores da série de sólidos.

Dessa forma, foram considerados apenas os seguintes parâmetros: temperatura da água, pH, condutividade elétrica, oxigênio dissolvido, demanda bioquímica de oxigênio, amônia, fósforo total, coliformes termotolerantes, salinidade, turbidez, *Daphnia* e sólidos totais. Os resultados foram então comparados com os limites estabelecidos para Águas Doces Classe 2, conforme a Resolução CONAMA nº 357/2005 (Tabela 2), permitindo identificar as estações que apresentam maior violação aos padrões de qualidade da água estabelecidos pela legislação. O percentual de violação foi calculado considerando a proporção de amostras que excederam

os limites permitidos em relação ao total de amostras analisadas em cada estação de amostragem para um determinado parâmetro.

Figura 13 - Porcentagem de dados faltantes por parâmetro no banco de dados bruto



Fonte: Elaborado pela autora a partir dos dados de monitoramento disponibilizados pela CPRH (2025)

Tabela 2 - Parâmetros de qualidade da água utilizados na pesquisa que possuem valores máximos permitidos de acordo com a Resolução CONAMA nº 357/2005

Parâmetros	Valor máximo permitido
Salinidade	0,50 ‰
Oxigênio dissolvido (OD)	≥ 5 mg/L O ₂
pH	6 a 9
Turbidez	100 UNT
Demanda bioquímica de oxigênio (DBO)	5 mg/L O ₂
Fósforo total	0,030 mg/L P em ambientes lênticos 0,050 mg/L P em ambientes intermediários 0,1 mg/L P em ambientes lóticos
Sólidos dissolvidos totais	500 mg/L
Coliformes termotolerantes	1.000 NMP/100 mL

Fonte: CONAMA (2005)

3.2.1.3 Análise estatística

Com o banco de dados definido, inicialmente foram calculadas as estatísticas descritivas dos dados de cada parâmetro para as estações da BHRI que foram selecionadas usando o pacote *rstatix* da linguagem R. Isso incluiu os valores mínimos e máximos, mediana, média e desvio padrão de cada parâmetro. Em seguida, a distribuição dos dados foi avaliada usando o teste de normalidade de Shapiro-Wilk com um nível de significância de 5% (SHAPIRO; WILK, 1965), com as seguintes hipóteses:

Hipótese nula (H_0): A amostra é proveniente de uma população normal;

Hipótese alternativa (H_1): A amostra não é proveniente de uma população normal.

Para examinar a variação espacial da qualidade da água na bacia do rio Ipojuca, foram empregados gráficos *boxplot*, também conhecidos como diagramas de caixa, para os dados de cada estação de amostragem. Esses gráficos, construídos usando o pacote *ggplot2* da linguagem R, oferecem uma representação visual que revela a distribuição de um conjunto de dados. Cada *boxplot* consiste em uma caixa que engloba o intervalo entre o primeiro quartil (Q1) e o terceiro quartil (Q3) dos dados, representando 25% e 75% das observações, respectivamente. No centro da caixa está a mediana. As "hastes" que se estendem para fora da caixa indicam a amplitude dos dados, excluindo os valores atípicos (*outliers*), que são pontos que se encontram além dos limites definidos pelas hastes, representados individualmente.

A interpretação de um *boxplot* envolve a observação da posição da mediana em relação à caixa (indicando a simetria dos dados), a extensão das hastes (sugerindo a dispersão dos dados) e a identificação de *outliers* (podendo indicar valores extremos ou discrepantes no conjunto de dados). Os quartis dividem as observações em quatro grupos de tamanhos iguais. Uma caixa pequena sugere que os dados estão altamente concentrados em torno da mediana, enquanto uma caixa grande indica maior heterogeneidade nos dados (CAPP; NIENOV, 2020).

Para avaliar se há diferença significativa entre os resultados dos parâmetros de qualidade da água, precipitação e vazão durante os períodos seco e chuvoso, foi aplicado o teste de Mann-Whitney (WILCOXON, 1945; MANN; WHITNEY, 1947), com nível de significância de 5%. Os parâmetros de qualidade da água foram analisados individualmente para cada estação amostral, enquanto os dados de precipitação e vazão foram agrupados com base nas quatro unidades de análise da bacia do rio Ipojuca.

Também conhecido como teste de Wilcoxon-Mann-Whitney, ou *Wilcoxon rank-sum test*, esse teste é utilizado para comparar dois grupos independentes de amostras, sob as seguintes hipóteses:

Hipótese nula (H_0): As distribuições das duas amostras são iguais;

Hipótese alternativa (H_1): As distribuições das duas amostras são diferentes.

Posteriormente, foram construídas matrizes de diagramas de dispersão separadamente para os períodos seco e chuvoso, incluindo a correlação de Spearman para avaliar a relação entre as variáveis de qualidade da água, precipitação e vazão. As matrizes foram construídas utilizando o pacote *psych* da linguagem R. O método de correlação de Spearman foi escolhido porque não assume nenhuma distribuição específica dos dados (SPEARMAN, 1904). O coeficiente de correlação varia entre os valores de -1 (correlação negativa perfeita) a +1 (correlação positiva perfeita), e quanto mais próximo de 1 for o valor absoluto do coeficiente, mais forte será a relação entre as variáveis.

Com o objetivo de identificar padrões espaço-temporais e determinar os parâmetros que exercem maior influência na qualidade da água da BHRI, foi utilizada a Análise de Componentes Principais (ACP). Os dados foram divididos em dois períodos: de 2005 a 2019 e de 2020 a 2023. Essa divisão foi necessária devido à descontinuidade na sequência de dados de alguns parâmetros, o que exigiria a imputação de um grande volume de dados para a aplicação da ACP, comprometendo a qualidade do resultado obtido. Enquanto fósforo total e amônia possuem registros apenas entre os anos de 2005 e 2019, sólidos totais e turbidez estão disponíveis apenas entre 2020 e 2023.

Para a aplicação da Análise de Componentes Principais, os dados devem ser padronizados para a mesma escala, a fim de evitar que algumas variáveis se tornem dominantes apenas devido às suas grandes unidades de medida. Nesta pesquisa, a ACP foi realizada por meio do pacote *FactoMineR*, da linguagem R, que executa automaticamente a padronização dos dados. Além disso, o pacote *factoextra* foi utilizado para gerar os gráficos *biplot* das variáveis com as amostras.

A seleção das componentes principais mais relevantes foi baseada no critério de reter as que explicam mais de 70% da variabilidade total dos dados (JOLLIFFE; CADIMA, 2016). Após essa seleção, foram consideradas, dentro dessas componentes, as variáveis que apresentaram coeficiente de correlação maior ou igual a $\pm 0,5$, garantindo uma significância prática (HAIR et al., 2009).

A análise de tendência nas séries temporais dos parâmetros foi realizada por meio do teste de Pettitt e do teste de Mann-Kendall. Para essa análise, foram utilizadas as médias anuais de cada parâmetro para cada estação de amostragem. Isso se justifica para que seja uniformizada a sequência dos dados ao longo dos anos, devido ao fato de o monitoramento ter variado entre campanhas bimestrais e trimestrais na maioria dos anos. Este procedimento também tem o efeito de eliminar a autocorrelação na série temporal, que é uma premissa para atender ao teste de Mann-Kendall, e reduzir efeitos de variação sazonal dos dados.

O teste de Pettitt (PETTITT, 1979) é um método não paramétrico, e foi utilizado para identificar possíveis pontos de mudança de tendência na série das médias anuais dos parâmetros de qualidade da água. Esse teste é baseado no teste de Mann-Whitney (baseado em *rankings*), e permite detectar uma mudança significativa em um momento desconhecido t , sob as seguintes hipóteses:

Hipótese nula (H_0): não há alteração na distribuição de uma sequência de variáveis aleatórias;

Hipótese alternativa (H_1): a função de distribuição $F_1(x)$ das variáveis aleatórias de X_1 a X_t é diferente da função de distribuição $F_2(x)$ das variáveis aleatórias de X_{t+1} a X_n .

Um ponto de mudança ocorre no tempo t quando a estatística K_t é significativamente diferente de zero para um nível pré-determinado. O nível de significância aproximado é calculado e, se o p -valor for inferior ao nível de significância pré-definido α , a hipótese nula é rejeitada. Assim, os dados podem ser divididos em duas subséries (antes e depois do ponto de mudança), cada uma com uma distribuição distinta. Nesta pesquisa, foi utilizado o nível de significância de 5%.

Em seguida, as tendências da qualidade da água nas estações de amostragem foram avaliadas utilizando o teste de Mann-Kendall (KENDALL, 1975; MANN, 1945), também não paramétrico. Esse teste é amplamente utilizado para avaliar estatisticamente dados ambientais e determinar se há uma tendência monotônica ascendente ou descendente ao longo do tempo, sob as seguintes hipóteses:

Hipótese nula (H_0): há ausência de tendência;

Hipótese alternativa (H_1): existe tendência significativa.

Com as estatísticas do teste calculadas, um valor negativo indica uma tendência decrescente e um valor positivo representa uma tendência ascendente. A significância estatística foi considerada quando $p < 0,05$. Séries temporais com valores ausentes foram

incluídas na análise, desde que não houvesse lacunas de dois ou mais anos consecutivos ou não consecutivos. Nos casos em que o teste de Pettitt identificou pontos de mudança na série temporal, testes individuais de Mann-Kendall foram aplicados para cada segmento antes e depois do ponto de mudança. As análises de tendência foram realizadas utilizando o pacote *trend* da linguagem R. Os gráficos de série temporal de todos os parâmetros avaliados, com base em todos os dados disponíveis (não agregados em médias anuais), estão apresentados no Apêndice A.

3.2.2 Relação entre o uso e cobertura do solo e a qualidade da água

A caracterização do uso e ocupação do solo na BHRI foi conduzida por meio dos dados disponíveis na Coleção 8 do Projeto de Mapeamento Anual da Cobertura e Uso da Terra no Brasil - MapBiomias. Esta coleção engloba mapas anuais de cobertura e uso do solo de todo o Brasil em uma escala de 30 metros, abrangendo o período entre os anos de 1985 e 2022, e está continuamente sendo aprimorada.

O MapBiomias é um projeto que congrega uma rede colaborativa de especialistas nas áreas de Sensoriamento Remoto, Ciência da Computação, Ciência de Dados, usos da terra e biomas. Todos os mapas são gerados por meio da classificação pixel a pixel de imagens dos satélites Landsat. Esse processo é conduzido utilizando extensivos algoritmos de Aprendizado de Máquina por meio da plataforma *Google Earth Engine*, que proporciona uma vasta capacidade de processamento na nuvem (MAPBIOMAS, 2024). Nesta pesquisa, foram utilizados mapas do uso e cobertura do solo referentes aos anos de 2005 e 2019.

Com o uso do *software* de código aberto QGIS, as imagens foram reprojatadas para o sistema de coordenadas planas Universal Transversa de Mercator (UTM), utilizando o DATUM oficial do Brasil, o SIRGAS2000, na zona 24 Sul. Em seguida, as imagens foram recortadas para se adequarem à área da BHRI. Posteriormente, foi realizada a reclassificação do uso do solo. Isso se faz necessário devido aos arquivos gerados pelo MapBiomias apresentarem níveis de classes muito específicos, os quais não são relevantes para este estudo. Assim, o uso e cobertura do solo na bacia do rio Ipojuca foi agrupado nas seguintes classes: formações florestais, formações savânicas, formações naturais não florestais, agropecuária, áreas de cultivo de cana-de-açúcar, áreas urbanizadas e corpos hídricos. Após a reclassificação, foram calculadas as áreas e percentuais de cada classe para os anos avaliados em relação à área total da bacia e em relação à área de um raio de 500 metros dos pontos de amostragem de qualidade da água.

O percentual de classes de uso do solo em um raio de 500 metros ao redor de cada estação de amostragem foi utilizado na análise estatística para avaliar os efeitos dos padrões de uso do solo sobre os parâmetros de qualidade da água. Análises de correlação foram realizadas para identificar os tipos de uso do solo que impactam significativamente esses parâmetros. Para essa análise, foram utilizados coeficientes de correlação de Spearman com p-valor inferior a 0,05.

Foi realizada uma Análise de *Cluster* (ou Análise de Agrupamentos) para classificar as estações de amostragem com base nos padrões de uso do solo no raio de 500 metros ao seu redor, bem como nos parâmetros de qualidade da água nos períodos seco e chuvoso. A Análise de Cluster é um método utilizado para agrupar dados com base em sua similaridade (HAIR et al., 2009). Para a formação da matriz de similaridade, adotou-se a distância euclidiana quadrada (*Squared Euclidean Distance*), enquanto o método de *Ward* foi empregado para definir os agrupamentos no dendrograma. As médias dos anos de 2005 e 2019 de todos os parâmetros de qualidade da água foram incluídos na análise, após padronização para eliminar o efeito de escala e unidades. Os tipos de uso do solo foram selecionados com base nos resultados da análise de correlação, a fim de minimizar a influência de fatores não críticos na clusterização. O pacote *factoextra* foi usado para gerar a visualização do dendrograma.

A relação entre padrões de uso do solo e indicadores de qualidade da água nos períodos seco e chuvoso foi examinada por meio da Análise de Redundância (RDA), usando o pacote *vegan* da linguagem R. A RDA é uma técnica estatística multivariada utilizada para explorar a relação entre dois conjuntos de variáveis: variáveis de resposta (qualidade da água) e variáveis preditoras (uso do solo). O principal propósito da RDA é determinar as combinações lineares das variáveis preditoras que explicam a máxima quantidade de variação nas variáveis de resposta. Isso ajuda a identificar quais variáveis preditoras estão mais fortemente associadas às variáveis de resposta e a visualizar as relações entre os dois conjuntos de variáveis (SONG et al., 2020; CHEN et al., 2021).

Nesta pesquisa, a RDA foi aplicada inicialmente às duas matrizes de dados completas, ou seja, todas as classes de uso do solo e todos os parâmetros de qualidade da água. Após os resultados do modelo completo terem sido obtidos, foi realizada a seleção pelo método de seleção progressiva (*forward selection*) das melhores variáveis explicativas para simplificar o modelo. Nesse método, o processo se inicia com um conjunto vazio e, a cada iteração, uma nova variável (classe de uso do solo) é adicionada. A estatística da RDA é representada pelo valor de R^2 e F. O valor de R^2 indica a força da relação linear entre variáveis de resposta e preditoras e o valor do F representa o teste global de significância.

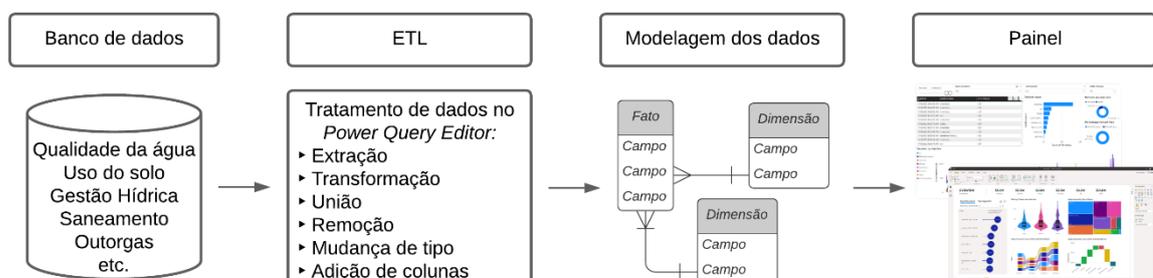
As médias anuais dos parâmetros de qualidade da água foram padronizadas utilizando transformação logarítmica antes de realizar as análises. O teste de permutação de Monte Carlo foi aplicado para medir a significância do efeito dos tipos de uso do solo na qualidade da água (JONGMAN; JONGMAN, 1995). Os gráficos de ordenação (*triplots*) foram gerados para os modelos completos apresentando os resultados da análise RDA, onde podem ser observadas as classes de uso do solo, os parâmetros de qualidade da água e as amostras individuais.

3.2.3 Desenvolvimento de painel de visualização da qualidade da água

A análise dos dados do monitoramento da qualidade da água é essencial na gestão dos recursos hídricos. Ela ajuda a entender os processos que impactam a qualidade da água, possibilitando o planejamento de melhorias, a identificação de áreas críticas de degradação, fornecendo suporte para fiscalização e formulação de políticas ambientais. Para desenvolver o painel de visualização dos dados do monitoramento da BHRI foi utilizado o *software Power BI*, que é uma ferramenta de *Business Intelligence*.

O *Power BI* é uma solução de análise de negócios em nuvem desenvolvida pela Microsoft que facilita a integração de diversas fontes de dados, permitindo análises detalhadas e apresentação visual por meio de relatórios, painéis e visualizações. Entre suas vantagens, destacam-se a inclusão de uma ampla variedade de gráficos e modelos de relatórios pré-definidos, além do recente acréscimo de recursos de Aprendizado de Máquina. A ferramenta está disponível tanto em *desktops* quanto em dispositivos móveis, e sua relação custo-benefício é considerada mais vantajosa em comparação com outras opções do mercado (MICROSOFT, 2024). Neste trabalho foi utilizada a versão *desktop*, que é gratuita. Para a construção do painel, foram realizadas quatro etapas, conforme indicado na Figura 14.

Figura 14 - Etapas da elaboração do painel de visualização da qualidade da água



Fonte: Elaborado pela autora (2025)

O objetivo é organizar e programar a interação entre os dados, permitindo a automatização de sua configuração e simplificando sua futura atualização e manutenção. A

primeira etapa consiste na coleta e definição do banco de dados utilizado, composto pelos seguintes conjuntos:

- Qualidade da água
- Uso do solo
- Gestão de recursos hídricos na bacia
- Características populacionais
- Indicadores de saneamento
- Estações de tratamento de esgotos
- Outorgas

Foi utilizado o conjunto completo de dados brutos de qualidade da água fornecidos pela CPRH, organizados em planilhas Excel e abrangendo os anos de 2005 a 2023. Esse conjunto inclui todas as estações e os parâmetros que não foram analisados em seções anteriores desta pesquisa devido à grande quantidade de dados faltantes. A escolha desse conjunto se justifica pela natureza do painel como uma ferramenta de visualização, que permite demonstrar as mudanças ocorridas no sistema de monitoramento ao longo do tempo, incluindo adições e remoções de parâmetros e estações.

Com relação ao uso do solo, foram obtidas da Plataforma Ecológica-Econômica de Pernambuco as séries históricas referentes a toda a bacia do rio Ipojuca e aos municípios que a compõem. Essa plataforma utiliza dados do MapBiomas para todo o estado, cobrindo o período de 1985 a 2023. Além disso, foram incluídos os dados do raio de 500 metros ao redor das estações de amostragem, utilizados na Análise de Redundância.

Foram também compiladas informações sobre eventos relevantes relacionados à gestão de recursos hídricos na BHRI, permitindo cruzar esses eventos com os dados de qualidade da água em uma linha do tempo. Além disso, foram extraídas características populacionais e indicadores de saneamento dos municípios da bacia a partir das bases de dados do Sistema Nacional de Informações sobre Saneamento (SNIS), enquanto as informações sobre as estações de tratamento de esgotos foram compiladas do estudo de Soares (2023). Os dados de outorgas foram retirados do banco de dados da ANA.

A segunda etapa é denominada ETL (do termo em inglês *Extraction, Transformation and Load*), e consiste no processo de extrair, transformar e carregar os dados no *Power Query Editor* do *Power BI*. É um processo fundamental para a preparação dos dados antes de sua análise e visualização. A etapa de extração envolve a obtenção dos dados de diversas fontes, como bancos de dados, arquivos locais ou na nuvem. Em seguida, na fase de transformação, os

dados são manipulados e limpos para garantir sua qualidade e consistência, incluindo a remoção de duplicatas, a correção de erros e a padronização de formatos. Por fim, na etapa de carga, os dados preparados são inseridos no modelo de dados do *Power BI* para serem utilizados na criação de relatórios e visualizações.

Com as tabelas de dados já formatadas, foi iniciada a etapa de modelagem dos dados, fundamental para o desenvolvimento de modelos semânticos no *Power BI* otimizados para desempenho e usabilidade. Nessa fase, são estabelecidos os relacionamentos entre as diferentes tabelas de acordo com um Esquema em Estrela, que organiza os dados em uma tabela de fatos (tabela principal) e tabelas de dimensões (tabelas secundárias).

O planejamento dessa modelagem é essencial, pois a relação entre as tabelas determina como os filtros aplicados a uma coluna podem ser propagados para outras tabelas no modelo. Essa propagação ocorre desde que exista um caminho de relacionamento estabelecido, podendo envolver múltiplas tabelas no processo.

A última etapa do processo corresponde à criação do painel de visualizações interativas, no qual os dados podem ser filtrados com base em critérios específicos, como datas ou localizações. Essa funcionalidade permite ao usuário acessar diferentes indicadores e informações de forma gráfica e tabular. Sempre que um filtro é aplicado, os resultados são atualizados automaticamente em cada gráfico, garantindo uma análise dinâmica e eficiente.

4 RESULTADOS E DISCUSSÃO

Neste capítulo, são apresentados os resultados obtidos com a aplicação da metodologia descrita anteriormente, começando pela avaliação da evolução espaço-temporal da qualidade da água na BHRI. Em seguida, são expostos os produtos da análise da relação entre o uso e cobertura do solo e a qualidade da água. Por fim, apresentam-se os resultados do desenvolvimento do painel de visualização da qualidade da água.

4.1 EVOLUÇÃO ESPAÇO-TEMPORAL DA QUALIDADE DA ÁGUA NA BACIA DO RIO IPOJUCA

4.1.1 Padrões espaciais de qualidade da água

Para avaliar a evolução espaço-temporal da qualidade da água na BHRI, foram analisadas as concentrações de 12 parâmetros físico-químicos e biológicos resultantes de coletas realizadas entre 2005 e 2023 em oito estações de monitoramento.

Os resultados do teste de Shapiro-Wilk, realizado com um nível de significância de 5% e apresentados na Tabela 3, indicaram a rejeição da hipótese nula para todos os parâmetros analisados. Isso significa que nenhum deles segue uma distribuição normal.

Tabela 3 - Resultados da estatística W e valores de p do teste de Shapiro-Wilk

Parâmetro	Estatística W	p-valor
OD	0,943	<0,001
DBO	0,554	<0,001
Fósforo total	0,588	<0,001
Amônia	0,284	<0,001
pH	0,944	<0,001
Temperatura da água	0,984	<0,001
Salinidade	0,565	<0,001
Turbidez	0,607	<0,001
Sólidos dissolvidos totais	0,577	<0,001
Condutividade elétrica	0,612	<0,001
Coliformes termotolerantes	0,664	<0,001
<i>Daphnia</i>	0,041	<0,001

Fonte: Elaborado pela autora (2025)

O resumo das concentrações dos parâmetros, apresentado por meio de estatísticas descritivas, é mostrado na Tabela 4, enquanto os *boxplots* são exibidos na Figura 15. Os dados foram comparados com os limites estabelecidos pela Resolução CONAMA nº 357/2005 para águas doces classe 2.

Tabela 4 - Resumo dos parâmetros de qualidade da água analisados na BHRI de 2005 a 2023

Parâmetro	Unidade	Mínimo	Máximo	Mediana	Média	Desvio Padrão	CONAMA 357/2005*
OD	(mg O ₂ /L)	0	11,9	4,2	4,2	2,74	≥ 5
DBO	(mg O ₂ /L)	0,36	211	4,1	13,2	23,26	< 5
Fósforo total	(mg P/L)	0,001	9,57	0,33	0,8	1,36	< 0,1
Amônia	(mg NH ₃ /L)	0,07	83	0,13	2,7	9,72	-
pH	-	4,5	9,2	7,18	7,0	0,71	6 a 9
Temperatura da água	(°C)	16	34	26	26,0	2,85	-
Salinidade	‰	0,04	6,60	0,2	0,5	0,68	< 0,5
Turbidez	(UNT)	0,01	300	20	29,9	38,47	< 100
Sólidos dissolvidos totais	(mg/L)	42	8111	219,5	582,7	844,99	< 500
Condutividade elétrica	(μS/cm)	61,8	11680	440,5	975,2	1299,59	-
Coliformes termotolerantes	(NMP/100mL)	20	160000	8000	45206,3	63550,48	< 1000
<i>Daphnia</i>	FDd	1	128	1	1,5	7,44	-

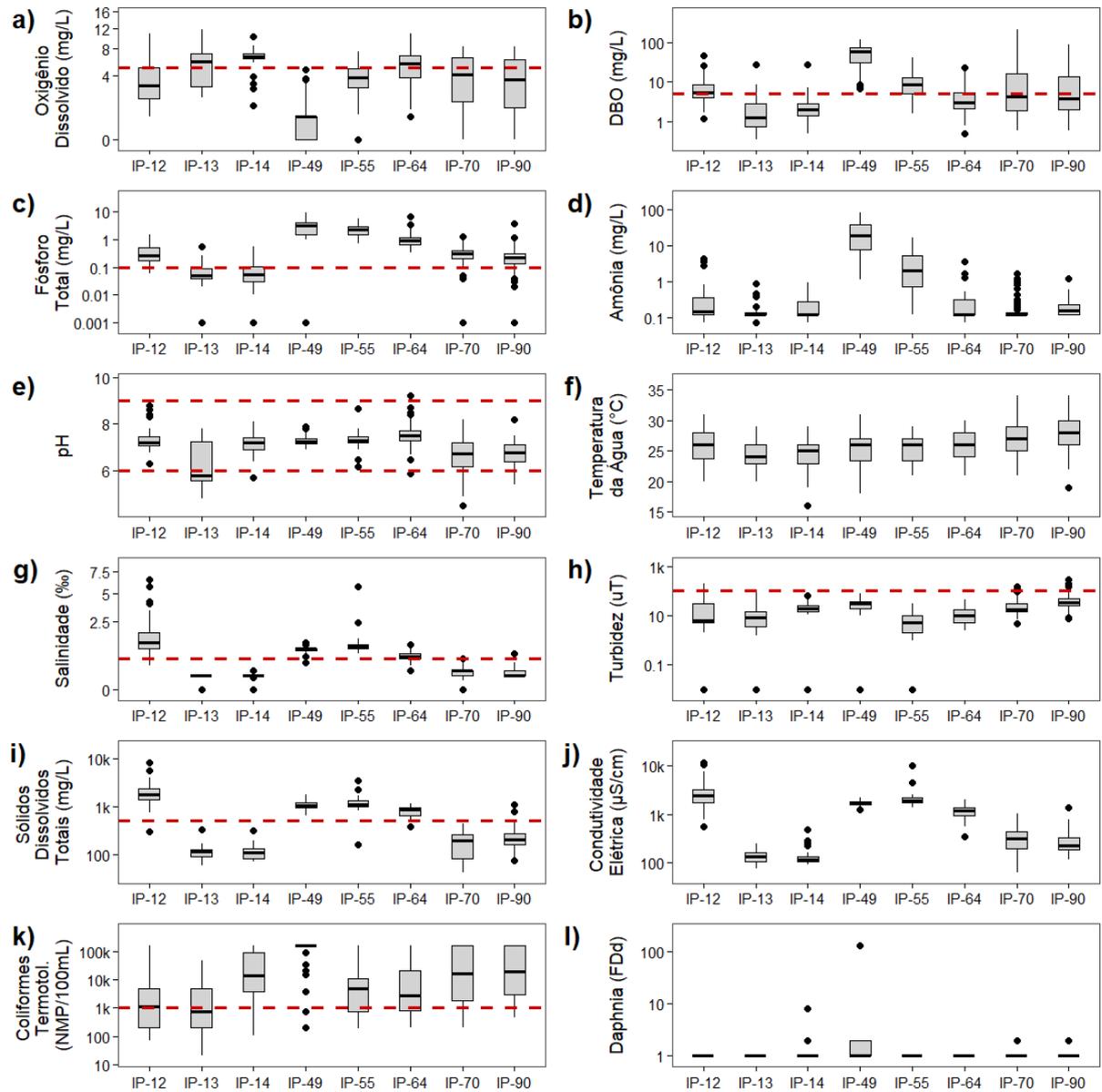
Fonte: Elaborado pela autora (2025)

*Águas doces Classe 2.

Conforme mostrado na Tabela 4, a ampla variação observada na maioria dos parâmetros resulta em valores elevados de desvio padrão e em uma distribuição assimétrica. Essa assimetria, comumente encontrada em dados ambientais (CALAZANS et al., 2018), foi confirmada pelos resultados do teste de Shapiro-Wilk.

Os níveis de oxigênio dissolvido variaram de 0 mg O₂/L a 11,9 mg O₂/L, com mediana e média iguais a 4,2 mg O₂/L, estando abaixo do limite mínimo estabelecido pelo CONAMA (≥ 5 mg O₂/L). Pode-se notar na Figura 15a que o OD é um parâmetro crítico na bacia do Ipojuca, com todas as estações apresentando valores abaixo do recomendado pela legislação. Os pontos IP-12 e IP-55 se destacam com mais de 75% das amostras coletadas estando abaixo de 5 mg O₂/L, enquanto no IP-49 isso ocorreu em 100% das amostras.

Figura 15 - Variação dos parâmetros de qualidade da água nas estações de monitoramento da BHRI entre 2005 e 2023



Fonte: Elaborado pela autora (2025)

Nota: Escala do eixo Y ajustada para melhor visualização: raiz quadrada em a) e g), logarítmica (base 10) nos demais, exceto e) e f).

Legenda: IP-12: Rio Ipojuca, na ponte da PE-180, que liga Belo Jardim e São Bento do Una; IP-13: Nascente do Rio Bitury em Belo Jardim; IP-14: Rio Bitury, à montante do reservatório, em Belo Jardim; IP-49: Rio Ipojuca, à jusante de Caruaru; IP-55: Rio Ipojuca, a montante de Gravatá; IP-64: Rio Ipojuca, à jusante de Chã Grande; IP-70: Rio Ipojuca, à jusante da Usina União Indústria, em Primavera; IP-90: Rio Ipojuca, à jusante da Usina Ipojuca, em Ipojuca.

As condições de oxigenação significativamente baixas das águas da bacia do rio Ipojuca vêm sendo diretamente influenciadas pelos aportes de carga orgânica, especialmente provenientes de esgotos não tratados de áreas urbanas e de efluentes industriais, incluindo os de usinas de cana-de-açúcar. Em outras palavras, vastas extensões da bacia têm sido afetadas há muitos anos pelo lançamento contínuo de resíduos orgânicos.

Essa situação se torna particularmente evidente nas estações localizadas nos centros urbanos ao longo do rio Ipojuca, onde há uma entrada expressiva de matéria orgânica proveniente do descarte de esgotos domésticos, como observado nas estações IP-49 (Caruaru), IP-55 (Gravatá) e IP-64 (Chã Grande), e/ou de efluentes industriais, como nas estações IP-70 (Primavera) e IP-90 (Ipojuca), situadas imediatamente a jusante de usinas de cana-de-açúcar.

A demanda bioquímica de oxigênio variou de 0,36 mg O₂/L a 211 mg O₂/L (Figura 15b), excedendo o limite estabelecido pela Resolução CONAMA nº 357/2005 (5 mg O₂/L) em todas as estações avaliadas. As maiores concentrações foram observadas nos pontos IP-49, IP-70 e IP-90.

O aumento da DBO em um corpo d'água geralmente resulta do lançamento de resíduos predominantemente orgânicos. A presença de uma quantidade significativa de matéria orgânica pode levar à completa ausência de oxigênio dissolvido na água, resultando na mortalidade de peixes e outras formas de vida aquática (NACAR; METE; BAYRAM, 2020). Além disso, um alto valor de DBO pode indicar um aumento na população microbiana presente na água, perturbando o equilíbrio do ecossistema aquático.

Outro impacto relevante está relacionado à qualidade da água utilizada para consumo humano. Concentrações elevadas de DBO podem causar sabores e odores desagradáveis, além de comprometer a eficiência dos processos usados nas estações de tratamento de água pela obstrução dos filtros de areia (CETESB, 2023).

As concentrações de fósforo total variaram entre 0,001 mg P/L e 9,57 mg P/L (Figura 15c), com todas as estações de monitoramento excedendo o limite de 0,1 mg P/L estabelecido pela Resolução CONAMA nº 357/2005. Esses resultados evidenciam a influência das fontes poluidoras ao longo da BHRI.

Em particular, destacaram-se as estações IP-49, IP-55, IP-64, IP-70 e IP-90, que refletem o aporte contínuo de efluentes provenientes de esgotos domésticos e de atividades industriais ao rio Ipojuca, conforme mencionado anteriormente.

Além disso, os resultados da estação IP-12, localizada em uma zona de pecuária leiteira, evidenciam a potencial influência das atividades rurais na qualidade da água, como a dessedentação de animais e a presença de fauna silvestre, que também podem intensificar a carga poluente difusa na bacia.

A concentração de amônia apresentou variação entre 0,07 mg NH₃/L e 83 mg NH₃/L (Figura 15d), com os valores mais elevados registrados na estação IP-49, localizada na cidade de Caruaru. É importante ressaltar que a amônia é um composto tóxico que impõe restrições significativas à vida dos peixes, sendo letal para muitas espécies em concentrações acima de 5

mg/L. Além disso, sua presença nas águas naturais intensifica o consumo de oxigênio dissolvido devido aos processos de oxidação biológica (CETESB, 2023). Esse fenômeno pode contribuir para a redução dos níveis de oxigênio na água, afetando negativamente a fauna aquática e a qualidade geral do ecossistema.

O pH variou de 4,5 a 9,2, com uma mediana de 7,18 e uma média de 7,0 (Figura 15e). Durante o período avaliado, o pH ultrapassou o valor máximo recomendado pela Resolução CONAMA nº 357/2005 (9,0) apenas uma vez, na estação IP-64, durante a coleta realizada em outubro de 2009, quando foi registrado o valor de 9,2.

Entretanto, em diversas ocasiões, esse parâmetro ficou abaixo do limite mínimo estabelecido pela resolução (6,0), especialmente nas estações IP-13, IP-70 e IP-90. As possíveis causas para a acidificação da água nesses pontos são distintas. As estações IP-70 e IP-90 estão situadas a jusante de usinas de produção de cana-de-açúcar, indicando a influência de efluentes industriais na alteração do pH do rio. Os estudos de Gunkel et al. (2007), no rio Ipojuca, e Silva et al. (2023), na bacia do rio Ivinhema, também reportaram acidificação das águas em áreas próximas às lavouras de cana-de-açúcar.

Já na estação IP-13, localizada na nascente do rio Bitury, os valores baixos de pH podem estar relacionados às características geoquímicas do solo da região. Essa tendência foi observada por Agrizzi et al. (2018), que relataram que, na ausência de fontes externas de poluição, a acidez da água das nascentes pode estar relacionada ao tipo de solo, especialmente em áreas com solos naturalmente ácidos, e à influência de águas subterrâneas, que tendem a apresentar pH mais baixo do que as superficiais.

A temperatura da água variou de 16°C a 34°C, com mediana e média iguais a 26°C. A Figura 15f mostra que as temperaturas nas estações IP-70 e IP-90, localizadas nos trechos mais à jusante na bacia, apresentam valores mais altos que as observadas nos trechos à montante. Gunkel et al. (2007) também registraram um aumento da temperatura da água na região, associando esse efeito à proximidade do rio Ipojuca com áreas de cultivo de cana-de-açúcar, onde a vinhaça era continuamente utilizada para fertirrigação.

A salinidade nas oito estações de amostragem variou entre 0,04‰ e 6,60‰ (Figura 15g), ultrapassando o limite de 0,5‰ estabelecido pela Resolução CONAMA nº 357/2005 para águas doces classe 2 na maioria dos pontos, com exceção das estações IP-13, IP-14 e IP-70. As estações IP-12, IP-49, IP-55 e IP-64 apresentaram medianas acima desse limite, indicando características mais próximas de águas salobras. Esses resultados evidenciam um déficit hídrico nessas partes do rio, onde prevalece o trecho intermitente da bacia, resultando em uma maior concentração de íons na água. Essa salinidade pode comprometer usos da água para consumo

humano e para irrigação, por exemplo. A salinização do solo, resultante da irrigação com água de alta salinidade, pode reduzir a produtividade agrícola, dificultar o crescimento das culturas e alterar as propriedades físicas e químicas do solo, além de contribuir para a contaminação das águas subterrâneas (MOHANAVELU et al., 2021).

A turbidez variou de 0,01 UNT a 300 UNT ao longo das estações monitoradas (Figura 15h), todas apresentando medianas abaixo do limite máximo estabelecido pela legislação (100 UNT). Foram observados apenas casos isolados de violação da resolução em algumas estações, como IP-12, IP-70 e IP-90, conforme demonstrado pelos *outliers* nos *boxplots* da Figura 15h. A alta turbidez pode resultar na redução da fotossíntese de vegetações aquáticas submersas e algas. Esse impacto na base da cadeia alimentar pode afetar negativamente a produtividade de peixes e outros organismos aquáticos. Além disso, altos níveis de turbidez dificultam os processos de tratamento da água para abastecimento público, aumentando os custos operacionais das estações de tratamento. Do ponto de vista ambiental e recreacional, águas altamente turvas também podem comprometer a estética e a qualidade para usos recreativos e industriais (CETESB, 2023).

A concentração de sólidos dissolvidos totais excedeu o limite de 500 mg/L (Figura 15i), estabelecido pela Resolução CONAMA nº 357/2005, em diversas estações de monitoramento ao longo da BHRI, especificamente nas estações IP-12, IP-49, IP-55, IP-64 e IP-90, com concentrações variando de 42 mg/L a 8111 mg/L. Esse era um resultado esperado, uma vez que essas estações apresentam características de águas salobras, especialmente IP-12, IP-49, IP-55 e IP-64.

Assim como os sólidos dissolvidos totais, a condutividade elétrica está diretamente relacionada à salinidade, variando entre 61,8 $\mu\text{S}/\text{cm}$ e 11680 $\mu\text{S}/\text{cm}$ (Figura 15j). A tendência observada para a condutividade segue a mesma dos sólidos dissolvidos, com valores mais elevados registrados nas estações com maior salinidade.

Foram detectadas amostras com concentrações de coliformes termotolerantes acima do limite máximo recomendado pela Resolução CONAMA nº 357/2005 para águas doces Classe 2 (1000 NMP/100 mL) em todas as estações de monitoramento ao longo da bacia do rio Ipojuca, na maioria das amostras analisadas (Figura 15k). As concentrações variaram entre 20 NMP/100 mL e 160.000 NMP/100 mL.

A presença desses coliformes indica contaminação recente por fezes de animais de sangue quente, além de sugerir a possível presença de bactérias e vírus patogênicos associados aos resíduos fecais, evidenciando o lançamento de esgoto doméstico no rio (VON SPERLING, 2005). Esse resultado destaca a necessidade urgente de implementação de medidas para reduzir

a contaminação fecal nos recursos hídricos da BHRI, visando à proteção da saúde pública e da qualidade ambiental.

Em um estudo realizado na bacia do rio Capibaribe, Silva et al. (2025) identificaram a presença de enterobactérias na água do rio, indicando contaminação de origem fecal. Os autores alertaram para os riscos à saúde pública, pois as águas do rio Capibaribe podem atuar como vetores na disseminação de bactérias resistentes, afetando tanto os seres humanos quanto o meio ambiente. Foram reportadas concentrações de coliformes totais e *Escherichia coli* que excederam os limites estabelecidos pela legislação em pelo menos 1350 vezes, independentemente da diluição. Esses resultados indicam uma contaminação fecal contínua, provavelmente associada a descargas de esgoto não tratado ou tratado de forma inadequada no rio.

O monitoramento da qualidade da água na BHRI inclui o controle de agentes tóxicos por meio de testes de ecotoxicidade. Esses testes envolvem a exposição de organismos a diferentes concentrações de substâncias químicas ou a diferentes diluições dessas substâncias durante um período de exposição predefinido. Neste estudo, foram avaliados os resultados de toxicidade aguda usando o microcrustáceo *Daphnia magna*, cujos resultados são expressos em FDd (Fator de Diluição no qual não há lesão ou morte).

Os resultados obtidos variaram de 1 FDd a 128 FDd na BHRI (Figura 151). Essa variação reflete a diversidade de concentrações da substância testada na água em relação ao grupo controle, onde a substância não estava presente. Um resultado de 1 FDd indica que a concentração da substância na água de teste não teve efeito adverso nos organismos testados em comparação com as condições de controle, ou seja, a amostra foi considerada não tóxica. Por outro lado, um resultado de 128 FDd indica uma concentração 128 vezes superior ao limite máximo tolerado pelos organismos em relação ao controle, sugerindo uma forte toxicidade da substância para os organismos aquáticos testados (MENDONÇA, 2005).

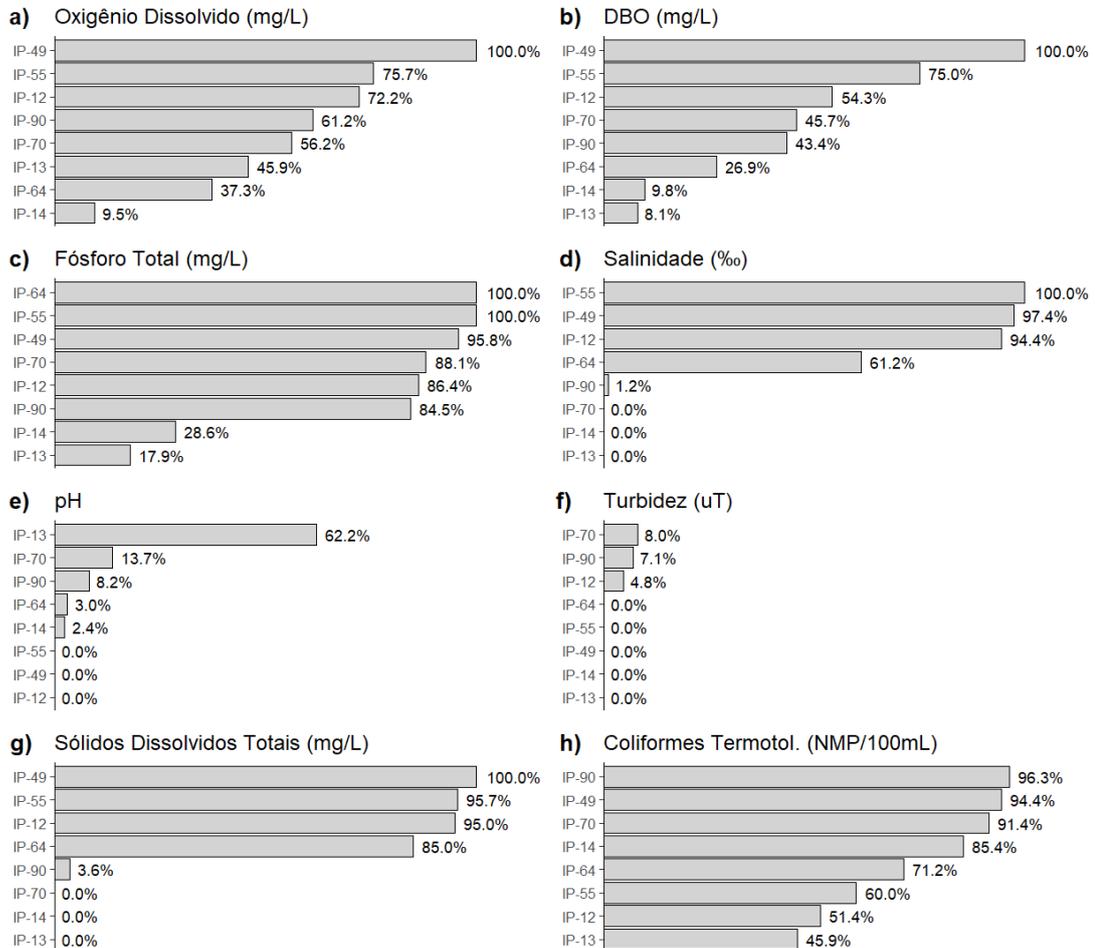
A maioria das amostras não apresentou toxicidade, com exceção das estações IP-14, IP-49, IP-70 e IP-90. Destaca-se o caso de IP-49 (Caruaru), que apresentou a maior incidência de amostras tóxicas, incluindo o valor máximo encontrado na bacia (128 FDd).

Com relação ao município de Caruaru, vale destacar que nele está localizado o Polo Têxtil de Pernambuco. De acordo com Ayed et al. (2021), os efluentes têxteis poluídos com corantes são geralmente considerados altamente tóxicos para a biota aquática e conhecidos por afetar significativamente a atividade fotossintética, resultando em perturbação do equilíbrio ecológico. No caso da *Daphnia magna*, o corante vermelho básico causa inibição da reprodução, de acordo com Koroglu et al. (2019).

O efluente da maioria das operações de tingimento têxtil possui uma cor marrom-avermelhada escura, que é esteticamente desagradável quando despejado em águas receptoras. Embora existam diversos métodos de remoção de cor, nenhum deles funciona em todos os casos. Além dos problemas já citados, a cor nos efluentes impede a penetração da luz na água, causando e acelerando o processo de eutrofização, o que resulta no consumo de oxigênio dissolvido e dificulta a reoxigenação da água. Dessa forma, águas residuais têxteis contendo corantes devem ser tratadas antes de sua descarga no meio ambiente (ROY et al., 2018).

A Figura 16 mostra que os parâmetros que tiveram violações à legislação em todas as estações foram precisamente os indicadores de poluição: OD (Figura 16a), DBO (Figura 16b), fósforo total (Figura 16c) e coliformes termotolerantes (Figura 16h). Levando em conta a soma das percentagens de violação apenas desses quatro parâmetros, a estação de amostragem com a pior qualidade da água foi a IP-49, seguida das estações IP-55, IP-90, IP-70, IP-12, IP-64, IP-14 e IP-13. Dessa forma, fica evidente que a qualidade da água na BHRI está mais comprometida nas proximidades das áreas de intensa intervenção antrópica, como regiões com maior densidade populacional e zonas afetadas por atividades agrícolas e industriais.

Figura 16 - Percentuais de violação da Resolução CONAMA n° 357/2005 para águas doces Classe 2 nas estações de amostragem analisadas



Fonte: Elaborado pela autora (2025)

No caso específico dos nutrientes, a densidade populacional foi apontada como o principal fator contribuinte para o aumento do nitrogênio na água superficial em seis bacias hidrográficas do litoral de Pernambuco (incluindo a BHRI) em uma série de dados analisada entre 1990 e 2016 no estudo de Noriega et al. (2021). Dentro desse fator antropogênico, os autores incluíram as contribuições das águas residuais, do escoamento urbano e da indústria.

A literatura aponta amplamente os impactos negativos do lançamento de efluentes não tratados nos rios sobre a qualidade da água (PAK et al., 2021; PEREIRA et al., 2021; SILVA et al., 2021). Esses estudos apontam que o aporte de matéria orgânica e nutrientes contribui para a redução do oxigênio dissolvido e para o desequilíbrio do ecossistema, causando a morte de peixes e estimulando o crescimento excessivo do fitoplâncton. Dessa forma, os resultados obtidos para o rio Ipojuca estão alinhados com padrões observados em outros sistemas fluviais impactados pela poluição orgânica.

A ausência de infraestrutura adequada de saneamento básico nos municípios da bacia reforça essa tendência, como evidenciado pelos dados do Sistema Nacional de Informações sobre Saneamento (SNIS, 2024). A maior cobertura de coleta de esgoto entre os municípios da bacia do Ipojuca é de apenas 51,69%, registrada em Tacaimbó. A cidade de Caruaru, que apresentou a pior qualidade da água entre as estações analisadas, possui a terceira maior cobertura, com apenas 40,87%, evidenciando a necessidade urgente de melhorias nesses índices. Uma alternativa para isso seria o aprimoramento das tecnologias de tratamento de esgoto para o nível terciário, que poderia reduzir significativamente as concentrações de efluentes provenientes de fontes pontuais nesses corpos d'água, conforme também verificado por Wang et al. (2015).

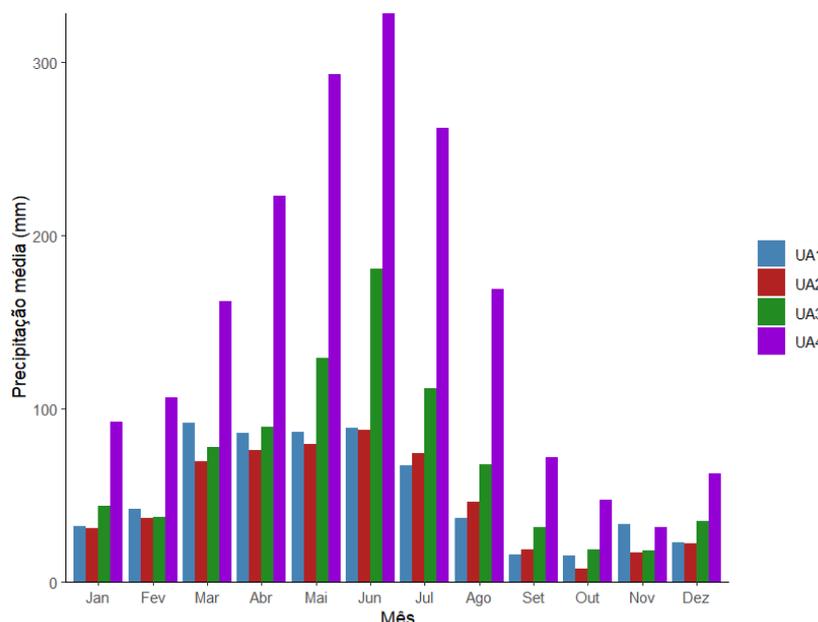
4.1.2 Influência dos períodos seco e chuvoso

A análise dos dados de precipitação e vazão na BHRI entre os anos de 2005 e 2023 revelou padrões distintos ao longo do período, refletindo a variabilidade climática e hidrológica da região. Para calcular as médias mensais de precipitação em cada unidade de análise (UA) da bacia foram usados dados das seguintes estações pluviométricas levantadas:

- UA1: Belo Jardim e Sanharó
- UA2: Caruaru (IPA) e São Caetano
- UA3: Gravatá e Chã Grande
- UA4: Primavera, Escada e Ipojuca

A Figura 17 mostra que a distribuição da precipitação não é homogênea em toda a bacia, com os volumes de chuva aumentando à medida que se aproxima do litoral. No período analisado, a estação chuvosa se concentrou aproximadamente entre março e julho na UA1 e na UA2, enquanto o período chuvoso foi registrado nas UA3 e UA4 entre março e agosto. A UA4, correspondente à Zona da Mata e à Zona Costeira do estado de Pernambuco, apresenta uma variação mais acentuada dos índices pluviométricos entre os meses secos e chuvosos, enquanto as outras UAs mostram um comportamento mais regular da precipitação ao longo do ano.

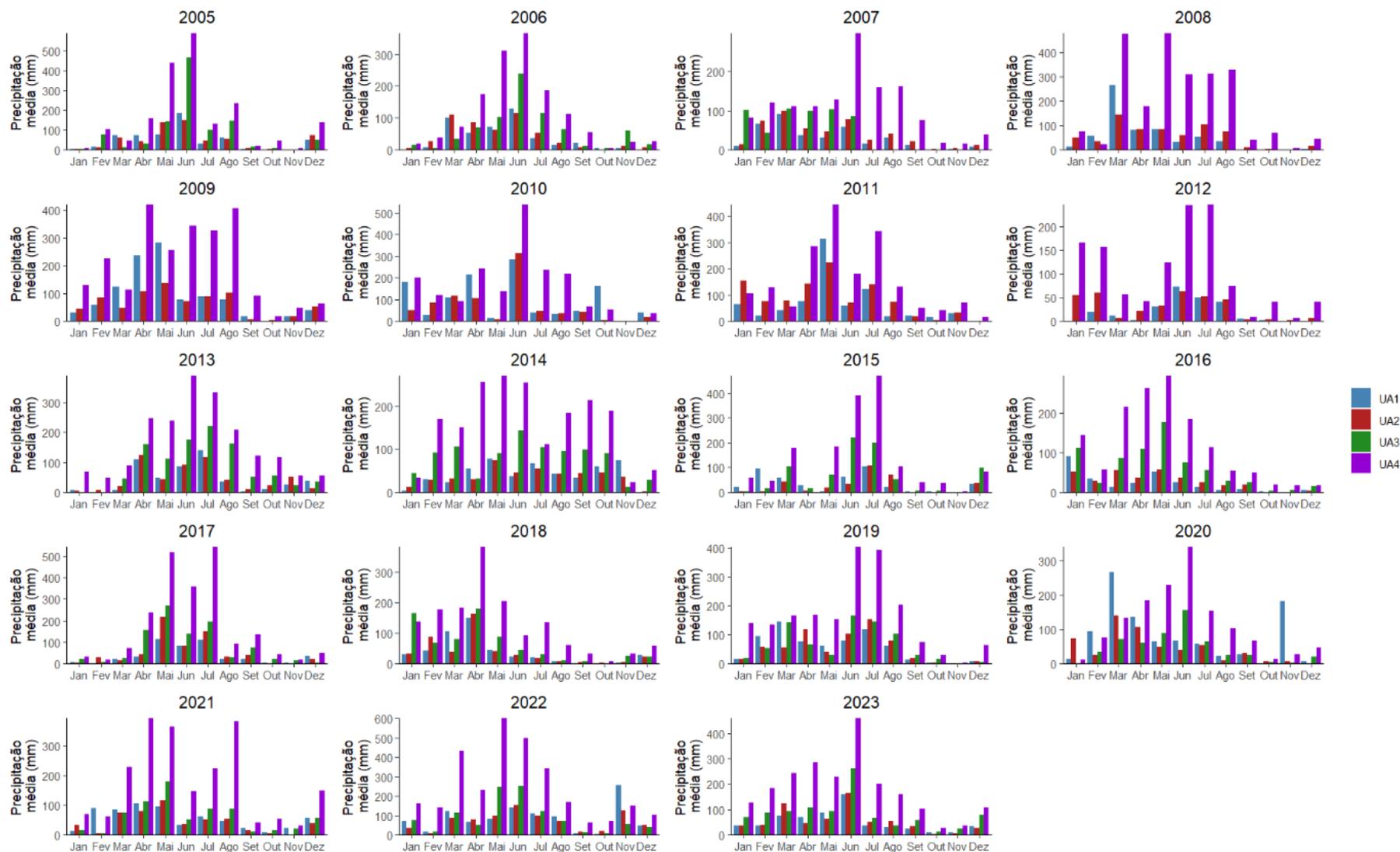
Figura 17 - Precipitação média mensal observada nas unidades de análise da bacia do rio Ipojuca no período de 2005 a 2023



Fonte: Elaborado pela autora a partir dos dados disponibilizados pela APAC (2025)

Ao se realizar uma análise ano a ano no período considerado, os dados de precipitação revelaram certa variabilidade em relação à média mensal dos 19 anos avaliados. Observa-se, na Figura 18, que os anos de 2022, 2017, 2010 e 2005 apresentaram os maiores índices pluviométricos, concentrados principalmente entre os meses de maio e julho. Por outro lado, os anos de 2016, 2014, 2012 e 2007 registraram os menores volumes de precipitação, embora esses valores tenham sido mais distribuídos ao longo dos meses. Silva et al. (2023) analisaram as tendências pluviométricas na bacia do Ipojuca entre 1962 e 2015 e atribuíram as variações observadas à atuação de sistemas meteorológicos de micro e meso escala na região. Os autores não identificaram uma tendência de longo prazo para a precipitação na BHRI, mas observaram uma variabilidade interdecenal, caracterizada por décadas mais secas sucedendo décadas mais chuvosas, e vice-versa.

Figura 18 - Precipitação média mensal observada para cada ano entre 2005 e 2023, nas unidades de análise da bacia do rio Ipojuca



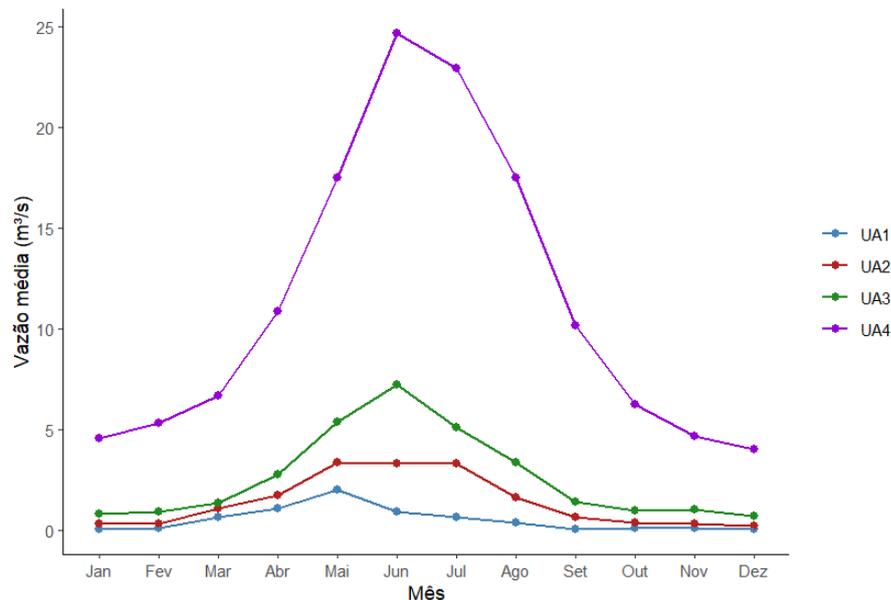
Fonte: Elaborado pela autora a partir dos dados disponibilizados pela APAC (2025)

Considerando essas variações no regime de chuvas e sua influência direta sobre a disponibilidade hídrica, foi realizada a estimativa das médias mensais de vazão em cada UA da bacia, com base nos dados das seguintes estações fluviométricas levantadas:

- UA1: Sanharó
- UA2: Caruaru
- UA3: Gravatá
- UA4: Engenho Tabocas e Engenho Maranhão

As médias mensais dos dados de vazão obtidos para os anos de 2005 a 2023 (Figura 19) reforçam que o rio Ipojuca transiciona de um regime intermitente para um regime perene nas proximidades de Gravatá, na UA3. Os picos de vazão na UA3 e na UA4 ocorreram no mês de junho, enquanto, na UA1, foram registrados em maio. Já na UA2, os picos de vazão se mantiveram constantes entre maio e julho. Assim como observado para a precipitação, a UA4 apresenta os maiores valores de vazão da bacia, com o rio mantendo o escoamento nessa região mesmo durante o período seco.

Figura 19 - Vazão média mensal observada nas unidades de análise da bacia do rio Ipojuca no período de 2005 a 2023



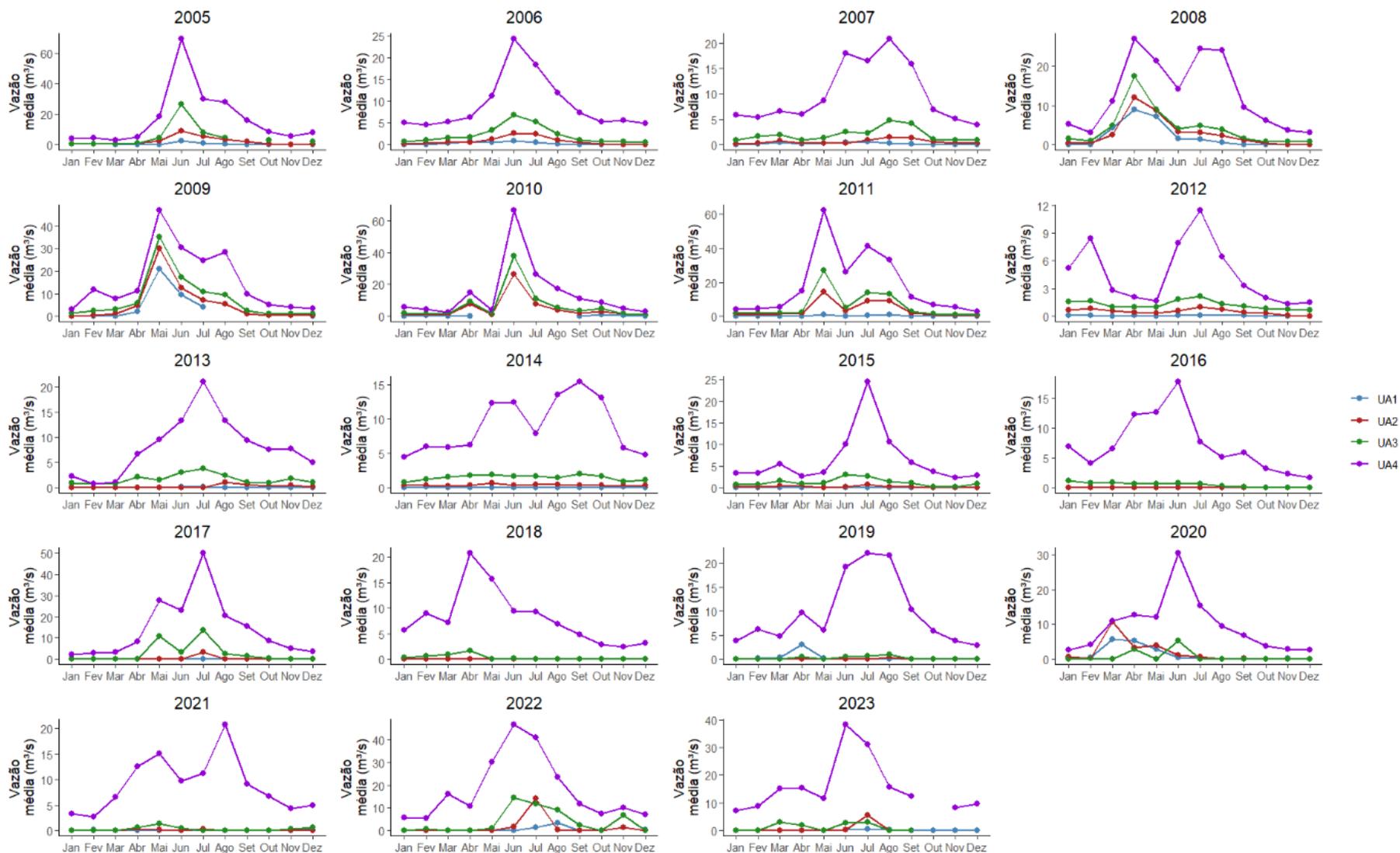
Fonte: Elaborado pela autora a partir dos dados disponibilizados pela ANA (2025)

As condições fluviométricas de uma bacia hidrográfica são influenciadas pela precipitação e pelas descargas dos lençóis freáticos, resultando em componentes de escoamento superficial e subterrâneo. Parte da chuva é interceptada, parte infiltra no solo e o restante esco pela superfície, contribuindo para a vazão dos cursos d'água (HANSFORD et al., 2020).

Almeida et al. (2017) analisaram as curvas de permanência das estações fluviométricas do rio Ipojuca e comprovaram a permanência do curso d'água a partir da estação localizada no município de Gravatá, onde a vazão de referência Q90 não foi nula. As declividades acentuadas dessas curvas indicaram uma maior contribuição da componente superficial para o escoamento do rio Ipojuca, sendo esse comportamento mais perceptível nas estações de Sanharó e Caruaru. Diante dessas informações, os autores inferiram que a permanência da vazão do rio Ipojuca está diretamente associada aos índices pluviométricos da bacia, sendo que, na porção intermitente, a contribuição da componente subterrânea do escoamento é provavelmente insignificante.

Na comparação ano a ano, seguindo o padrão observado na precipitação, as vazões também apresentaram variabilidade em relação à média dos 19 anos. A Figura 20 mostra que os maiores valores de vazão do período foram registrados nos anos de 2017, 2011, 2010 e 2005. Em contrapartida, as menores vazões ocorreram em 2016, 2014 e 2012. Destaca-se, ainda, a ocorrência de mais de um pico de vazão ao longo dos anos de 2021, 2017, 2014, 2012, 2011, 2009, 2008 e 2007.

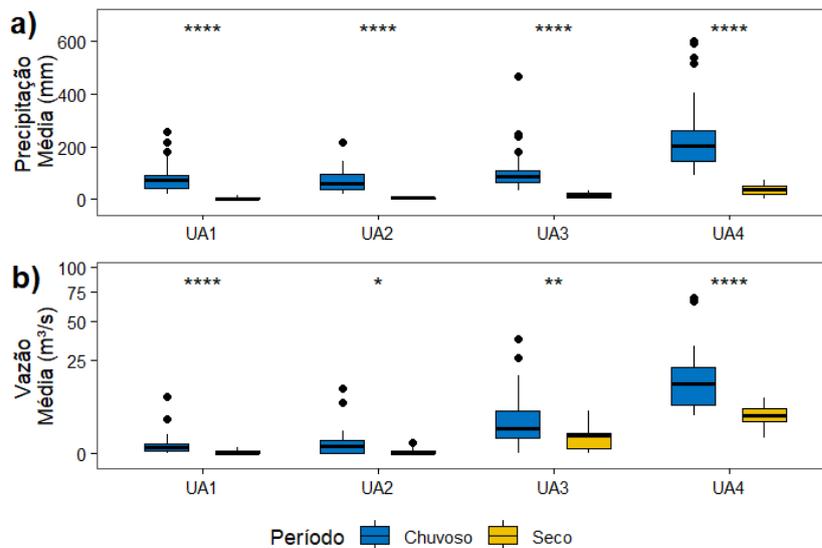
Figura 20 - Vazão média mensal observada para cada ano entre 2005 e 2023, nas unidades de análise da bacia do rio Ipojuca



Fonte: Elaborado pela autora a partir dos dados disponibilizados pela ANA (2025)

Para a avaliação das diferenças entre os períodos seco e chuvoso, foi utilizado o teste de Mann-Whitney (não paramétrico), com nível de significância de 5% ($\alpha = 0,05$). Os resultados foram apresentados graficamente, acompanhados dos *boxplots* das amostras correspondentes a cada período. Os dados de precipitação e vazão utilizados são os mesmos das Figuras 19 e 20, agrupados por unidade de análise da bacia. As diferenças entre os períodos seco e chuvoso são apresentadas na Figura 21.

Figura 21 - Diferenças entre a precipitação (a) e a vazão (b) médias nas unidades de análise da BHRI



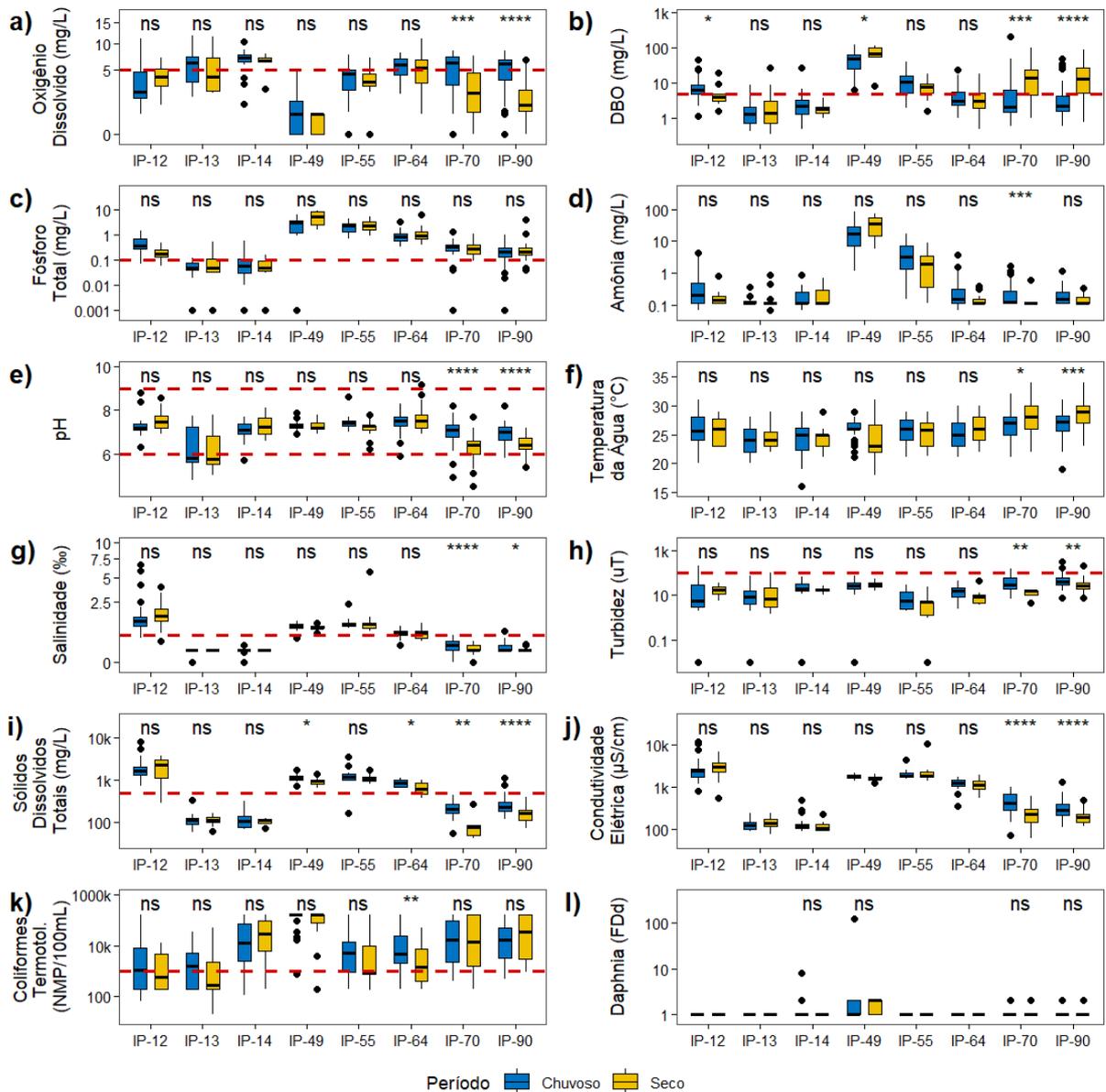
Fonte: Elaborado pela autora (2025)

Notas: Os níveis de significância reconhecidos foram * $p < 0,05$; ** $p < 0,01$; *** $p < 0,001$; **** $p < 0,0001$. Escala do eixo Y em b) ajustada para raiz quadrada para melhor visualização.

Observam-se diferenças significativas entre os períodos seco e chuvoso em todas as unidades de análise para os dados de vazão e precipitação. Os *boxplots* refletem a dinâmica hidrológica da bacia, com a UA1 e a UA2 apresentando menores volumes de chuva e vazão, caracterizando o trecho intermitente, enquanto a UA3 e a UA4 exibem maiores valores, correspondendo ao trecho perene.

As variações nos parâmetros de qualidade da água entre os períodos seco e chuvoso na BHRI são apresentadas na Figura 22. Observa-se que as estações IP-70 e IP-90, situadas na UA4, apresentaram diferenças sazonais significativas em quase todos os 12 parâmetros analisados, com exceção de fósforo total, coliformes termotolerantes e *Daphnia*. Nessas duas estações, os maiores valores para a maioria dos parâmetros foram registrados no período chuvoso, exceto para DBO e temperatura da água, que tiveram valores mais altos no período seco.

Figura 22 - Variação dos parâmetros de qualidade da água nas estações de monitoramento da bacia do rio Ipojuca entre 2005 e 2023, nos períodos seco e chuvoso



Fonte: Elaborado pela autora (2025)

Notas: Os níveis de significância reconhecidos foram * $p < 0,05$; ** $p < 0,01$; *** $p < 0,001$; **** $p < 0,0001$; ns denota diferenças não significativas. Escala do eixo Y ajustada para melhor visualização: raiz quadrada em a) e g), logarítmica (base 10) nos demais, exceto e) e f).

Legenda: IP-12: Rio Ipojuca, na ponte da PE-180, que liga Belo Jardim e São Bento do Una; IP-13: Nascente do Rio Bitury em Belo Jardim; IP-14: Rio Bitury, à montante do reservatório, em Belo Jardim; IP-49: Rio Ipojuca, à jusante de Caruaru; IP-55: Rio Ipojuca, a montante de Gravatá; IP-64: Rio Ipojuca, à jusante de Chã Grande; IP-70: Rio Ipojuca, à jusante da Usina União Indústria, em Primavera; IP-90: Rio Ipojuca, à jusante da Usina Ipojuca, em Ipojuca.

No caso da amônia (Figura 22d), apenas a estação IP-70 apresentou diferença significativa, com maiores valores registrados no período chuvoso. Segundo Freire et al. (2021), o aumento nas concentrações de nitrogênio na estação chuvosa pode estar relacionado à ressuspensão de matéria orgânica acumulada no sedimento ou ao seu transporte pelo

escoamento superficial. Os autores afirmam que as diferenças sazonais dependem do equilíbrio entre a poluição da água do rio, a degradação do nitrogênio (especialmente na estação seca) e o aporte de nitrogênio na estação chuvosa. Aplicando essa perspectiva ao caso da bacia do rio Ipojuca, os dados sugerem que a poluição mantém uma fonte elevada e contínua de nitrogênio ao longo do ano, mostrando-se muito mais relevante do que os outros dois fatores. Isso explicaria a ausência de diferenças sazonais significativas para a amônia na maioria dos locais de amostragem.

Além disso, a estação IP-12 apresentou níveis de DBO significativamente maiores no período chuvoso (Figura 22b), enquanto em IP-49 os maiores valores ocorreram no período seco (assim como em IP-70 e IP-90). Nos estudos de Freire et al. (2021) e Le et al. (2023), as maiores concentrações de DBO ocorreram no período seco e nos trechos urbanizados dos rios que receberam lançamentos de efluentes sanitários e industriais, um cenário comparável ao observado em IP-49, IP-70 e IP-90. O caso de IP-12, por outro lado, reflete o impacto difuso da pecuária leiteira na região, onde a poluição está mais associada ao escoamento superficial de resíduos animais do que a lançamentos pontuais de esgoto (ANDRADE et al., 2021).

Os sólidos dissolvidos totais foram significativamente maiores no período chuvoso em IP-49 e IP-64, assim como em IP-70 e IP-90 (Figura 22i). Altas concentrações de sólidos na estação chuvosa também foram relatadas na literatura, associadas a efluentes domésticos e ao escoamento superficial de campos cultivados (ANDEM et al., 2014; FREIRE et al., 2021). Xiong et al. (2022) apontaram que maiores teores de sólidos e turbidez na estação chuvosa também podem reduzir a penetração de luz na coluna de água e limitar o crescimento de algas.

Quanto aos demais parâmetros, a estação IP-64 foi a única a apresentar diferenças significativas para os coliformes termotolerantes (Figura 22k), com maiores valores registrados no período chuvoso. Fósforo total (Figura 22c) e *Daphnia* (Figura 22l) foram os únicos parâmetros a não apresentarem diferenças significativas em nenhuma das estações monitoradas.

As estações IP-70 e IP-90 tiveram os piores resultados de OD no período seco, exibindo por vezes condições anóxicas (Figura 22a). No estudo de Pak et al. (2021), o OD também apresentou maior concentração durante a estação chuvosa e, correspondentemente, exibiu menor taxa de excedência aos limites da legislação local nesse período, o que os autores atribuíram à maior turbulência e mistura decorrentes da maior vazão do rio.

Adicionalmente, um dos principais subprodutos da indústria sucroalcooleira local é o vinhoto (ou vinhaça), que possui uma elevada carga orgânica. Barros (2008) avaliou a concentração de OD em pontos a jusante de usinas na bacia do Ipojuca entre 1995 e 2006, verificando uma redução significativa nos níveis de OD, com muitos valores iguais a zero,

especialmente entre os meses de outubro e março. Esse período coincidia com a moagem da cana-de-açúcar e com a estação seca, sugerindo uma forte influência dessa atividade industrial na depleção do oxigênio.

Finalmente, em relação à temperatura, os resultados mostraram que as medianas nos trechos mais à jusante na bacia são mais altas que as observadas nos trechos à montante, principalmente no período seco (Figura 22f). Essa observação corrobora o estudo de Barros (2008) na mesma bacia.

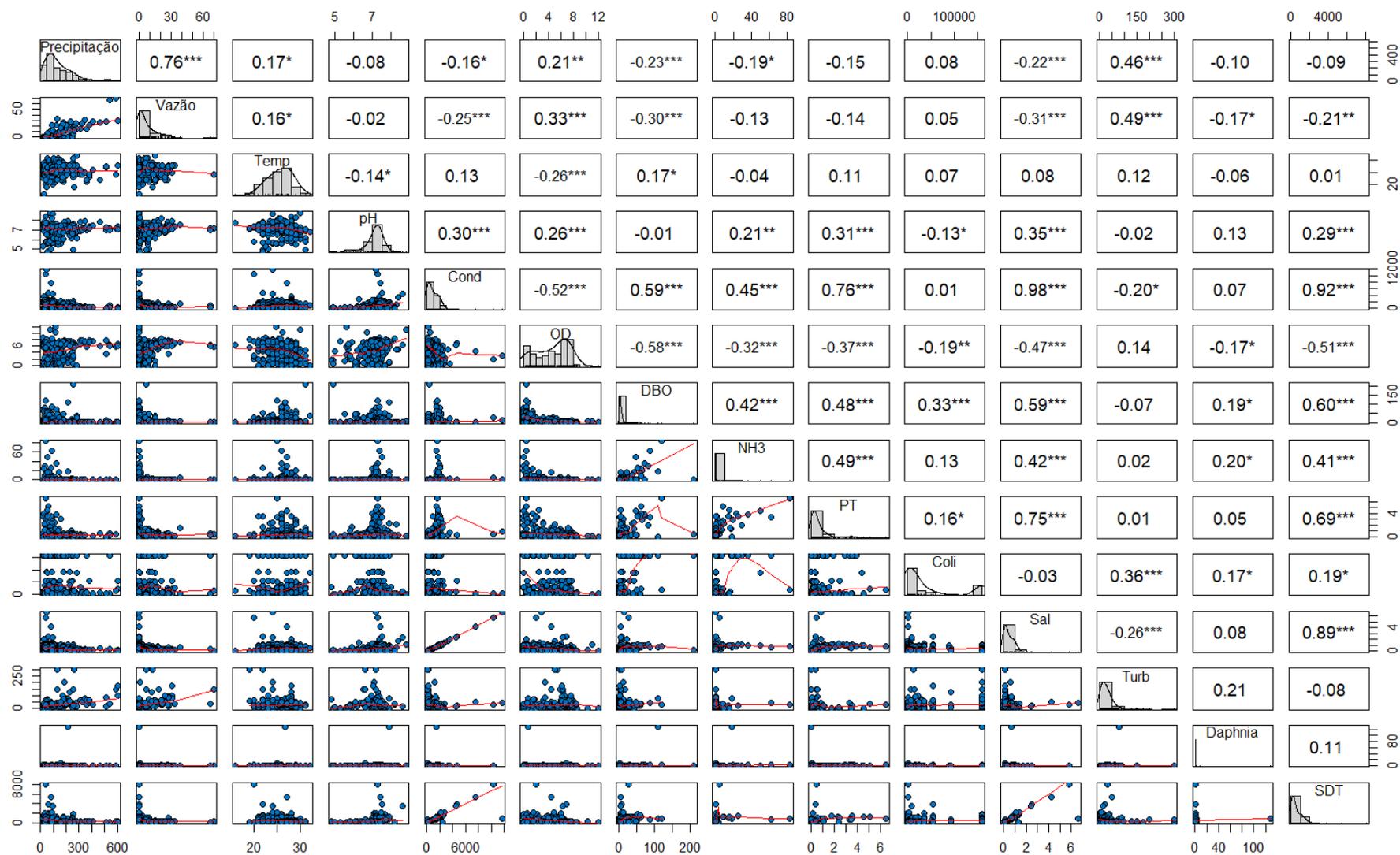
Para quantificar a relação entre os múltiplos parâmetros de qualidade da água, a precipitação e a vazão, foram calculados os coeficientes de correlação de Spearman. Estes coeficientes, juntamente com as matrizes de dispersão dos dados, são apresentados para os períodos chuvoso (Figura 23) e seco (Figura 24).

Durante o período chuvoso, observou-se uma correlação positiva forte ($\rho = 0,76$) entre precipitação e vazão, indicando que o aumento das chuvas eleva significativamente o volume de água nos corpos hídricos, conforme esperado devido ao maior escoamento superficial e contribuição de águas pluviais. Em contrapartida, a condutividade elétrica (Cond) apresentou uma correlação negativa moderada ($\rho = -0,52$) com o OD, e correlações positivas, sendo moderada com a DBO ($\rho = 0,59$) e forte com o fósforo total (PT) ($\rho = 0,76$). Isso indica que locais com maior condutividade também tendem a apresentar maior carga orgânica (DBO) e concentração de nutrientes, como o fósforo. Esse padrão é compatível com o observado principalmente nas estações IP-12, IP-49 e IP55 na BHRI e pode estar associado ao carreamento de efluentes urbanos, agrícolas ou dejetos animais durante as chuvas, que transportam simultaneamente sais, matéria orgânica e nutrientes para o rio (ANDRADE et al., 2021; FREIRE et al., 2021).

O OD apresentou correlações negativas moderadas com a DBO ($\rho = -0,58$) e com os sólidos dissolvidos totais (SDT) ($\rho = -0,51$), reforçando que ambientes com maior demanda bioquímica de oxigênio e maior concentração de sólidos dissolvidos tendem a ter menor oxigenação. Essa relação pode ocorrer seja pelo consumo microbiano na decomposição da matéria orgânica, seja pela interferência de íons na solubilidade do oxigênio (CETESB, 2023). Por fim, a DBO e o PT mostraram correlações positivas, sendo moderada e forte com a salinidade (Sal) ($\rho = 0,59$ e $\rho = 0,75$, respectivamente) e moderadas com os SDT ($\rho = 0,60$ e $\rho = 0,69$). Estes resultados evidenciam que tanto a matéria orgânica quanto o fósforo estão associados a águas com maior teor de sais e outros constituintes dissolvidos. No estudo de Xiong et al. (2022) também foram observadas correlações positivas entre fósforo total e sólidos

na estação chuvosa, e os autores sugeriram que o fósforo estaria principalmente na sua forma particulada nessas condições.

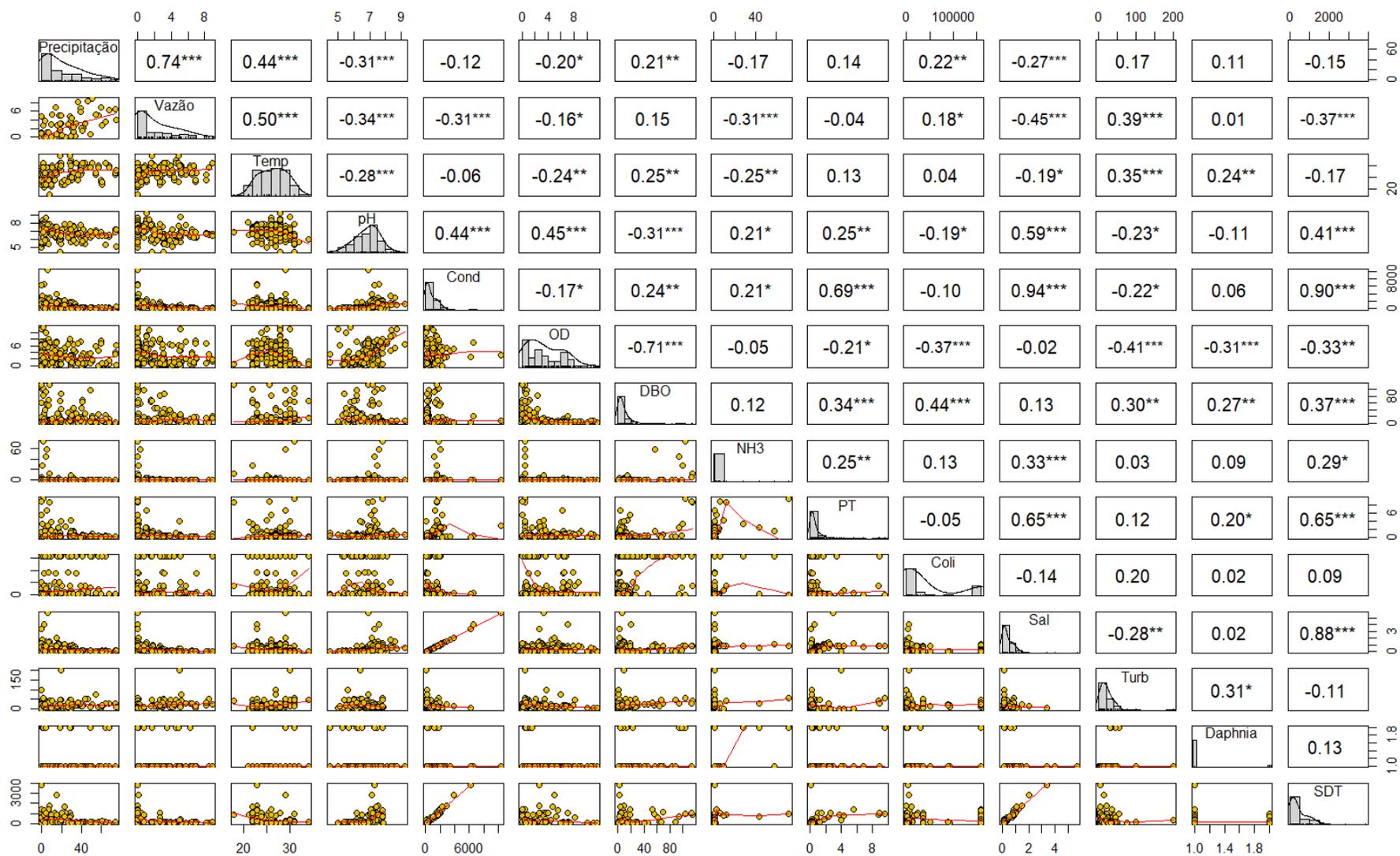
Figura 23 - Matriz de gráficos de dispersão dos dados do período chuvoso, com gráficos de dispersão bivariados abaixo da diagonal, histogramas na diagonal e correlação de Spearman acima da diagonal



Fonte: Elaborado pela autora (2025)

Nota: Os níveis de significância reconhecidos foram * $p < 0,05$; ** $p < 0,01$; *** $p < 0,001$.

Figura 24 - Matriz de gráficos de dispersão dos dados do período seco, com gráficos de dispersão bivariados abaixo da diagonal, histogramas na diagonal e correlação de Spearman acima da diagonal



Fonte: Elaborado pela autora (2025)

Nota: Os níveis de significância reconhecidos foram * $p < 0,05$; ** $p < 0,01$; *** $p < 0,001$.

Durante o período seco, a alta correlação positiva entre precipitação e vazão ($\rho = 0,74$) manteve-se significativa, porém ligeiramente menor do que a observada no período chuvoso ($\rho = 0,76$). Isso indica que, mesmo na estiagem, eventos de chuva ainda influenciam a vazão dos corpos hídricos, embora com menor intensidade devido à redução no volume de escoamento superficial. Além disso, a vazão apresentou uma correlação positiva moderada com a temperatura ($\rho = 0,50$), relação que era fraca no período chuvoso. Contrariamente ao esperado pela literatura, onde vários estudos demonstraram que a temperatura da água é inversamente relacionada à vazão do rio (WEBB, 1996; WEBB et al., 2003), a correlação positiva encontrada na BHRI sugere a influência de fatores locais não abordados nesses estudos.

O pH demonstrou uma correlação positiva moderada com a salinidade ($\rho = 0,59$), relação inexistente ou fraca no período chuvoso. Essa associação pode ser atribuída à maior influência de processos como a evaporação durante o período seco, que concentra sais e pode elevar o pH em águas com baixo fluxo, ou à redução da diluição de substâncias alcalinas presentes naturalmente no ambiente (CETESB, 2023).

As correlações envolvendo a condutividade elétrica mostraram mudanças significativas entre os períodos. A relação entre condutividade e OD ($\rho = -0,17$) e entre condutividade e DBO ($\rho = 0,24$), que eram moderadas no período chuvoso ($\rho = -0,52$ e $\rho = 0,59$, respectivamente), tornam-se fracas no período seco. Isso indica que o impacto das substâncias dissolvidas (representadas pela condutividade) na disponibilidade de oxigênio e na DBO é menos pronunciado na estiagem. Tal fato pode ocorrer devido à menor entrada de matéria orgânica e poluentes por escoamento superficial, reduzindo a DBO associada a fontes difusas. Por outro lado, a correlação entre condutividade e fósforo total ($\rho = 0,69$) manteve-se moderada, ainda que menor do que no período chuvoso ($\rho = 0,76$). Essa persistência sugere que fontes pontuais de poluição (como esgotos ou efluentes industriais) podem continuar a liberar nutrientes e sais mesmo com a redução do fluxo de água.

A correlação negativa entre OD e DBO ($\rho = -0,71$) tornou-se mais forte no período seco do que no chuvoso ($\rho = -0,58$). Isso indica que, com a menor diluição e vazão características do período seco, a decomposição de matéria orgânica passa a consumir oxigênio de forma mais acentuada, agravando a redução do OD (LEE et al., 2016). Em contraste, a relação entre OD e SDT ($\rho = -0,33$) tornou-se mais fraca do que na estação chuvosa ($\rho = -0,51$), provavelmente porque outros fatores, como a DBO, passam a ter influência dominante na oxigenação, diminuindo a importância relativa dos sólidos totais dissolvidos na ausência de grande aporte por chuvas.

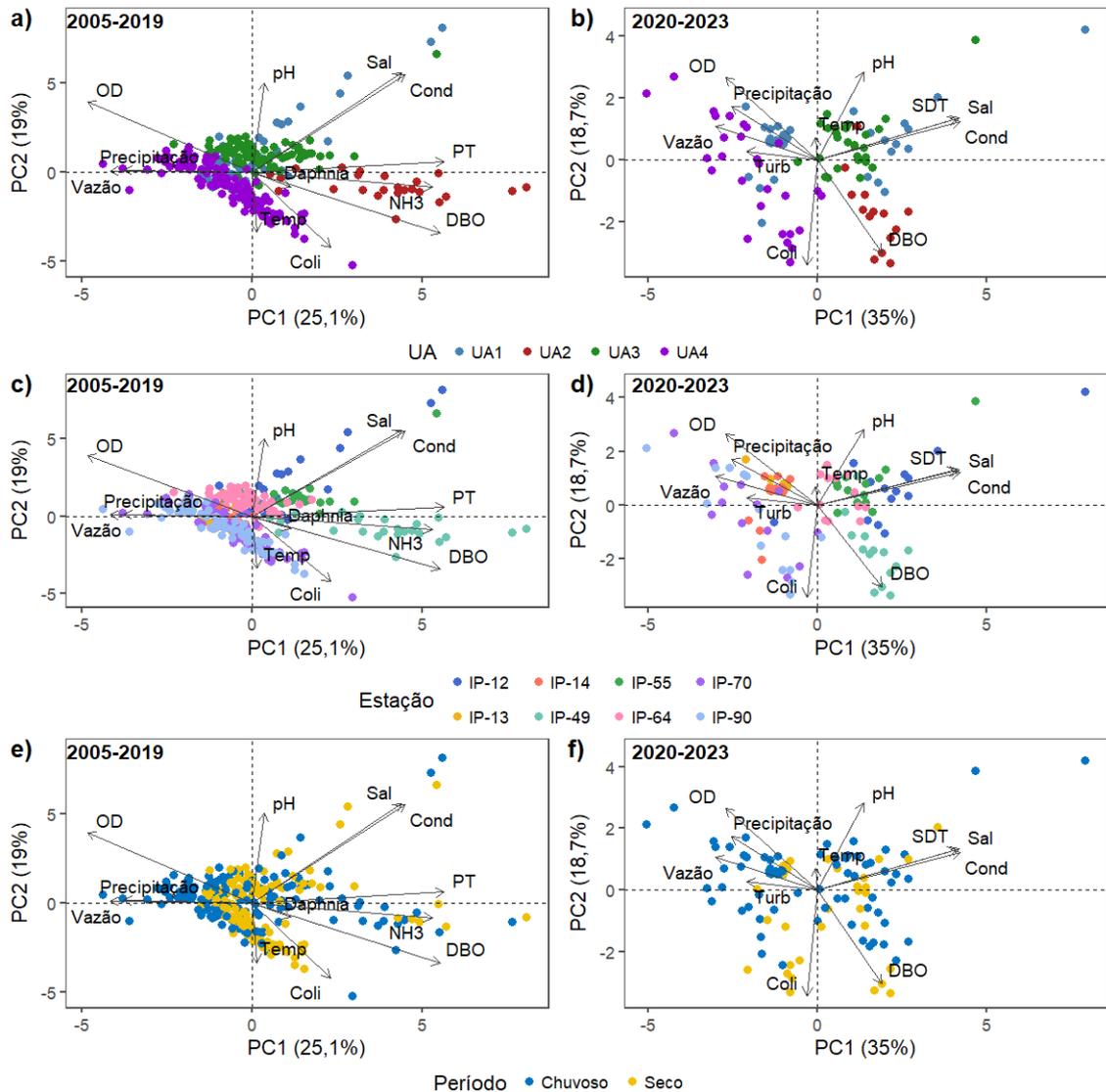
Quanto à DBO, suas correlações com salinidade ($\rho = 0,13$) e SDT ($\rho = 0,37$) são consideravelmente mais fracas do que no período chuvoso ($\rho = 0,59$ e $\rho = 0,60$, respectivamente). Isso reforça que a matéria orgânica dissolvida está menos associada a íons e outros constituintes no período seco, possivelmente devido à redução no transporte de poluentes difusos. Por fim, o fósforo total manteve correlações positivas moderadas com salinidade ($\rho = 0,65$) e SDT ($\rho = 0,65$), ainda que menores do que no período chuvoso ($\rho = 0,75$ e $\rho = 0,69$).

Conforme observado nos *boxplots* apresentados anteriormente, verificou-se um comportamento semelhante entre os sólidos dissolvidos totais, a condutividade elétrica e a salinidade, parâmetros considerados indicadores da mineralização da água. Essa relação é corroborada pelas altas correlações encontradas entre eles, tanto no período seco quanto no período chuvoso. Os gráficos de dispersão entre condutividade e salinidade exibem uma forte relação linear positiva, indicando uma associação muito estreita entre essas duas variáveis, com coeficientes de correlação de $\rho = 0,98$ no período chuvoso e $\rho = 0,94$ no período seco. Além disso, foram observadas altas correlações positivas entre a condutividade e os sólidos dissolvidos totais ($\rho = 0,92$ no chuvoso; $\rho = 0,90$ no seco) e entre a salinidade e os sólidos dissolvidos ($\rho = 0,89$ no chuvoso; $\rho = 0,88$ no seco). Essas correlações representam as relações mais fortes encontradas entre os parâmetros avaliados na BHRI e são condizentes com a literatura (SUN et al., 2023; XIONG et al., 2022).

4.1.3 Parâmetros que exercem maior influência na qualidade da água

A Análise de Componentes Principais permitiu visualizar a importância relativa dos parâmetros de qualidade da água na BHRI em relação a cada componente principal (Figura 25). A divisão dos dados em dois períodos distintos (2005-2019 e 2020-2023) possibilitou uma análise mais adequada às características do conjunto de dados, respeitando as limitações impostas pela disponibilidade de informações para certos parâmetros em parte do período total.

Figura 25 - *Biplot* da Análise de Componentes Principais para os dois primeiros eixos nos períodos de 2005-2019 e 2020-2023



Fonte: Elaborado pela autora (2025)

Nota: As figuras a) e b) apresentam as amostras agrupadas por unidades de análise; c) e d), por estações de amostragem; e e) e f), por períodos chuvoso e seco.

Legenda: IP-12: Rio Ipojuca, na ponte da PE-180, que liga Belo Jardim e São Bento do Una; IP-13: Nascente do Rio Bitury em Belo Jardim; IP-14: Rio Bitury, à montante do reservatório, em Belo Jardim; IP-49: Rio Ipojuca, à jusante de Caruaru; IP-55: Rio Ipojuca, a montante de Gravatá; IP-64: Rio Ipojuca, à jusante de Chã Grande; IP-70: Rio Ipojuca, à jusante da Usina União Indústria, em Primavera; IP-90: Rio Ipojuca, à jusante da Usina Ipojuca, em Ipojuca.

Para o período de 2005-2019, os resultados da ACP indicaram cinco componentes principais (PCs) que, conjuntamente, explicaram 76,74% da variância total dos dados (Tabela 5). A primeira componente representou 25,12% da variância total dos dados e mostrou correlação positiva com os parâmetros fósforo total, DBO, amônia, condutividade e salinidade, e negativa com o OD e a vazão. Essa combinação de variáveis na PC1 sugere a influência de esgotos domésticos e efluentes urbanos, especialmente na estação IP-49 (Caruaru, pertencente à UA2), onde esses parâmetros apresentaram valores mais elevados. Esta interpretação está

condizente com diversos resultados na literatura, nos quais a primeira componente está frequentemente associada à degradação da qualidade da água por fontes de poluição (FREIRE et al., 2021; GONZÁLEZ et al., 2024; ZORAL-ALMEIDA et al., 2018).

A segunda componente (PC2) explicou 18,96% da variância total e foi positivamente relacionada aos parâmetros salinidade, condutividade e pH, e negativamente aos coliformes termotolerantes. Similarmente, a terceira componente (PC3) explicou 14,20% da variância e foi positivamente influenciada pela precipitação e pela vazão. A quarta componente (PC4) representou 9,83% da variância total, sendo positivamente influenciada pela temperatura da água. Por fim, a quinta componente (PC5) representou 8,63% da variância total e foi positivamente influenciada pela *Daphnia*.

Em relação à distribuição das amostras, observa-se uma clara distinção entre as unidades de análise (Figura 25a). Entretanto, observa-se também que, dentre as estações da UA1, as amostras de IP-13 e IP-14 ficam sobrepostas pelas amostras das estações IP-70 e IP-90 (Figura 25c), que fazem parte da UA4. Isso provavelmente se deve ao fato de que essas quatro estações de amostragem são as únicas dentre os locais analisados a apresentarem águas predominantemente doces. As estações da UA4 foram as únicas que mostraram maior distinção entre as amostras coletadas nos períodos seco e chuvoso (Figura 25e), o que corrobora os resultados do teste de Mann-Whitney apresentados anteriormente. Segundo o teste, essas estações apresentaram diferenças sazonais significativas em quase todos os 12 parâmetros analisados, sugerindo que a resposta hidrológica nessa região interfere significativamente na variabilidade dos parâmetros de qualidade da água.

Tabela 5 - Matriz de componentes principais derivada da ACP com dados sobre parâmetros de qualidade da água na bacia do rio Ipojuca, entre 2005-2019

Parâmetro	PC1	PC2	PC3	PC4	PC5
PT	0,708	0,075	0,208	-0,322	-0,344
DBO	0,692	-0,429	0,304	-0,084	0,058
NH3	0,664	-0,107	0,347	-0,405	-0,156
Cond	0,564	0,690	0,080	0,419	0,060
Sal	0,551	0,698	0,063	0,423	0,063
Coli	0,290	-0,531	0,208	0,133	0,195
<i>Daphnia</i>	0,143	-0,104	0,254	-0,095	0,880
pH	0,044	0,630	0,360	-0,312	0,053
Temp	0,018	-0,423	0,182	0,539	-0,172
Precipitação	-0,472	0,014	0,781	0,115	-0,039
Vazão	-0,522	0,009	0,745	0,144	-0,168
OD	-0,607	0,493	0,043	-0,320	0,080
Autovalor	3,01	2,28	1,70	1,18	1,04
% de variância explicada	25,12	18,96	14,20	9,83	8,63

Parâmetro	PC1	PC2	PC3	PC4	PC5
% de variância acumulada	25,12	44,08	58,28	68,11	76,74

Fonte: Elaborado pela autora (2025)

Para o período de 2020-2023, a ACP resultou em apenas três componentes principais, responsáveis por 70,66% da variância total dos dados (Tabela 6). A primeira componente (PC1) representou 34,99% da variância total e foi positivamente correlacionada com os parâmetros condutividade, salinidade e sólidos dissolvidos totais, e negativamente com a vazão, o OD e a precipitação. Este resultado indica que a mineralização e a diluição foram os fatores primários de variação neste período. A segunda componente (PC2) explicou 18,75% da variância total e foi positivamente relacionada aos parâmetros pH e OD, e negativamente aos coliformes termotolerantes e à DBO. A terceira componente (PC3) explicou 16,93% da variância e foi positivamente influenciada pela turbidez, precipitação e pela vazão.

Quanto à distribuição das amostras neste segundo no período (2020-2023), fica ainda mais clara a separação entre as estações com águas predominantemente doces (IP-13 e IP14, na UA1; IP-70 e IP-90, na UA4) e aquelas com água predominantemente salobra (IP-12, na UA1; IP-49, na UA2; IP-55 e IP-64, na UA3) (Figura 25b e d). Na ausência de dados recentes sobre o fósforo total e a amônia, que eram dominantes na estação IP-49 (Caruaru, na UA2) juntamente com a DBO entre 2005-2019, os indicadores de mineralização da água se tornaram os fatores mais explicativos da variabilidade da qualidade da água na bacia, conforme refletido pela PC1.

Observou-se comportamento sazonal semelhante aos anos anteriores (2005-2019) para as estações da UA4, que continuaram mostrando distinção mais visível entre as amostras coletadas nos períodos seco e chuvoso. Porém, percebe-se também que, no período 2020-2023, o vetor da DBO no *biplot* aponta diretamente para as amostras da estação IP-49 coletadas no período seco (Figura 25f), demonstrando que essas amostras apresentam altos valores para esse parâmetro. Isso reforça a hipótese de permanência de fontes pontuais de carga orgânica em Caruaru, mesmo em um contexto em que os indicadores de salinidade ganharam maior destaque na explicação da variância global dos dados.

Tabela 6 - Matriz de componentes principais derivada da ACP com dados sobre parâmetros de qualidade da água na bacia do rio Ipojuca, entre 2020-2023

Parâmetro	PC1	PC2	PC3
Cond	0,898	0,262	0,314
Sal	0,892	0,285	0,307
SDT	0,875	0,281	0,327
DBO	0,407	-0,642	0,241
pH	0,293	0,601	-0,201
Temp	-0,011	0,151	0,234

Parâmetro	PC1	PC2	PC3
Coli	-0,066	-0,725	0,376
Turb	-0,441	0,059	0,690
Precipitação	-0,540	0,367	0,628
OD	-0,576	0,566	-0,267
Vazão	-0,638	0,230	0,570
Autovalor	3,85	2,06	1,86
% de variância explicada	34,99	18,75	16,93
% de variância acumulada	34,99	53,74	70,66

Fonte: Elaborado pela autora (2025)

4.1.4 Análise de tendências

A análise de ponto de mudança realizada na BHRI, por meio do teste de Pettitt, identificou alterações significativas nas séries temporais de parâmetros de qualidade da água em quase todas as estações de amostragem, com exceção da IP-13 (nascente do rio Bitury, em Belo Jardim), onde não foram detectadas mudanças (Tabela 7).

Entre os resultados significativos (Figura 26), observou-se, na estação IP-12, uma diminuição da média da temperatura da água a partir de 2015. Na IP-14, houve um aumento da condutividade elétrica a partir de 2013. A estação IP-49 apresentou duas mudanças: uma diminuição da temperatura da água a partir de 2018 e um aumento da DBO a partir de 2012. Na IP-55, registrou-se a diminuição da temperatura da água e o aumento dos coliformes termotolerantes, ambos a partir de 2018. Já na IP-64, houve redução dos níveis médios de oxigênio dissolvido a partir de 2012 e aumento dos coliformes termotolerantes a partir de 2017. A estação IP-70 apresentou múltiplas mudanças: redução da temperatura da água e da DBO a partir de 2017, diminuição da condutividade elétrica e da salinidade a partir de 2015, e um aumento do oxigênio dissolvido a partir de 2015. Por fim, na estação IP-90, foi identificada uma diminuição da média da temperatura da água a partir de 2018. Esses resultados evidenciam alterações importantes nas características da água ao longo do tempo, possivelmente associadas a mudanças no uso do solo, no clima regional ou em práticas de manejo ambiental.

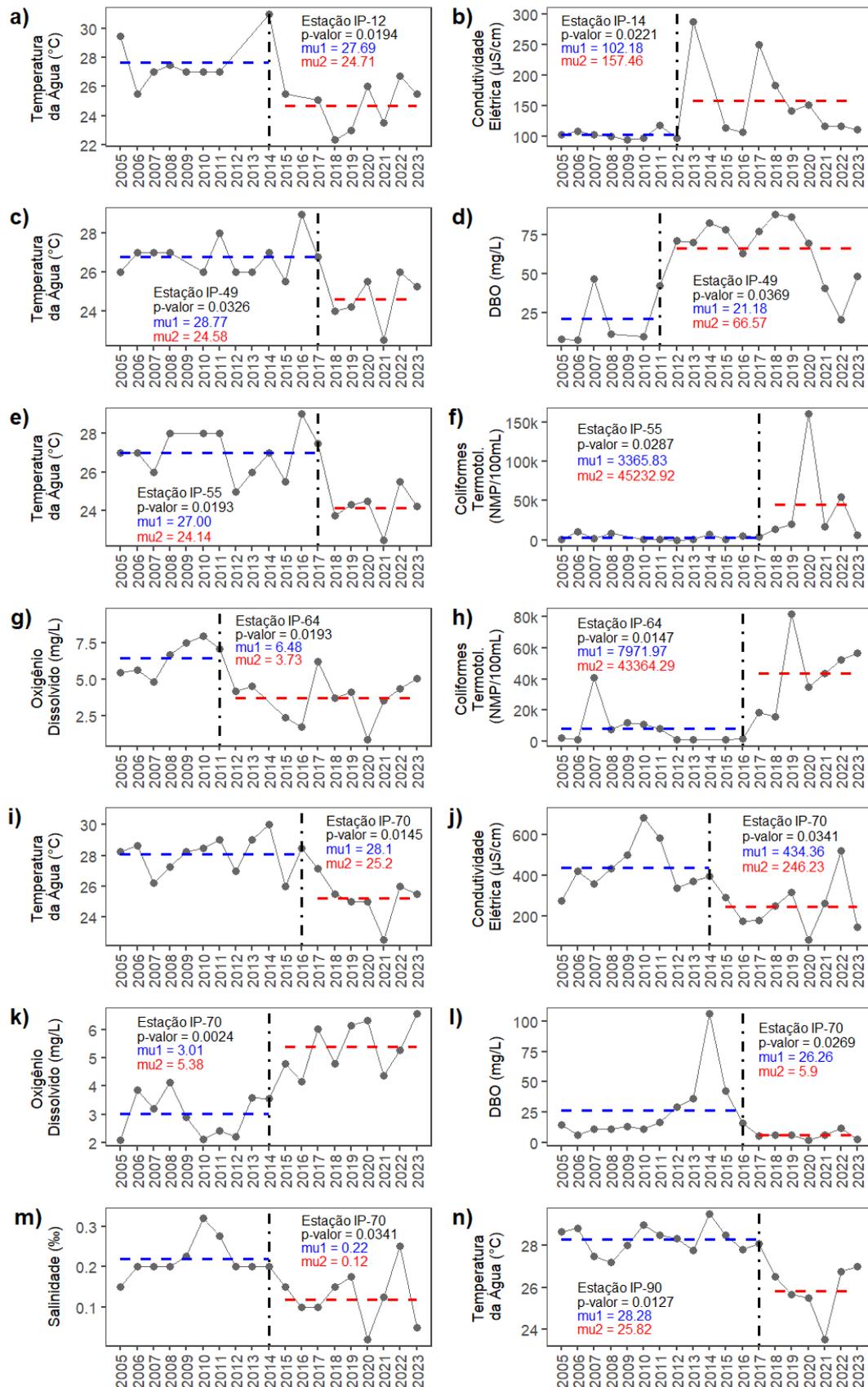
Tabela 7 - Resultados do teste de Pettitt para os parâmetros de qualidade da água na BHRI.

Estação	Teste de Pettit	Temp	pH	Cond	OD	DBO	NH3	PT	Coli	Sal
IP-12	U	58	46	42	38	44	33	17	38	40
	p-valor	0,0194	0,1082	0,1757	0,2732	0,1386	0,0610	0,7920	0,2732	0,2203
	K*	8	4	8	4	6	7	5	10	8
IP-13	U	34	51	45	58	44	24	37	37	32
	p-valor	0,6482	0,1585	0,2779	0,07535	0,3031	0,7658	0,2042	0,5267	0,7372
	K*	4	6	15	8	7	10	8	6	16
IP-14	U	26	27	68	41	33	37	25	25	32
	p-valor	1	0,9828	0,02207	0,3886	0,6919	0,1224	0,5586	1	0,7372
	K*	9	9	8	11	13	8	5	11	12
IP-49	U	65	24	22	42	64	35	36	37	25
	p-valor	0,0326	1	1	0,3584	0,0369	0,1642	0,142	0,5267	1
	K*	12	7	15	5	6	7	6	11	11
IP-55	U	69	38	28	42	44	25	38	66	38
	p-valor	0,0193	0,4896	0,8097	0,3584	0,3031	0,5586	0,1050	0,0287	0,4896
	K*	12	7	8	6	12	3	6	12	7
IP-64	U	50	40	44	69	48	22	21	71	35
	p-valor	0,1749	0,4205	0,3031	0,0193	0,2117	0,7448	0,8131	0,0147	0,6060
	K*	10	8	14	7	4	6	7	11	14
IP-70	U	77	45	70	90	72	26	37	36	70
	p-valor	0,0145	0,3717	0,0341	0,0024	0,0269	0,6482	0,2042	0,6812	0,0341
	K*	12	10	10	10	12	3	9	14	10
IP-90	U	78	48	42	28	50	26	38	44	39
	p-valor	0,0127	0,2948	0,4617	1	0,2505	0,6482	0,1802	0,4002	0,5650
	K*	13	8	7	6	5	9	6	14	11

Fonte: Elaborado pela autora (2025)

Notas: * Provável ponto de mudança na série temporal. Os valores estatisticamente significativos ($p < 0,05$) estão destacados em negrito.

Figura 26 - Parâmetros com pontos de mudança significativos nas estações de amostragem na BHRI



Fonte: Elaborado pela autora (2025)

Notas: A linha vertical tracejada-pontilhada marca o ano identificado como o ponto de mudança pelo teste de Pettitt. A linha tracejada azul representa o valor médio antes do ponto de mudança (mu1), e a linha tracejada vermelha representa o valor médio após o ponto de mudança (mu2).

Com base nos pontos de mudança significativos identificados pelo teste de Pettitt, essas séries temporais foram divididas em dois segmentos para avaliação de tendências: antes e depois dos pontos de mudança detectados. O teste de Mann-Kendall indicou tendência significativa (positiva) apenas para a DBO na estação IP-70, especificamente no segmento anterior ao ponto de mudança (Tabela 8). Contudo, vale ressaltar que, embora a média de DBO no primeiro segmento (26,26 mg/L entre 2005-2016) tenha sido consideravelmente maior que a do segundo segmento (5,9 mg/L entre 2017-2023), essa média pré-mudança é fortemente influenciada pelo pico de DBO registrado no ano de 2014 (acima de 100 mg/L), o que pode distorcer a percepção da tendência geral nesse período inicial.

Tabela 8 - Resultados do teste de Mann-Kendall para séries temporais segmentadas em antes e depois dos pontos de mudança identificados

Estação	Parâmetro	Ponto de Mudança	Segmento	Teste de Mann Kendall	
				Tau	p-valor
IP-12	Temp	2014	Antes	0,161	0,6902
			Depois	0,255	0,4544
IP-14	Cond	2012	Antes	-0,357	0,2655
			Depois	-0,378	0,1524
IP-49	Temp	2017	Antes	0,000	1,0000
			Depois	0,333	0,4524
	DBO	2011	Antes	0,333	0,4524
			Depois	-0,333	0,1499
IP-55	Temp	2017	Antes	0,048	0,8887
			Depois	0,200	0,7071
	Coli	2017	Antes	-0,092	0,7305
			Depois	-0,067	1,0000
IP-64	OD	2011	Antes	0,619	0,0715
			Depois	0,055	0,8763
	Coli	2016	Antes	-0,273	0,2758
			Depois	0,524	0,1331
IP-70	Temp	2016	Antes	0,171	0,4898
			Depois	-0,250	0,5387
	Cond	2014	Antes	0,200	0,4743
			Depois	0,056	0,9170
	OD	2014	Antes	0,022	1,0000
			Depois	0,423	0,1422
	DBO	2016	Antes	0,576	0,0112
			Depois	0,048	1,0000
	Sal	2014	Antes	0,272	0,3600
			Depois	0,000	1,0000
IP-90	Temp	2017	Antes	-0,013	1,0000
			Depois	0,200	0,7071

Fonte: Elaborado pela autora (2025)

Nota: O valor estatisticamente significativo ($p < 0,05$) está em negrito.

Para os parâmetros que não apresentaram pontos de mudança significativos na série temporal, aplicou-se o teste de Mann-Kendall ao conjunto completo de dados para investigar tendências monotônicas (Tabela 9). Os resultados indicaram algumas tendências significativas em diferentes estações de amostragem (Figura 27). Na estação IP-12, foi observada uma tendência ascendente nos valores de coliformes termotolerantes, sugerindo agravamento da contaminação fecal ao longo do tempo na zona da pecuária leiteira em Belo Jardim. Na IP-13, foram identificadas duas tendências: uma ascendente para o pH, indicando elevação da alcalinidade da água, e uma descendente para a salinidade, o que pode estar relacionado à redução de sais dissolvidos na nascente do rio Bitury. A estação IP-49 apresentou tendência ascendente para o fósforo total, provavelmente associada ao aumento de cargas poluentes de origem antrópica em Caruaru ao longo dos anos, reflexo do crescimento da cidade. Já na IP-64, foi identificada uma tendência de aumento nos valores de DBO, o que pode refletir maior presença de matéria orgânica na água à jusante de Chã Grande. Por fim, na estação IP-90, também foi registrada uma tendência ascendente para o pH.

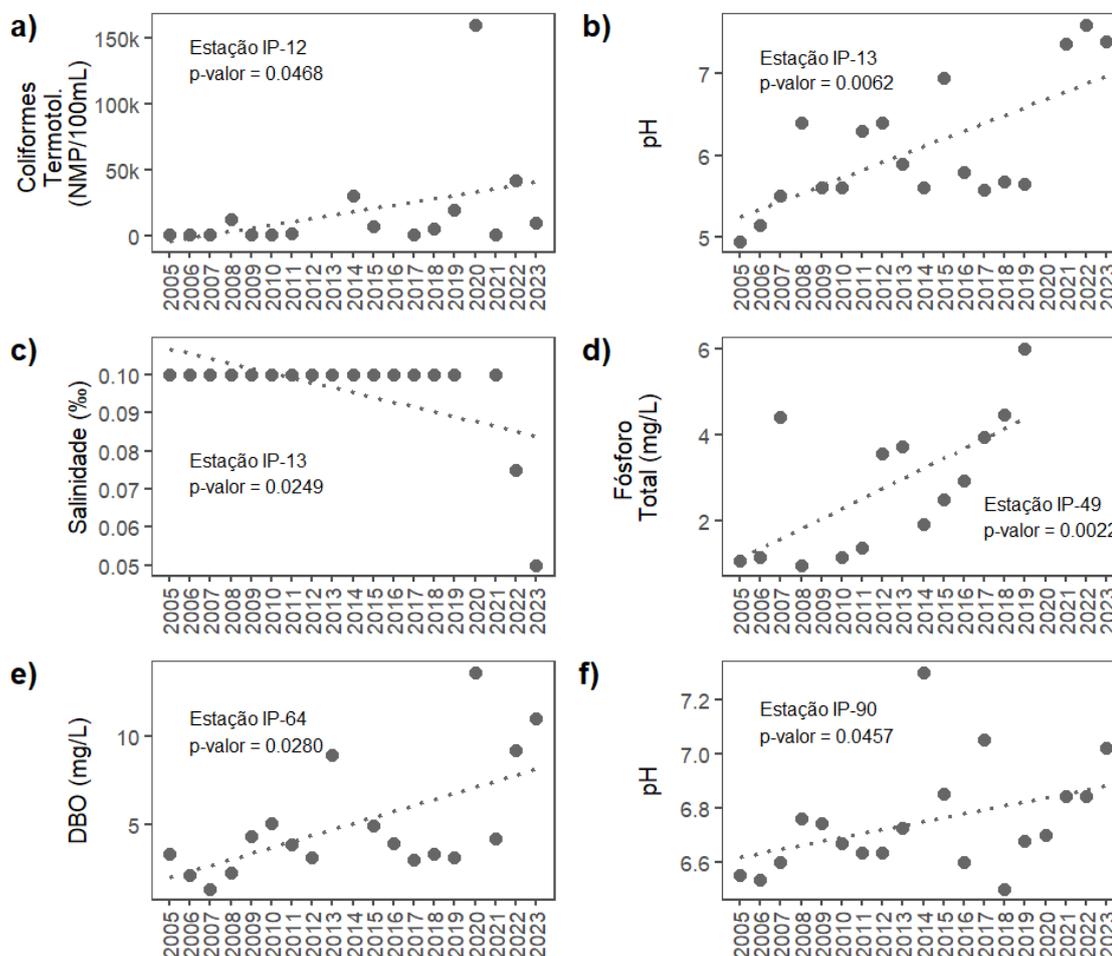
Tabela 9 - Resultados do teste de Mann-Kendall para séries temporais de parâmetros médios anuais nas estações de amostragem da BHRI

Estação	Teste de Mann Kendall	Temp	pH	Cond	OD	DBO	NH3	PT	Coli	Sal
IP-12	Tau	-	-0,192	0,200	-0,217	0,217	0,350	0,229	0,380	0,205
	p-valor	-	0,3214	0,3004	0,2604	0,2604	0,1415	0,3359	0,0468	0,2968
IP-13	Tau	0,175	0,483	-0,007	0,294	0,013	0,153	-0,298	0,273	-0,464
	p-valor	0,3401	0,0062	1,0000	0,0956	0,9698	0,4999	0,1367	0,1274	0,0249
IP-14	Tau	-0,020	0,172	-	0,320	-0,039	0,304	0,189	0,124	-0,185
	p-valor	0,9392	0,3426	-	0,0690	0,8497	0,1793	0,3796	0,4954	0,3725
IP-49	Tau	-	0,148	-0,059	-0,178	-	0,319	0,626	-0,196	0,000
	p-valor	-	0,4222	0,7731	0,3372	-	0,1253	0,0022	0,3339	1,0000
IP-55	Tau	-	-0,040	0,147	-0,020	0,118	-0,055	0,385	-	0,196
	p-valor	-	0,8495	0,4338	0,9396	0,5193	0,8267	0,0627	-	0,2845
IP-64	Tau	-0,335	-0,053	0,255	-	0,386	-0,155	0,253	-	0,179
	p-valor	0,0616	0,7908	0,1501	-	0,0280	0,4760	0,2284	-	0,3233
IP-70	Tau	-	0,300	-	-	-	0,000	-0,306	0,181	-
	p-valor	-	0,0799	-	-	-	1,0000	0,1245	0,2939	-
IP-90	Tau	-	0,342	-0,099	-0,053	0,240	0,153	-0,371	0,123	-0,142
	p-valor	-	0,0457	0,5756	0,7796	0,1617	0,4574	0,0600	0,4841	0,4337

Fonte: Elaborado pela autora (2025)

Notas: O valor estatisticamente significativo ($p < 0,05$) está em negrito. '-' indica que a tendência foi calculada antes e depois dos pontos de mudança na Tabela 8.

Figura 27 - Parâmetros que apresentaram tendências significativas na BHRI



Fonte: Elaborado pela autora (2025)

Nota: A linha pontilhada cinza representa a tendência linear.

Essas tendências observadas na BHRI alinham-se ao entendimento geral sobre os efeitos das atividades humanas nos corpos d'água. As intervenções humanas (por exemplo, agricultura, aquicultura, urbanização, construção de reservatórios) podem causar um aumento da carga de nutrientes, levar à deterioração da qualidade da água e resultar na eutrofização e na proliferação de fitoplâncton (ZHAO et al., 2019; CAO et al., 2023). Entretanto, ações de gestão ambiental podem ter o efeito contrário, melhorando a qualidade da água.

No estudo de Shi et al. (2019), por exemplo, foi identificado que o programa de coleta de esgoto em Melbourne, Austrália, como uma renovação de infraestrutura, levou à remoção de muitas descargas diretas de águas cinzas e à correção de vazamentos de fossas sépticas defeituosas, reduzindo o número de fossas sépticas de 1.486 em 2007 para 791 em 2016, o que potencialmente explicou a queda nos níveis de fósforo total nos pontos analisados.

Similarmente, no estudo de Wu et al. (2021), na China, observou-se que entre 2005 e 2018, as mudanças no uso do solo foram caracterizadas principalmente pelo aumento das áreas

construídas, enquanto as áreas agrícolas aumentaram entre 2005 e 2010, mas diminuíram após esse período. As áreas de floresta diminuíram consideravelmente entre 2005 e 2010, estabilizando-se posteriormente, e as áreas de pastagem apresentaram poucas alterações ao longo do tempo. Esses resultados refletiram o impacto positivo de políticas públicas implementadas nesse país, como o programa de conversão de terras agrícolas em florestas e pastagens, que conseguiu conter a perda contínua de cobertura vegetal nativa. No entanto, o avanço da urbanização, especialmente entre 2005 e 2010, substituiu grandes extensões de floresta por áreas urbanas, contribuindo para a degradação ambiental. Esse exemplo evidencia como ações políticas eficazes podem tanto mitigar quanto intensificar os impactos sobre a qualidade da água, dependendo da forma como o território é manejado. Políticas voltadas à preservação e recuperação de cobertura vegetal nativa, aliadas ao controle do crescimento urbano desordenado, são fundamentais para melhorar a qualidade da água em bacias hidrográficas.

4.2 ANÁLISE DA RELAÇÃO ENTRE O USO E COBERTURA DO SOLO E A QUALIDADE DA ÁGUA NA BACIA DO RIO IPOJUCA

Para analisar a relação entre a qualidade da água e o uso do solo na BHRI, utilizaram-se mapas do uso e cobertura do solo da plataforma MapBiomias referentes aos anos de 2005 e 2019. O ano de 2019 foi escolhido como ano final porque os dados para os parâmetros amônia e fósforo total, que são importantes indicadores de poluição, estão disponíveis somente até esse período, permitindo assim a análise nesse marco temporal. A evolução das áreas ocupadas pelas diferentes classes de uso do solo é apresentada na Tabela 10 e na Figura 28.

Tabela 10 - Área e proporção de área das classes de uso do solo na BHRI em 2005 e 2019

Classe de uso do solo	2005		2019	
	Área (km ²)	%	Área (km ²)	%
Agropecuária	1905,79	53,05	1714,19	47,68
Formação Savânica	1133,32	31,55	1267,768	35,26
Cana	302,71	8,43	292,26	8,13
Formação Florestal	127,24	3,54	162,04	4,51
Área Urbanizada	93,4	2,60	138,73	3,86
Corpo D'água	26,08	0,73	16,69	0,46
Formação Natural não Florestal	3,61	0,10	3,32	0,09
Total	3592,15	100,00	3595,00	100,00

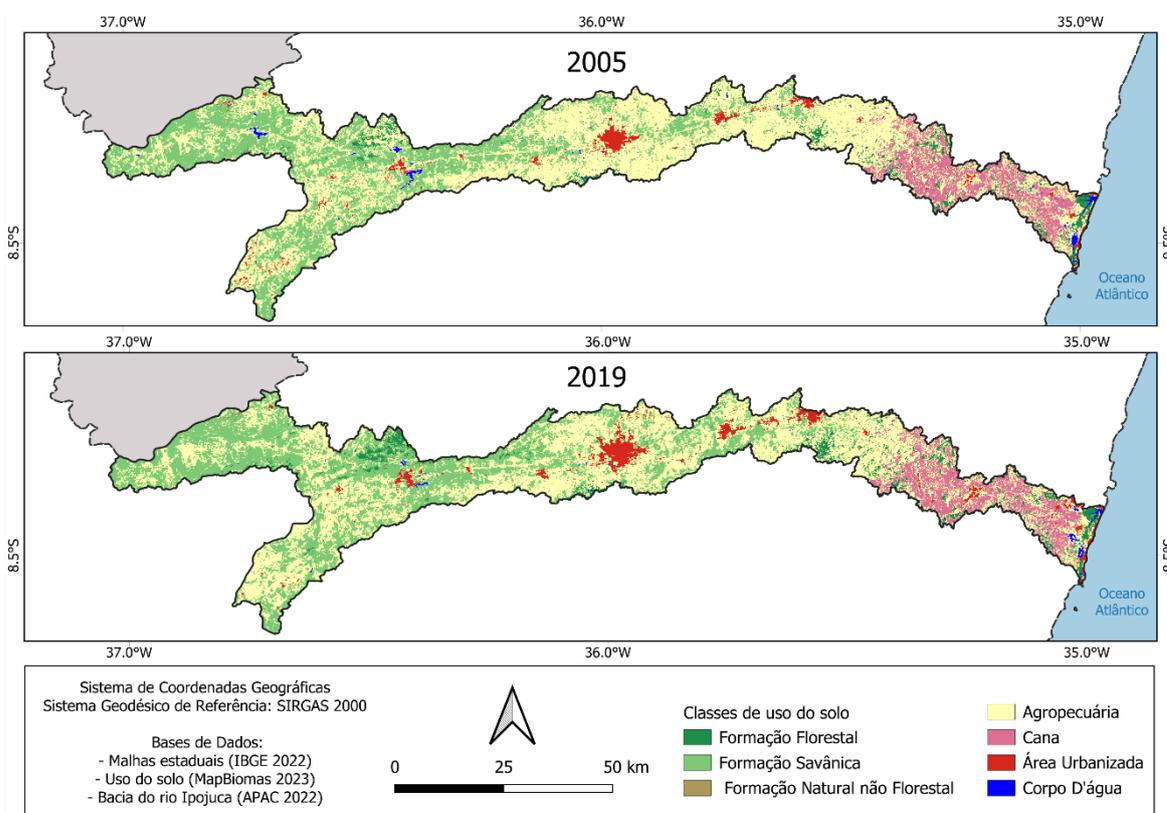
Fonte: Elaborado pela autora (2025)

Os resultados indicaram uma redução nas áreas destinadas à agropecuária, que passaram de 1.905,79 km² (53,05%) em 2005 para 1.714,19 km² (47,68%) em 2019, representando a principal mudança no uso do solo durante o período analisado. Por outro lado, a classe que representa a Formação Savânica apresentou um aumento, passando de 1.133,32 km² (31,55%) para 1.267,77 km² (35,26%).

As áreas destinadas ao cultivo de cana-de-açúcar na bacia apresentaram uma leve redução, de 302,71 km² (8,43%) para 292,26 km² (8,13%). Em contrapartida, a Formação Florestal aumentou de 127,24 km² (4,54%) para 162,04 km² (4,51%), indicando um leve crescimento na cobertura vegetal. A área urbanizada também apresentou crescimento, aumentando de 93,4 km² (2,60%) em 2005 para 138,73 km² (3,86%) em 2019, refletindo o avanço da urbanização no território na bacia.

Por fim, houve redução nas áreas ocupadas pela classe Corpos d'água, de 26,08 km² (0,73%) para 16,69 km² (0,46%), e uma ligeira diminuição nas Formações Naturais não Florestais, que passaram de 3,61 km² (0,10%) para 3,32 km² (0,09%).

Figura 28 - Evolução do uso do solo na BHRI nos anos de 2005 e 2019



Fonte: Elaborado pela autora (2025)

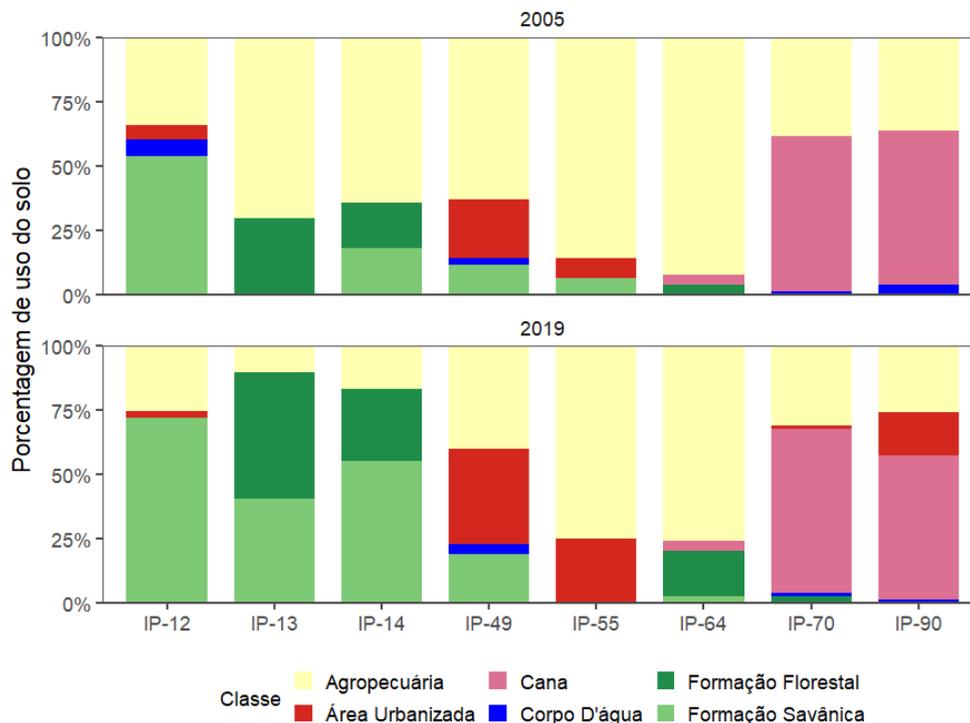
Em relação à composição do uso do solo no raio de 500 metros ao redor das estações de amostragem, comparando os anos de 2005 e 2019, os resultados revelaram alterações

expressivas em alguns dos pontos, refletindo as dinâmicas de uso e ocupação do solo em escala local (Figura 29).

Nas estações IP-12, IP-13 e IP-14, observou-se um aumento significativo da cobertura por Formação Savânica, com destaque para IP-12 e IP-14, onde essa classe passou a representar a maior parte da paisagem no entorno imediato das estações (mais de 70% em IP-12 e mais de 50% em IP-14). Ao mesmo tempo, notou-se uma redução na área de Agropecuária nessas mesmas estações. A estação IP-49, localizada em região urbana, manteve-se com predominância de Área Urbanizada e Agropecuária, embora a área de Formação Savânica também tenha aumentado levemente. Em IP-55 e IP-64, destacou-se a redução da Agropecuária, com o crescimento da Área Urbanizada ao redor de IP-55 e da Formação Florestal em IP-64. Já em IP-70, a configuração do uso do solo se manteve relativamente estável entre os dois anos analisados, com forte presença da área de cultivo de Cana e Agropecuária. Na estação IP-90, observou-se um aumento da Área Urbanizada e leve redução da Agropecuária.

De maneira geral, a comparação entre os anos revela um padrão de substituição de áreas agropecuárias por formação savânica em várias estações, o que pode estar associado a processos naturais de regeneração da vegetação ou a mudanças no manejo do solo. Essas alterações têm implicações diretas sobre a dinâmica da qualidade da água, uma vez que o uso do solo influencia fortemente os processos de erosão, escoamento e carga de poluentes nas áreas adjacentes aos corpos hídricos (CAI et al., 2020; WEISSTEINER et al., 2013).

Figura 29 - Composição do uso do solo em raio de 500 metros ao redor das estações de amostragem de qualidade da água da BHRI



Fonte: Elaborado pela autora (2025)

A partir dos dados de uso do solo em um raio de 500 metros ao redor das estações de amostragem, foi realizada uma análise de correlação entre os padrões de uso do solo e os parâmetros de qualidade da água na BHRI utilizando-se o método de Spearman (Tabela 11). Os resultados indicam relações significativas entre determinados tipos de uso do solo e as variáveis físico-químicas e biológicas da água, revelando como a ocupação do território influencia a qualidade dos recursos hídricos.

Tabela 11 - Resultados da análise de correlação entre os padrões de uso do solo e os parâmetros de qualidade da água na BHRI

Parâmetros	Formação Florestal	Formação Savânica	Agropecuária	Cana	Área Urbanizada	Corpo d'água
Temperatura	-0,527*	-0,282	0,265	0,367	0,046	0,569*
pH	-0,483	0,410	0,311	-0,369	0,565*	0,049
OD	0,311	0,223	-0,049	-0,195	-0,297	-0,218
DBO	-0,803***	-0,295	0,200	0,221	0,746***	0,452
NH3	-0,698**	-0,271	0,251	0,125	0,770***	0,303
PT	-0,433	-0,353	0,688**	-0,003	0,471	0,179
Coli	-0,136	0,061	-0,271	0,22	0,237	0,293
Sal	-0,596*	0,228	0,245	-0,215	0,653**	0,112
<i>Daphnia</i>	-0,216	0,205	0,028	-0,193	0,449	0,341

Fonte: Elaborado pela autora (2025)

Nota: Os níveis de significância reconhecidos foram * $p < 0,05$; ** $p < 0,01$; *** $p < 0,001$.

A classe Formação Florestal apresentou correlações negativas significativas com a DBO ($\rho = -0,803$), a amônia ($\rho = -0,698$), a salinidade ($\rho = -0,596$) e a temperatura ($\rho = -0,527$). Essas correlações evidenciam o papel protetor da vegetação para os ecossistemas aquáticos, reduzindo a carga de poluentes e a variabilidade térmica dos corpos d'água. A cobertura florestal atua como um filtro, retendo os sedimentos e poluentes transportados pelo escoamento superficial, além de aumentar a concentração de oxigênio dissolvido (CHEN et al., 2021; MELLO et al., 2018). Neste estudo, a Formação Florestal apresentou correlação positiva com o oxigênio dissolvido, entretanto, o resultado não foi estatisticamente significativo.

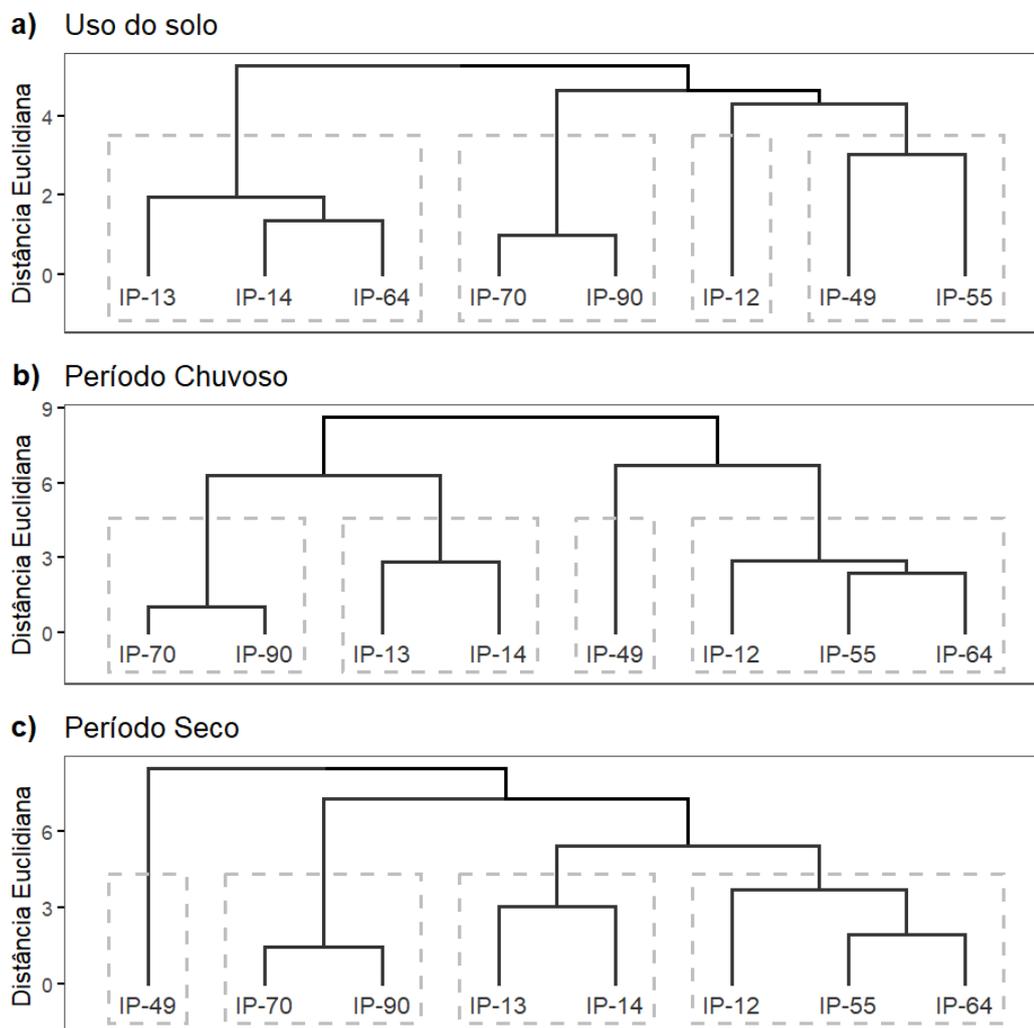
Por outro lado, as áreas agropecuárias apresentaram correlação positiva significativa com o fósforo total ($\rho = 0,688$), sugerindo que esse tipo de uso do solo contribui para a entrada de nutrientes nos corpos hídricos da bacia. Essa classe de uso do solo já foi identificada em diversos estudos na literatura como uma importante fonte difusa de nutrientes, como o fósforo, por meio de fertilizantes, fezes de animais e matéria orgânica (SHI et al., 2019; ZHOU et al., 2016).

A classe Área Urbanizada mostrou-se fortemente associada à degradação da qualidade da água, com correlações positivas e altamente significativas com a DBO ($\rho = 0,746$), a amônia ($\rho = 0,770$), a salinidade ($\rho = 0,653$) e o pH ($\rho = 0,565$), indicando o impacto das atividades urbanas no aumento da carga orgânica e da concentração de compostos químicos nos corpos d'água da bacia, o que é uma observação comum na literatura (HAN et al., 2023; XIONG et al., 2022).

Além disso, o uso do solo classificado como Corpo d'água também apresentou correlação significativa positiva com a temperatura ($\rho = 0,569$), o que pode refletir a maior exposição direta à radiação solar e o menor sombreamento nessas áreas (MELLO et al., 2020). Já as formações savânicas e as áreas de cana-de-açúcar não apresentaram correlações estatisticamente significativas com os parâmetros de qualidade da água analisados.

A Figura 30 apresenta os resultados da Análise de Cluster (agrupamento) das estações de amostragem com base em três critérios distintos: (a) uso do solo em um raio de 500 metros ao redor das estações, (b) parâmetros de qualidade da água no período chuvoso e (c) parâmetros de qualidade da água no período seco. Cada um desses critérios resultou em diferentes padrões de agrupamento em quatro grupos, sugerindo variações espaciais e sazonais na qualidade da água das estações monitoradas.

Figura 30 - Agrupamentos de acordo com (a) uso do solo e (b) indicadores de qualidade da água no período chuvoso e (c) no período seco



Fonte: Elaborado pela autora (2025)

Em relação ao uso do solo, o primeiro grupo contém as estações IP-13, IP-14 e IP-64. Esta última (IP-64) está localizada mais à jusante das duas primeiras, mas as três possuem em comum a dominância das classes Formação Florestal e Agropecuária em seu entorno. Por conter uma pequena área da classe Cana na sua adjacência, a estação IP-64 aparece, no dendrograma (Figura 30), próxima do segundo grupo, formado pelas estações IP-70 e IP-90. Ao redor dessas últimas estações (IP-70 e IP-90), predominam áreas de cultivo de cana-de-açúcar, seguidas por áreas agropecuárias. Apenas a estação IP-12 foi atribuída ao terceiro grupo, devido à sua baixa similaridade com as demais, resultado da alta proporção da classe Formação Savânica em seu entorno. O quarto grupo incluiu as estações IP-49 e IP-55, que possuem alta proporção das classes Área Urbanizada e Agropecuária em seu entorno.

Os agrupamentos baseados nos parâmetros de qualidade da água resultaram nas mesmas composições de grupos nos períodos seco e chuvoso, formando-se também quatro grupos, compostos pelas seguintes estações:

- IP-70 e IP-90
- IP-13 e IP-14
- IP-49
- IP-12, IP-55 e IP-64

No entanto, observa-se que esses grupos diferiram daqueles baseados no uso do solo, com exceção das estações IP-70 e IP-90, que permaneceram juntas nos três cenários analisados (uso do solo, período chuvoso, período seco), constituindo um grupo estável.

Embora a composição dos grupos fosse a mesma entre os períodos chuvoso e seco, a análise de cluster indicou uma variação na similaridade entre os grupos. Já no período seco, a principal diferença observada foi a maior proximidade (menor distância no dendrograma) entre o grupo da estação IP-49 e o grupo estável formado por IP-70 e IP-90, o que pode ser explicado pela redução do fluxo hídrico e maior concentração de contaminantes no período de estiagem.

De acordo com esses resultados, a análise de cluster introduz uma camada de complexidade relevante. O fato de os agrupamentos baseados na qualidade da água não espelharem perfeitamente aqueles baseados no uso do solo no entorno das estações de amostragem, à exceção do grupo associado à cana-de-açúcar/agropecuária (IP-70 e IP-90), sugere que outros fatores, além da composição do uso do solo nos 500m, afetam as características da água em cada estação. Isso pode incluir a influência de fontes pontuais específicas, a conectividade hidrológica na paisagem, efeitos de usos do solo mais distantes ou mesmo características intrínsecas do trecho do rio. Adicionalmente, a dinâmica sazonal mostrou-se determinante; a maior similaridade observada entre a estação de influência urbana (IP-49) e o grupo de IP-70 e IP-90 durante o período seco reforça como a redução da vazão pode intensificar o impacto relativo de certos poluentes, alterando as relações de similaridade entre locais e destacando a vulnerabilidade da bacia em períodos de estiagem.

A Análise de Redundância (RDA) foi utilizada para explorar mais a fundo as conexões entre os tipos de uso do solo e os parâmetros de qualidade da água (Tabela 12 e Figura 31). Para a RDA, é necessário ter duas matrizes de dados: uma para variáveis explicativas ou preditoras (uso do solo) e outra para variáveis resposta (qualidade da água). Inicialmente, a RDA foi realizada utilizando a matriz das variáveis explicativas completa (uso do solo ao redor das estações de amostragem com 6 classes), comparada à matriz de variáveis resposta tanto para o período chuvoso quanto para o seco.

A análise inicial com os modelos completos revelou uma capacidade explicativa considerável da variância nos dados de qualidade da água:

- No período chuvoso, o modelo explicou 58% da variância total (R^2 ajustado = 0,37 ou 37%; $F = 2,77$; $p = 0,009$).
- No período seco, essa explicação foi ainda maior, atingindo 69,8% da variância (R^2 ajustado = 0,51 ou 51%; $F = 3,69$; $p = 0,001$).

Contudo, em ambos os períodos, verificou-se multicolinearidade entre as variáveis preditoras, indicando redundância e justificando a necessidade de um processo de seleção de variáveis para obter modelos mais simples.

Para simplificar os modelos, aplicou-se o método de seleção progressiva (*forward selection*).

- Período Chuvoso: O modelo selecionado incluiu apenas três classes de uso do solo: Área Urbanizada, Formação Florestal e Formação Savânica. Este modelo simplificado explicou 50% da variância total nos dados de qualidade da água, com um R^2 ajustado de 0,375 (ou 37,5%) ($F = 4,00$; $p = 0,001$). Observa-se que o R^2 ajustado deste modelo foi praticamente idêntico ao do modelo completo (0,37), indicando que as três variáveis selecionadas retiveram a capacidade explicativa essencial com maior simplicidade.
- Período Seco: Para este período, o processo de seleção resultou em um modelo com quatro variáveis: Área Urbanizada, Formação Florestal, Formação Savânica e Cana. Este conjunto explicou 65,4% da variância total (R^2 ajustado = 0,50 ou 50%; $F = 4,25$; $p = 0,001$). Assim como no período chuvoso, o modelo simplificado do período seco manteve essencialmente a capacidade explicativa do modelo completo (comparando-se os valores de R^2 ajustado: 0,50 e 0,51), mas utilizando menos variáveis.

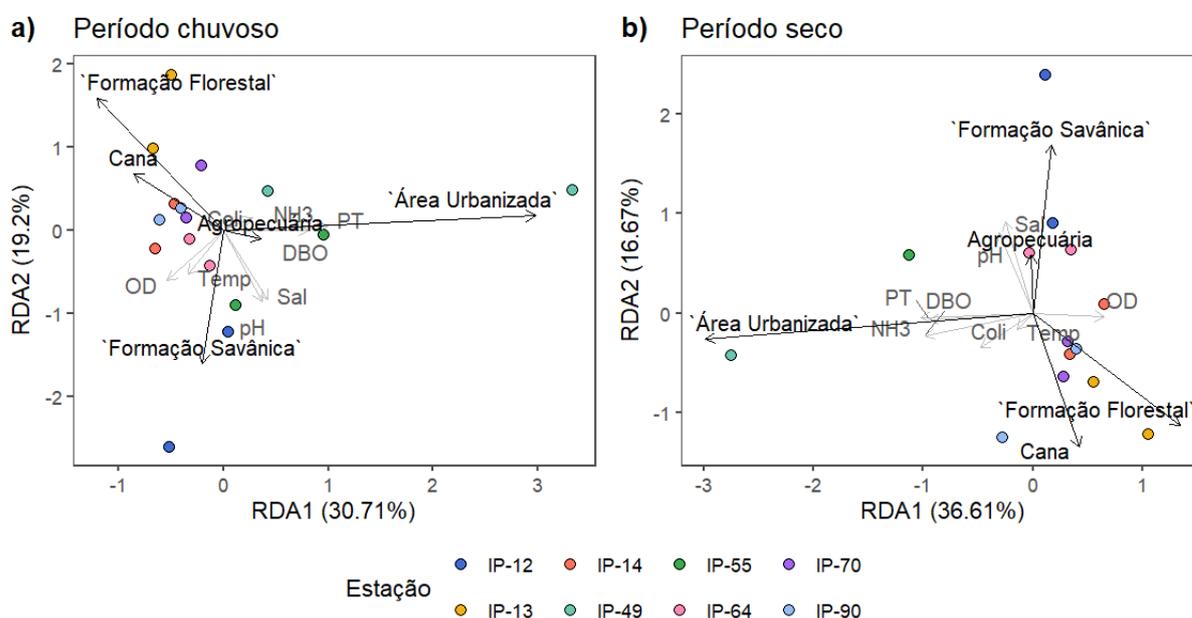
Vale ressaltar a inclusão da classe Cana no modelo selecionado para o período seco, o que não ocorreu no período chuvoso, sugerindo uma possível relevância sazonal desta classe na explicação da variância da qualidade da água na bacia.

Tabela 12 - Resultados da análise de redundância para a porcentagem da variância geral da qualidade da água explicada pelo uso do solo na BHRI

Período	Modelo	Variação explicada (%)			R ² ajustado	F	p-valor
		Eixo 1	Eixo 2	Todos			
Chuvoso	Completo	30,71	19,20	58,04	0,370	2,77	0,009
	Selecioneado	30,46	16,08	50,00	0,375	4,00	0,001
Seco	Completo	36,61	16,67	69,79	0,509	3,69	0,001
	Selecioneado	36,61	16,08	65,36	0,499	4,25	0,001

Fonte: Elaborado pela autora (2025)

Figura 31 - Gráficos de ordenação (*triplots*) mostrando a relação entre as classes de uso do solo (setas pretas), os parâmetros de qualidade da água (setas cinzas) e as amostras individuais nos períodos chuvoso e seco, para o modelo completo



Fonte: Elaborado pela autora (2025)

Os *triplots* da RDA (Figura 31) ilustram a relação entre as estações de amostragem (círculos), os parâmetros de qualidade da água (setas cinzas) e as classes de uso do solo (setas pretas). Os eixos canônicos do *triplot* da RDA para o período seco explicaram uma porcentagem maior da variação nos dados de qualidade da água em comparação com o período chuvoso. Isso sugere que o uso do solo exerce uma maior influência na qualidade da água da BHRI durante o período seco. As longas setas correspondentes às classes Área Urbanizada, Formação Florestal e Formação Savânica indicam que estas classes foram as mais influentes na qualidade da água, em ambos os períodos analisados. A classe Cana, por sua vez, mostra uma influência mais pronunciada no período seco. A Área Urbanizada apresenta um efeito antagônico ao das demais classes, indicado por sua seta apontar em direção oposta às destas últimas. A distribuição das estações (círculos) no gráfico indica que muitas são influenciadas pelas variações nas classes

Formação Florestal e Cana. Por outro lado, a Área Urbanizada mostra forte associação com a qualidade da água da estação IP-49, a qual se posiciona próxima às setas dos parâmetros amônia, fósforo total, DBO e coliformes termotolerantes no *tripilot*.

A dinâmica sazonal na parte mais à jusante da BHRI foi reforçada pelo processo de seleção de variáveis da RDA, que identificou a classe Cana como uma preditora importante da qualidade da água especificamente no modelo do período seco, enquanto as classes Área Urbanizada, Formação Florestal e Formação Savânica foram selecionadas em ambos os períodos. A maior evidência do impacto da cana-de-açúcar (e potencialmente outras fontes agrícolas associadas) na estação seca pode estar ligada à maior concentração relativa de nutrientes (como o fósforo total, conforme visto na análise de correlação com a classe Agropecuária) e à sua mobilização via fluxo de base ou escoamento superficial mais localizado quando a capacidade de diluição do rio é menor (CHEN et al., 2021, ZHANG et al., 2019). Simultaneamente, a inclusão consistente da Área Urbanizada nos modelos de ambas as estações, associada a parâmetros como amônia e DBO, demonstra a pressão contínua das fontes urbanas sobre a qualidade hídrica da bacia.

4.3 PAINEL DE VISUALIZAÇÃO DA QUALIDADE DA ÁGUA

Para o desenvolvimento do painel de visualização da qualidade da água foram utilizados dados armazenados em arquivos .csv (*comma separated values*) e planilhas Excel. Com as tabelas de dados já formatadas por meio do *Power Query* do Power BI, foi realizada a etapa da modelagem dos dados.

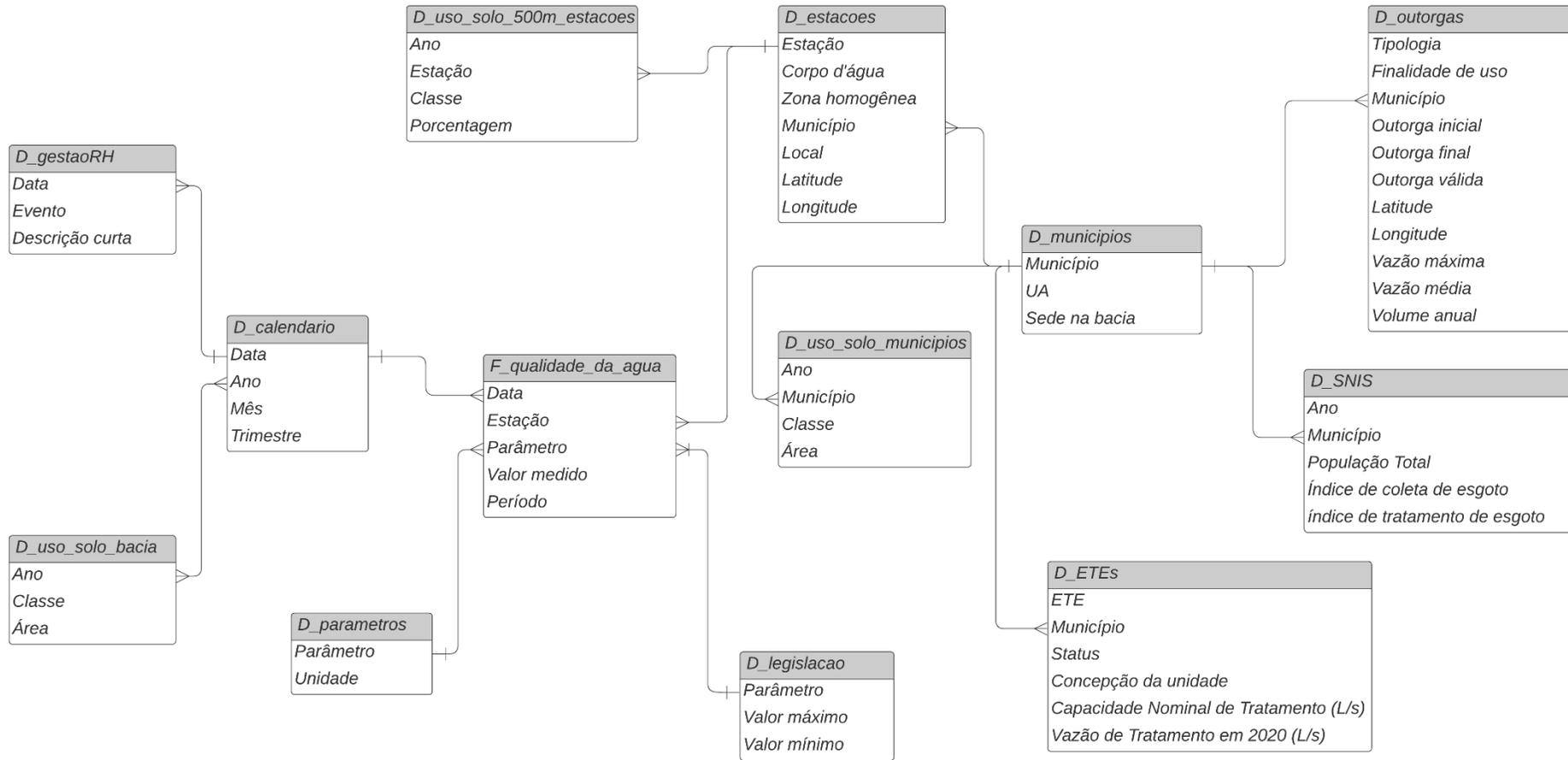
O modelo conceitual de dados foi elaborado para subsidiar a construção do painel de visualização da qualidade da água na BHRI e visa integrar informações provenientes de diferentes fontes temáticas, possibilitando análises multidimensionais sobre a dinâmica ambiental da região. A estrutura do modelo é fundamentada em uma tabela fato central e diversas tabelas dimensão, que fornecem o contexto necessário para a interpretação e visualização dos dados. Foi utilizada uma estrutura híbrida com um Esquema em Estrela bem definido, incorporando a normalização de determinadas dimensões (Esquema Floco de Neve) para maior organização dos dados. O esquema do modelo conceitual aplicado à BHRI é mostrado na Figura 32 a seguir.

A tabela fato principal, F_qualidade_da_agua, consolida os registros das campanhas de monitoramento da qualidade da água, contendo os seguintes atributos: data da coleta, estação de amostragem, parâmetro analisado, valor medido e o período hidrológico correspondente (chuvoso ou seco). Esta tabela está interligada a outras dimensões que contextualizam os dados observados.

Entre as dimensões mais relevantes, destaca-se a D_estacoes, que agrega informações espaciais e administrativas das estações de monitoramento, como o corpo hídrico associado, a zona homogênea na qual a estação de amostragem está inserida, o município, a localização descritiva da estação e as coordenadas geográficas. A tabela D_parametros descreve os parâmetros físico-químicos e biológicos monitorados, juntamente com suas unidades de medida, enquanto a tabela D_legislacao provê os valores máximos e mínimos estabelecidos para cada parâmetro, conforme a Resolução CONAMA nº 357/2005 para águas doces classes 2, permitindo a avaliação de conformidade das medições.

A dimensão temporal é representada pela tabela D_calendario, que desdobra a variável data em componentes como ano, mês e trimestre, contribuindo para análises sazonais e interanuais.

Figura 32 - Modelo conceitual de dados usado para elaboração do painel de visualização da qualidade da água da bacia do rio Ipojuca



Fonte: Elaborado pela autora (2025)

A caracterização do uso e ocupação do solo é abordada em três escalas complementares: no entorno imediato das estações de monitoramento, por meio da tabela D_uso_solo_500m_estacoes; na escala da bacia hidrográfica, representada pela tabela D_uso_solo_bacia; e no nível municipal, por meio da tabela D_uso_solo_municipios. Essas tabelas contêm informações sobre as classes de uso do solo e suas respectivas áreas ou porcentagens, permitindo inferências espaciais quanto à influência antrópica sobre a qualidade da água.

A tabela D_municipios fornece uma chave de ligação espacial entre as diferentes cidades e as unidades de análise da bacia, indicando ainda quais municípios possuem sede dentro dos limites da bacia hidrográfica. Informações relativas à infraestrutura de saneamento básico são armazenadas na tabela D_SNIS, que inclui indicadores como população total, índice de coleta de esgoto e índice de tratamento de esgoto por município e ano, com base em dados do Sistema Nacional de Informações sobre Saneamento. Os dados das estações de tratamento de esgotos que estão situadas em municípios que fazem parte da bacia são reunidos na tabela D_ETEs.

Aspectos legais e institucionais são contemplados pelas tabelas D_outorgas e D_gestaoRH. A primeira reúne informações sobre os atos de outorga de direito de uso da água, incluindo tipologia, finalidade, datas de vigência, localização geográfica, e dados de vazão e volume autorizado. A segunda registra importantes eventos relacionados à gestão de recursos hídricos na bacia principalmente dentro do escopo das Políticas Nacional e Estadual de Recursos Hídricos, identificados por data e descrição resumida.

Com a reunião dessas informações, o modelo conceitual de dados proporciona uma estrutura relacional abrangente, orientada à análise integrada de dados ambientais para a geração de conhecimento aplicado à gestão da qualidade da água na BHRI. O modelo proposto provê o suporte necessário para responder, dentre outras perguntas:

- Como evoluiu a turbidez (ou outro parâmetro) em determinada estação?
- Quais estações apresentam mais violações dos limites legais recomendados?
- Como variam os parâmetros de qualidade entre períodos chuvoso e seco?
- Quais municípios concentram maior uso urbano/agropecuário do solo?
- Qual o volume total de água outorgado por município ou corpo hídrico?
- Quais são os principais usos da água (irrigação, abastecimento humano, indústria etc.) por município, e como eles evoluíram ao longo do tempo?

- Onde estão localizadas espacialmente as outorgas concedidas, e como se distribuem por finalidade?
- Qual a situação do saneamento em municípios com baixa qualidade da água?
- Quais ações de gestão ocorreram em anos com alterações significativas na qualidade da água?

A estruturação eficiente de dados é um desafio recorrente na área de gestão hídrica, e algumas propostas destacam-se por sua flexibilidade e capacidade de integração de múltiplas fontes. Abdallah e Rosenberg (2019) desenvolveram um modelo de dados generalizável chamado *Water Management Data Model* (WaMDaM, ou Modelo de Dados de Gestão de Água), implementado em um banco de dados relacional com uso de metadados contextuais, vocabulários controlados e ferramentas de suporte para facilitar a organização e o armazenamento de dados provenientes de diferentes fontes e modelos. A proposta buscou tornar mais acessível a interação dos usuários com os dados de gestão hídrica. A aplicabilidade do modelo foi demonstrada por meio de cinco casos de uso envolvendo treze conjuntos de dados e modelos voltados à bacia do rio Bear, nos Estados Unidos, permitindo a comparação e seleção de diferentes tipos de dados, redes e elementos de cenário, além da alimentação de modelos com as informações selecionadas. O banco de dados resultante apresentou estrutura flexível e escalável, o que permite incorporar novos dados, modelos e metadados, adaptando-se a diferentes contextos de aplicação.

No contexto brasileiro, uma contribuição relevante é o modelo ACQUA, desenvolvido por Cysne (2004) para representação de recursos hídricos, oferecendo suporte à modelagem de redes complexas por meio da integração de informações hidrológicas, hidrográficas, geométricas e topológicas de objetos georreferenciados. O modelo permitiu representar e analisar seções de canais, tubulações ou rios, identificando locais com estruturas hidráulicas, como as utilizadas para irrigação. Além disso, possibilitou o cálculo de áreas irrigadas, o controle da oferta e demanda hídrica em períodos de seca e o registro de dados coletados por estações de monitoramento. O ACQUA também tinha capacidade de representar fenômenos geográficos relacionados à água, como lagos, poços, reservatórios, barragens, pontos de captação e lançamento, entre outros, integrando tanto dados espaciais quanto semânticos.

Essas propostas reforçam a necessidade de modelos que viabilizem a interoperabilidade entre diferentes tipos de dados, contribuindo para soluções mais adaptáveis em diferentes escalas e realidades. Dessa forma, com base no modelo conceitual de dados estabelecido neste trabalho e nos relacionamentos previamente definidos entre as tabelas, foi desenvolvido um

painel de visualização interativa com o objetivo de facilitar a análise integrada dos dados de qualidade da água na BHRI.

No painel, foram selecionados os tipos de visualizações mais adequados para representar as variáveis ambientais disponíveis, incluindo *boxplots*, gráficos de linha, gráficos de barras, tabelas e mapas georreferenciados. Com o intuito de promover a interatividade e a exploração dinâmica das informações, foram incorporados filtros que permitem aos usuários selecionar os dados com base em critérios específicos, como datas, estações de monitoramento e municípios.

A interface do painel conta, no lado esquerdo, com um navegador de páginas em formato de menu lateral, no qual cada botão representa um atalho para uma das abas temáticas. Ao todo, foram desenvolvidas seis abas, organizadas para refletir diferentes dimensões da análise ambiental. A primeira aba, ilustrada na Figura 33, apresenta uma introdução à ferramenta, destacando seus objetivos, funcionalidades e as principais fontes de dados utilizadas, tais como CPRH, APAC, ANA e SNIS.

Figura 33 - Aba de apresentação do painel de visualização da qualidade da água desenvolvido para a bacia do rio Ipojuca



Fonte: Elaborado pela autora (2025)

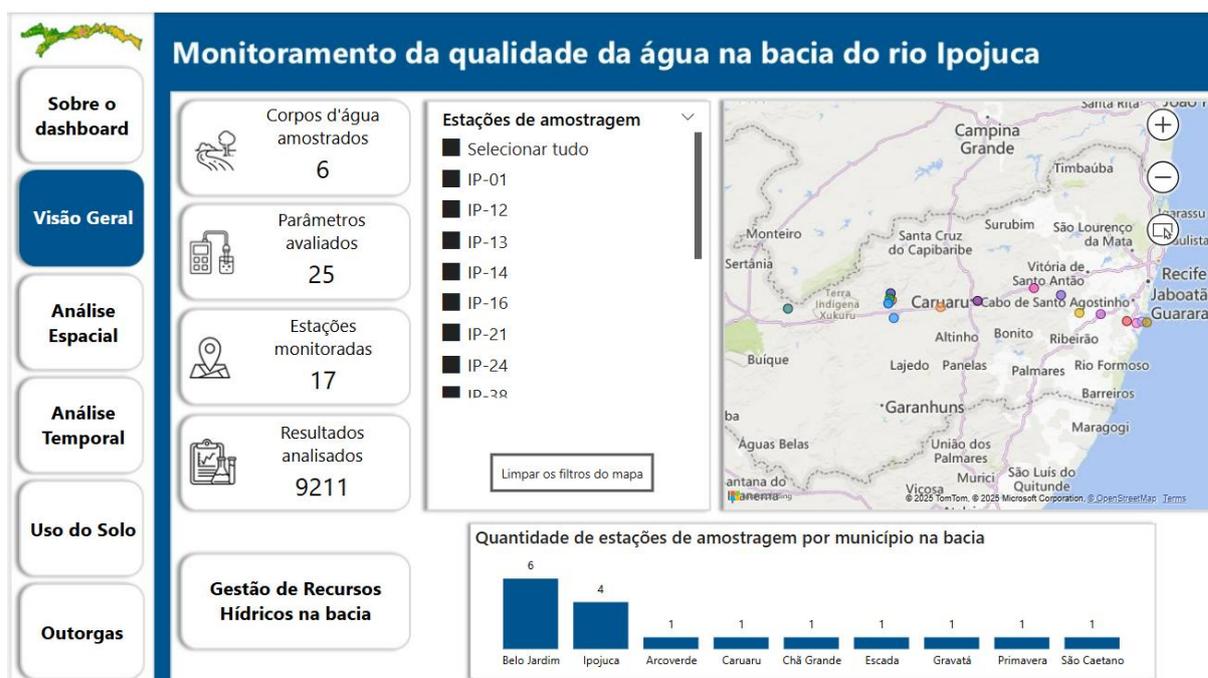
A segunda aba do painel apresenta uma visão geral do monitoramento da qualidade da água na BHRI (Figura 34). Essa seção oferece um panorama das principais métricas de monitoramento, incluindo o número de corpos hídricos amostrados (seis), a quantidade de parâmetros de qualidade da água analisados (25), o total de estações de amostragem

monitoradas (17) e o número de resultados registrados nas campanhas de monitoramento (9211).

Além dessas informações, a aba contempla um gráfico de barras que representa a quantidade de estações de amostragem distribuídas por município, bem como um mapa interativo que localiza espacialmente essas estações. O mapa permite ao usuário aplicar filtros e selecionar quais estações deseja visualizar, facilitando análises mais específicas.

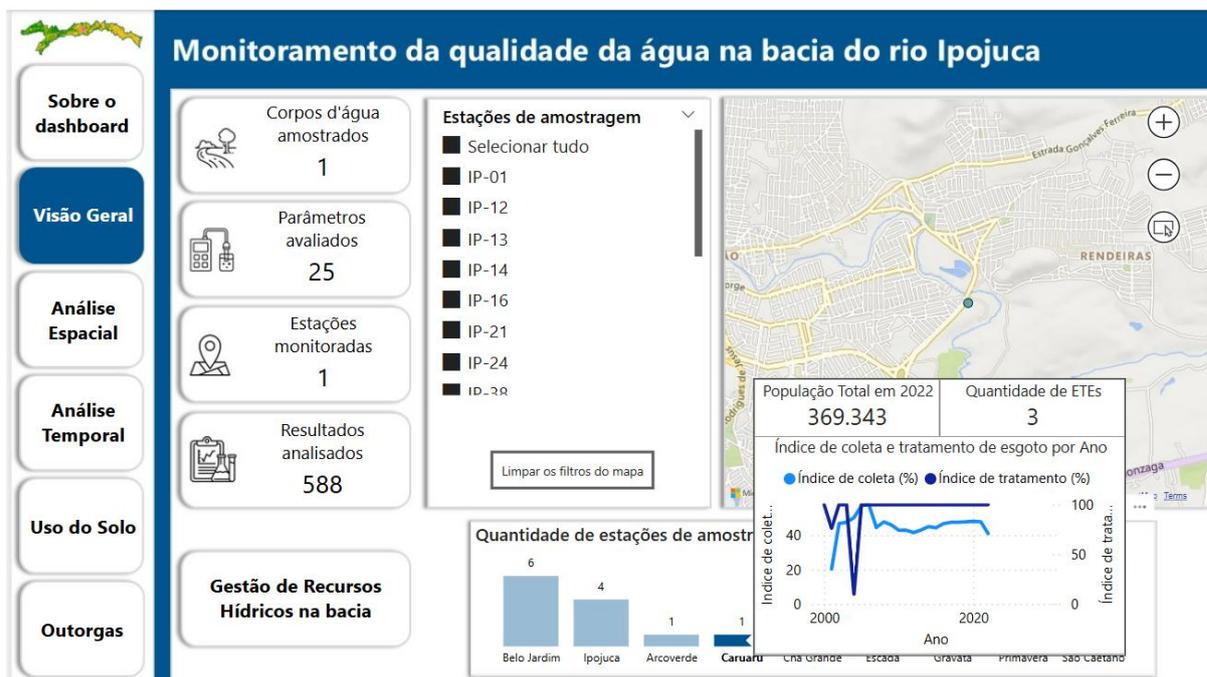
Outro destaque dessa aba é a funcionalidade interativa integrada ao gráfico de distribuição da quantidade de estações por município. Ao posicionar o ponteiro do mouse sobre as barras do gráfico, são exibidas informações demográficas e de saneamento para cada município, como a população total, o número de estações de tratamento de esgotos e a evolução temporal dos índices de coleta e tratamento de esgoto (na Figura 35, temos o exemplo de Caruaru, que contava com uma população total de 369.343 habitantes em 2022, possui três ETEs e apresentou tendência de aumento e estabilização nos níveis de coleta e queda brusca com posterior aumento e estabilização nos níveis de coleta de esgoto ao longo dos anos). Essa integração de dados ambientais e socioeconômicos permite uma análise mais ampla e contextualizada das condições da bacia.

Figura 34 - Aba de visão geral do monitoramento da qualidade da água da bacia do rio Ipojuca



Fonte: Elaborado pela autora (2025)

Figura 35 - Interatividade na aba sobre a visão geral, mostrando informações populacionais e de saneamento



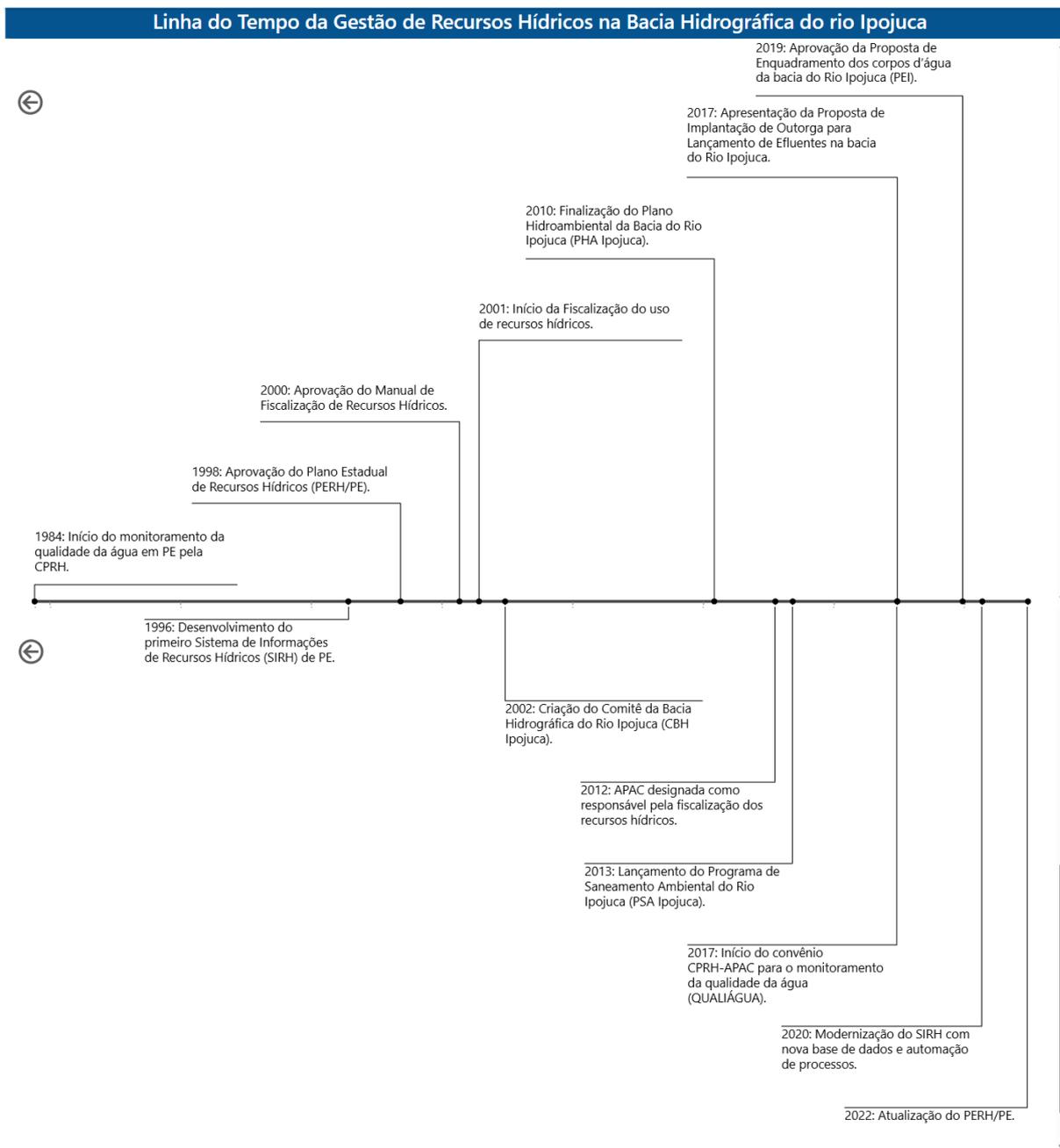
Fonte: Elaborado pela autora (2025)

Ainda na aba *Visão Geral*, há um botão que direciona o usuário para uma página dedicada à linha do tempo da gestão dos recursos hídricos na BHRI, conforme ilustrado na Figura 36. Nessa linha do tempo, são destacados eventos relevantes que marcaram a evolução da governança dos recursos hídricos na região, desde o início do monitoramento da qualidade da água pela CPRH, em 1984, até ações mais recentes, como a atualização do Plano Estadual de Recursos Hídricos de Pernambuco (PERH/PE), em 2022.

Entre os marcos apresentados, destacam-se aqueles diretamente relacionados à BHRI, como a criação do Comitê da Bacia Hidrográfica do Ipojuca (2002), a finalização do Plano Hidroambiental da bacia (2010), o lançamento do Programa de Saneamento Ambiental do rio Ipojuca (PSA Ipojuca) em 2013, além da apresentação da Proposta de Implantação da Outorga para Lançamento de Efluentes (2017) e a Proposta de Enquadramento dos corpos hídricos da bacia (2019).

Essa visualização fornece um panorama cronológico das políticas, planejamentos e ações implementadas, permitindo ao usuário compreender a trajetória institucional e os avanços na gestão da bacia hidrográfica ao longo dos anos.

Figura 36 - Linha do tempo da Gestão de Recursos Hídricos na Bacia do Ipojuca



Fonte: Elaborado pela autora (2025)

A terceira aba do painel, denominada *Análise Espacial*, permite explorar a distribuição espacial dos parâmetros de qualidade da água ao longo da bacia do rio Ipojuca. Nessa seção, o usuário pode selecionar as estações de amostragem, o parâmetro de interesse e o intervalo de tempo a ser analisado. A interface exibe um mapa interativo, onde as estações de amostragem estão georreferenciadas, facilitando a identificação da localização espacial dos pontos monitorados. No mapa, o tamanho das bolhas de cada estação representa a mediana do

parâmetro selecionado, o que possibilita a identificação visual de regiões potencialmente mais críticas quanto à qualidade da água.

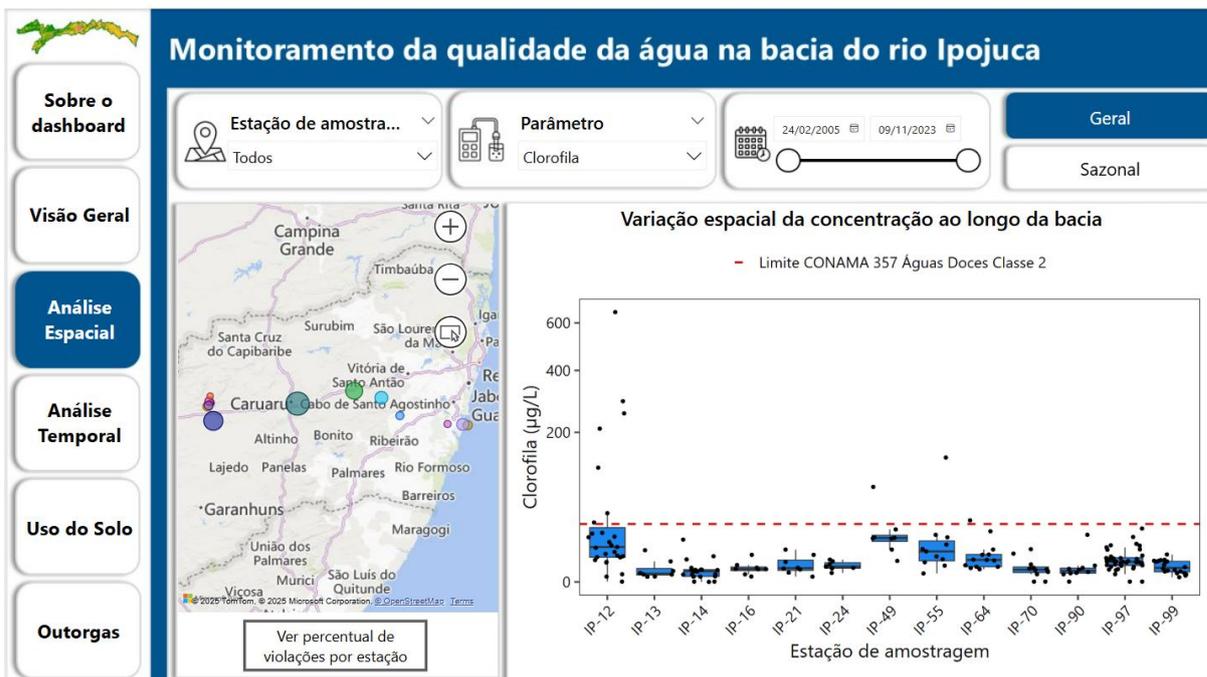
A Figura 37 ilustra a distribuição das concentrações de clorofila-a nas estações de amostragem da bacia onde esse parâmetro foi registrado, comparando os valores obtidos com o limite estabelecido pela Resolução CONAMA nº 357/2005 para corpos hídricos de águas doces de classe 2 (30 µg/L). Observa-se que apenas as estações IP-12, IP-49, IP-55 e IP-64 apresentaram valores acima do limite recomendado.

Um recurso adicional da aba permite acessar uma visualização específica do percentual de violações por estação (Figura 38). Nessa visualização, destaca-se a estação IP-12, onde aproximadamente 26% das amostras analisadas ultrapassaram o limite legal. Essa estação apresentou valores excepcionalmente elevados de clorofila-a, chegando a ultrapassar em mais de 20 vezes o valor de referência, evidenciando condições significativamente fora dos padrões normalmente observados em amostras de qualidade da água.

A clorofila-a é amplamente utilizada como indicador da biomassa algal, e seu aumento pode estar associado à eutrofização, muitas vezes provocada pela entrada excessiva de nutrientes nos corpos hídricos (LE MOAL et al.; 2019; NGAMILE et al., 2025). Assim, essa visualização contribui para uma análise mais detalhada e fundamentada da conformidade com os padrões legais de qualidade da água e potenciais problemas que podem estar associados a determinada estação de amostragem.

A aba também disponibiliza visualização sazonal, permitindo diferenciar os dados entre período chuvoso e seco (Figura 39). A partir desse detalhamento, verifica-se que os valores extremos de clorofila-a na estação IP-12 ocorreram predominantemente durante o período chuvoso, enquanto as medianas foram, em geral, mais elevadas no período seco na maioria das estações analisadas.

Figura 37 - Aba de análise espacial da qualidade da água ao longo da bacia do rio Ipojuca



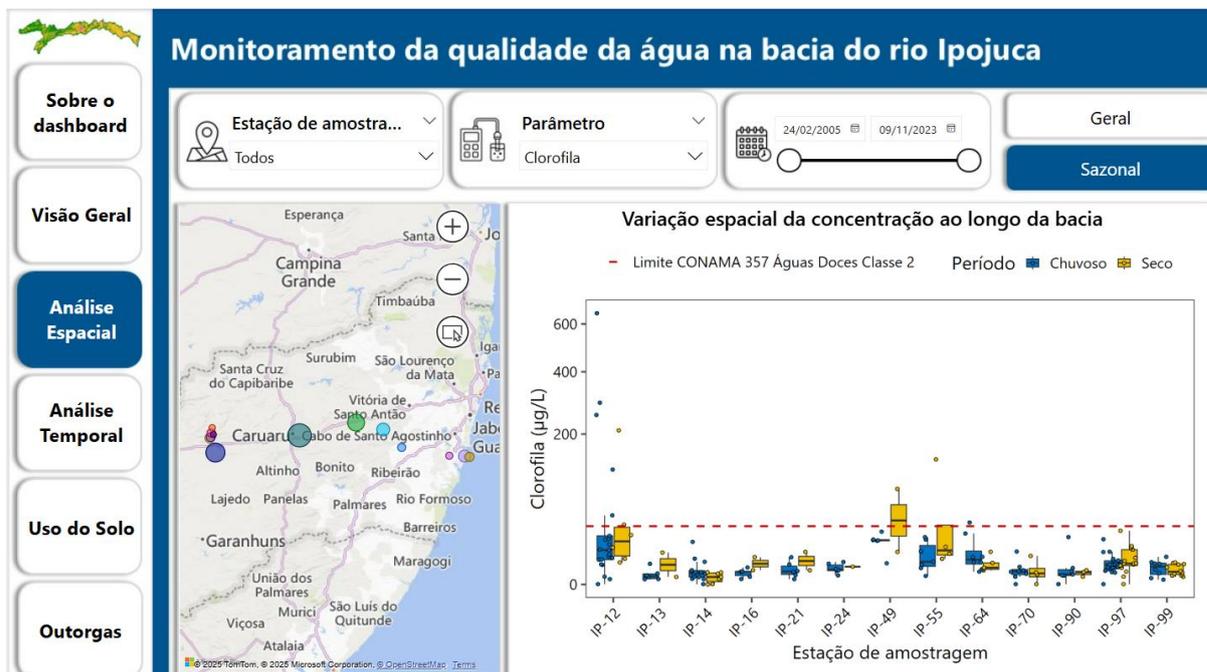
Fonte: Elaborado pela autora (2025)

Figura 38 - Aba de análise espacial da qualidade da água ao longo da bacia do rio Ipojuca com detalhamento da porcentagem de violação aos limites recomendados por estação de amostragem



Fonte: Elaborado pela autora (2025)

Figura 39 - Aba de análise espacial da qualidade da água ao longo da bacia do rio Ipojuca com detalhamento sazonal



Fonte: Elaborado pela autora (2025)

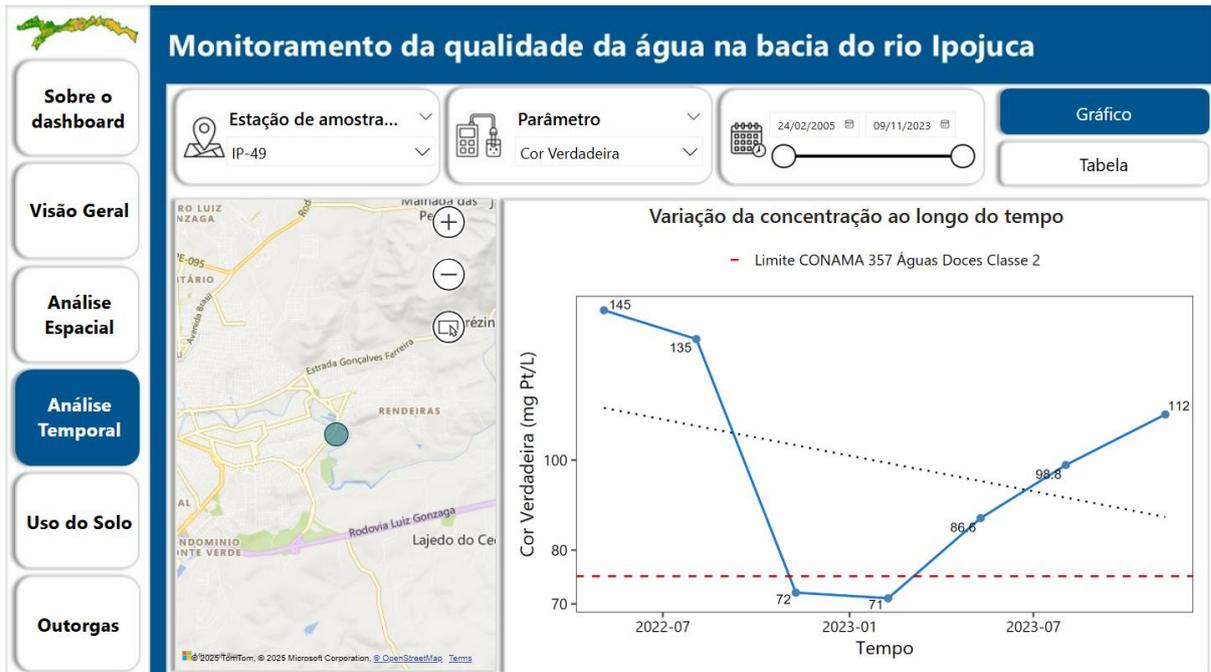
A quarta aba do painel, intitulada *Análise Temporal*, permite a visualização da variação dos parâmetros de qualidade da água ao longo do tempo nas diferentes estações de amostragem da BHRI. Nessa aba, é recomendado que o usuário selecione uma estação específica, o parâmetro desejado, e o intervalo temporal a ser analisado, além de escolher entre as formas de apresentação dos dados: gráfico ou tabela.

A Figura 40 mostra a visualização em formato gráfico, facilitando a identificação de tendências temporais e possíveis episódios de degradação ou melhoria na qualidade da água. No exemplo apresentado, observa-se a variação do parâmetro cor verdadeira na estação IP-49, com destaque para o fato de que esse parâmetro só começou a ser registrado a partir de maio de 2022 nessa estação. O gráfico também exibe uma linha de tendência e o limite estabelecido pela Resolução CONAMA nº 357/2005 para águas doces classe 2 (75 mg Pt/L para cor verdadeira), possibilitando uma análise comparativa em relação ao padrão legal. Essa análise mostra que, no período monitorado, a cor verdadeira esteve em conformidade com a legislação apenas em duas ocasiões. A cor verdadeira se refere à determinação de cor em amostras filtradas, sem turbidez, sendo, portanto, causada por substâncias dissolvidas na água e não por partículas em suspensão (XAVIER et al., 2022).

Já a Figura 41 apresenta os dados em formato de tabela, com a organização das informações por estação de amostragem selecionada e por trimestre. Essa visualização é útil para consultas específicas e análises mais detalhadas, permitindo observar, por exemplo,

valores pontuais, variações entre as estações e ausência de dados, além de mostrar visualmente as violações aos limites recomendados destacadas em laranja. Destaca-se na tabela que foi obtido registro de cor verdadeira uma vez em outubro do ano de 2005, apenas para a estação IP-13.

Figura 40 - Aba de análise temporal da qualidade da água da bacia do rio Ipojuca com detalhamento em forma de gráfico



Fonte: Elaborado pela autora (2025)

Figura 41 - Aba de análise temporal da qualidade da água da bacia do rio Ipojuca com detalhamento em forma de tabela

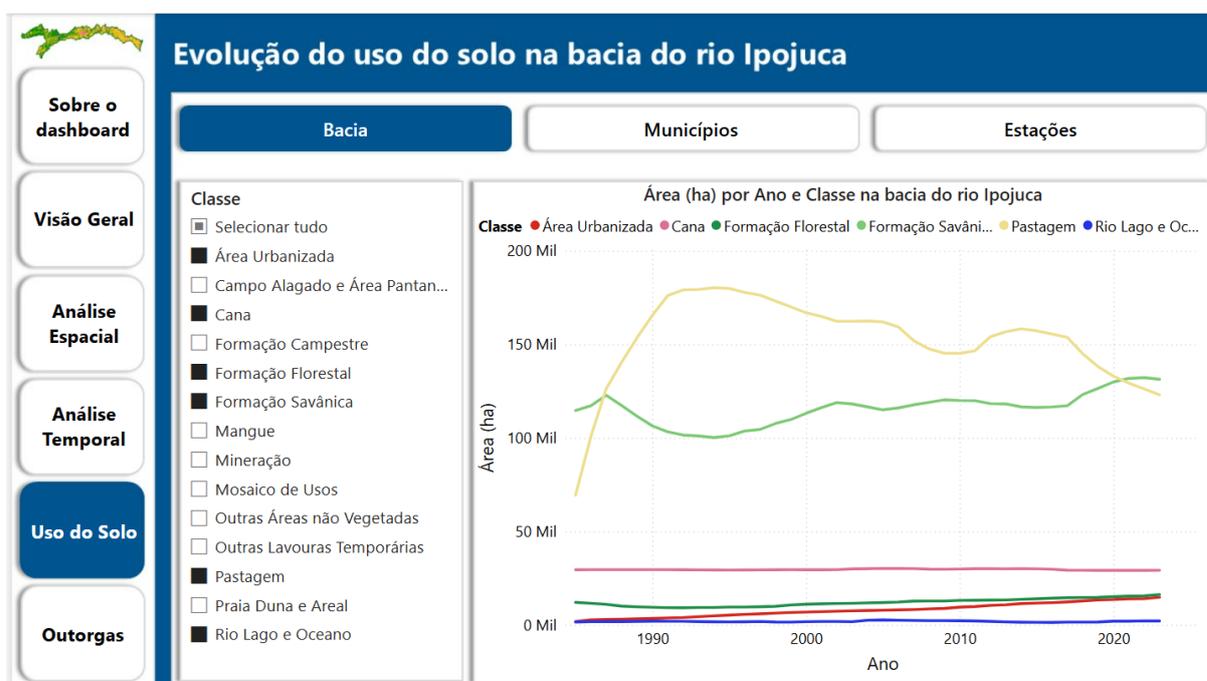


Fonte: Elaborado pela autora (2025)

A quinta aba do painel interativo trata da análise do uso do solo na BHRI e permite explorar as mudanças espaciais e temporais da ocupação do território em diferentes níveis: bacia, municípios e entorno das estações de amostragem.

A Figura 42 apresenta a visualização da evolução do uso do solo para toda a bacia, com dados anuais de 1985 a 2023. As classes de uso do solo podem ser selecionadas no painel lateral, o que permite observar, por exemplo, a predominância de áreas de formação savânica e o aumento progressivo de áreas de pastagem ao longo dos primeiros anos da série, seguido de uma redução menos acentuada até os dias atuais. Já as áreas urbanizadas, embora representem uma menor proporção da bacia, apresentam uma tendência de crescimento ao longo das décadas.

Figura 42 - Aba de análise do uso do solo com detalhamento para a bacia do rio Ipojuca

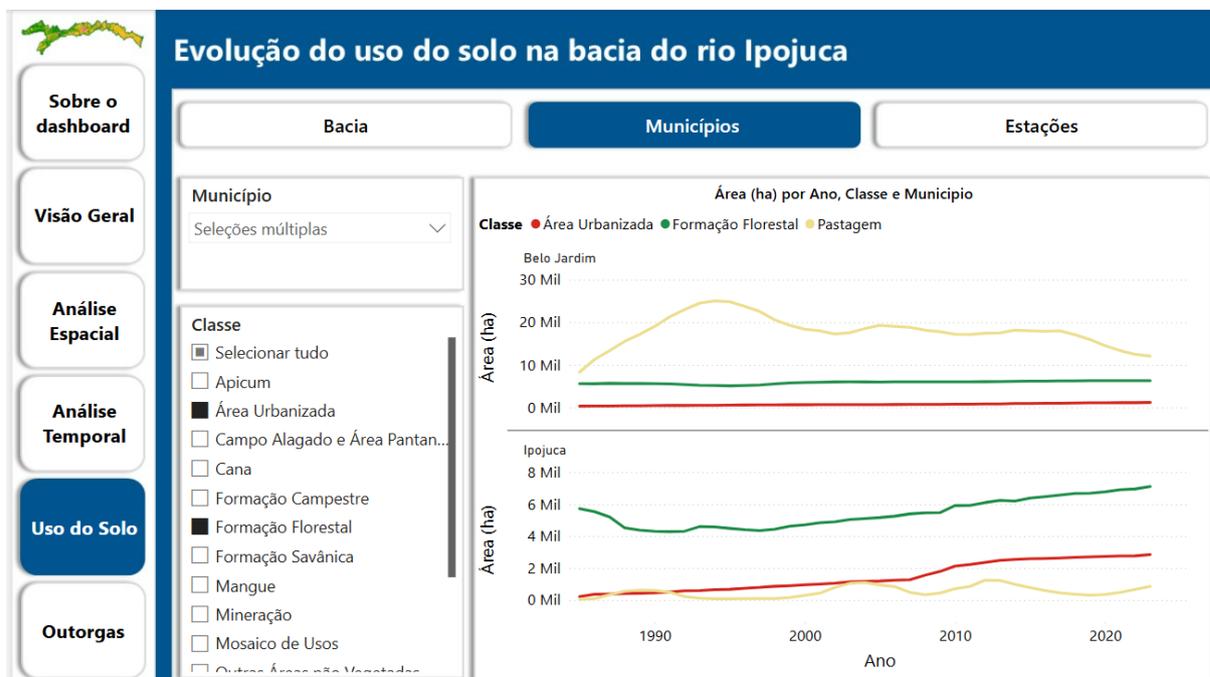


Fonte: Elaborado pela autora (2025)

A Figura 43 detalha essa análise para os municípios inseridos na bacia, possibilitando a comparação entre eles. Nesse exemplo, é possível observar que os municípios selecionados (Belo Jardim e Ipojuca) apresentam padrões distintos de cobertura do solo. Belo Jardim é caracterizado por grande extensão de pastagens, enquanto Ipojuca é caracterizado por formação florestal. Em Belo Jardim, houve um crescimento seguido por uma queda nas áreas de pastagens, que, no entanto, permaneceram em níveis elevados; já em Ipojuca, a tendência é de oscilação mais discreta, com essas áreas se mantendo em níveis baixos. O crescimento urbano foi mais intenso em Ipojuca, principalmente a partir dos anos 2010, enquanto em Belo Jardim

foi mais gradual e discreto. A seleção simultânea de diferentes municípios e classes de uso do solo permite análises comparativas mais aprofundadas.

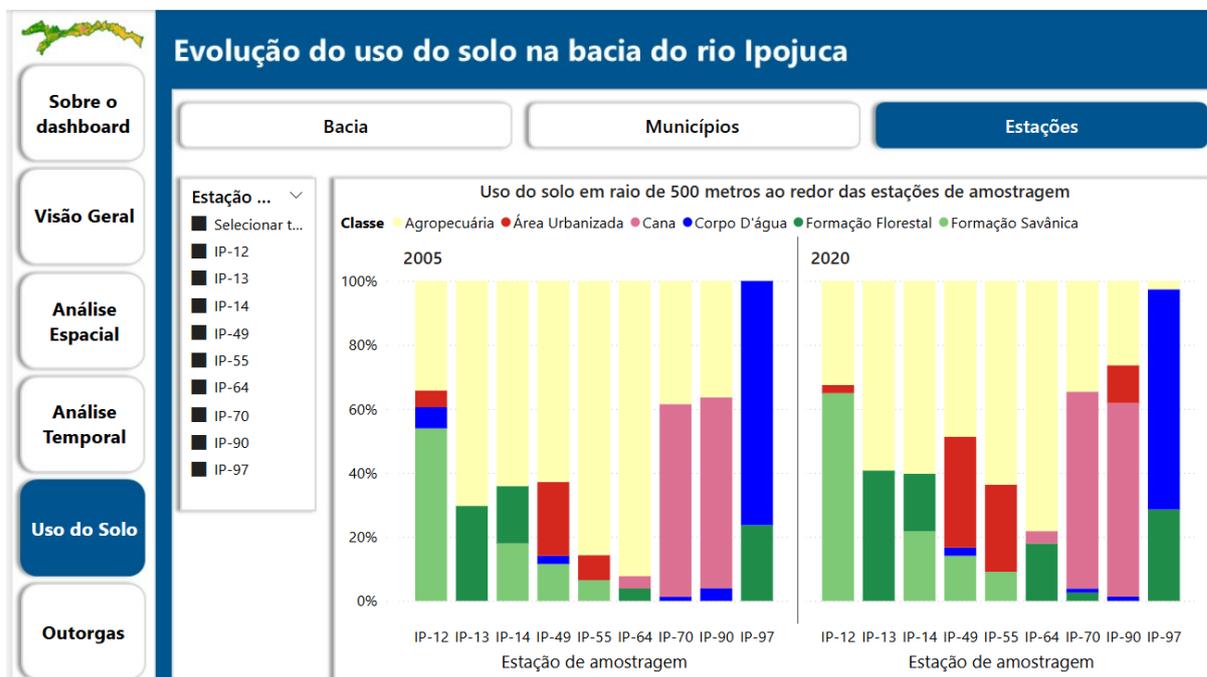
Figura 43 - Aba de análise do uso do solo com detalhamento para os municípios inseridos na bacia



Fonte: Elaborado pela autora (2025)

Por fim, a Figura 44 mostra a distribuição do uso do solo em um raio de 500 metros ao redor das estações de amostragem, permitindo avaliar a influência do entorno imediato sobre os dados de qualidade da água. O gráfico comparativo entre os anos de 2005 e 2020 permite identificar alterações significativas nas proporções de cobertura vegetal, áreas agrícolas e urbanizadas próximas às estações. Observa-se, por exemplo, o aumento de áreas urbanizadas no entorno das estações IP-49, IP-55 e IP-90, o que pode contribuir para pressões antrópicas locais que afetam diretamente os corpos hídricos monitorados. Nota-se também o aumento relevante da cobertura de formação florestal no entorno das estações IP-13 e IP-64. Ao passar o ponteiro do *mouse* sobre as barras, são mostradas ao usuário as porcentagens exatas de cada classe para cada estação.

Figura 44 - Aba de análise do uso do solo com detalhamento para o raio de 500 metros ao redor das estações de amostragem



Fonte: Elaborado pela autora (2025)

A sexta e última aba do painel corresponde à visualização das outorgas concedidas na BHRI, apresentada na Figura 45. Foram identificadas quatro tipologias de outorga na base de dados disponibilizadas pela Agência Nacional de Águas e Saneamento Básico (ANA): barragem, captação, lançamento e ponto de referência. As finalidades de uso registradas por tipologia foram:

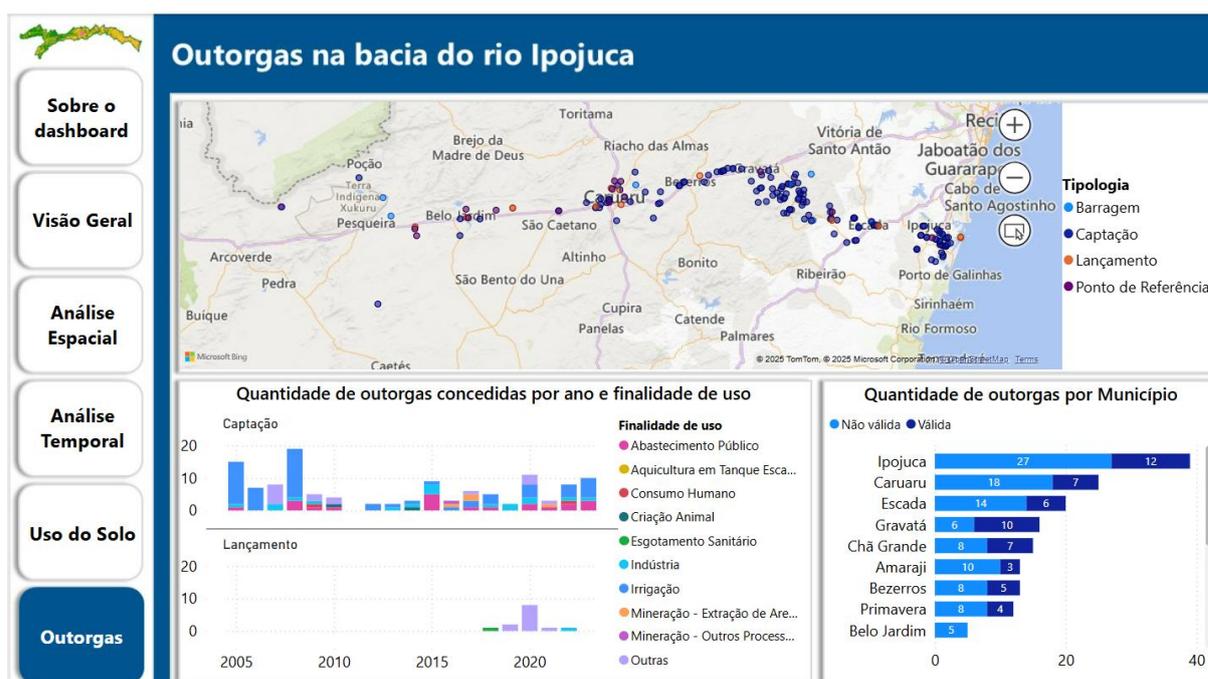
- Barragem: Aproveitamento Hidroelétrico; Reservatório/Barramento/Regularização de Vazões para Usos Múltiplos.
- Captação: Abastecimento Público; Aquicultura em Tanque Escavado; Consumo Humano; Criação Animal; Indústria; Irrigação; Mineração - Extração de Areia/Cascalho em Leito de Rio; Mineração - Outros Processos Extrativos; Outras.
- Lançamento: Esgotamento Sanitário; Indústria; Outras.
- Ponto de referência: Obras Hidráulicas; Serviços.

O mapa interativo mostra os pontos georreferenciados identificados pelas tipologias, o que facilita o reconhecimento de áreas com maior concentração de outorgas. Nota-se que a maior parte das outorgas está concentrada na região que se estende do município de Caruaru em direção à jusante do rio.

Abaixo do mapa, o gráfico à esquerda apresenta a quantidade de outorgas concedidas por ano e por finalidade de uso, considerando apenas as tipologias de captação e lançamento com data de início entre 2005 e 2023. Essa seleção é especialmente relevante para comparação com os dados de qualidade da água disponíveis no mesmo intervalo. Irrigação, indústria e abastecimento público são as finalidades predominantes nas outorgas de captação, sendo o município de Ipojuca destaque em 2005, com 13 outorgas concedidas para irrigação. Todas estiveram vigentes até 2010, período durante o qual as estações de amostragem IP-90 e IP-95, localizadas em Ipojuca, apresentaram níveis acima de 90% de violações ao limite de fósforo total estabelecido pela Resolução CONAMA 357/2005. Dentre as 25 outorgas registradas para o município de Caruaru, nove são destinadas a indústrias, oito não têm finalidade especificada (outras) e duas são para irrigação. Como visto anteriormente na seção 4.1.1, a estação de amostragem localizada nesse município (IP-49) apresenta a pior qualidade da água dentre os locais analisados.

Complementando a análise, o gráfico à direita resume a quantidade total de outorgas por município, diferenciando entre as vigentes e vencidas. O gráfico de barras horizontais facilita a comparação entre os municípios, revelando que Ipojuca concentra o maior número de outorgas vigentes, seguido por Gravatá, Caruaru e Chã Grande. A presença expressiva de outorgas vencidas também é um aspecto relevante, podendo apontar para desafios na fiscalização e na atualização dos registros de uso da água na região.

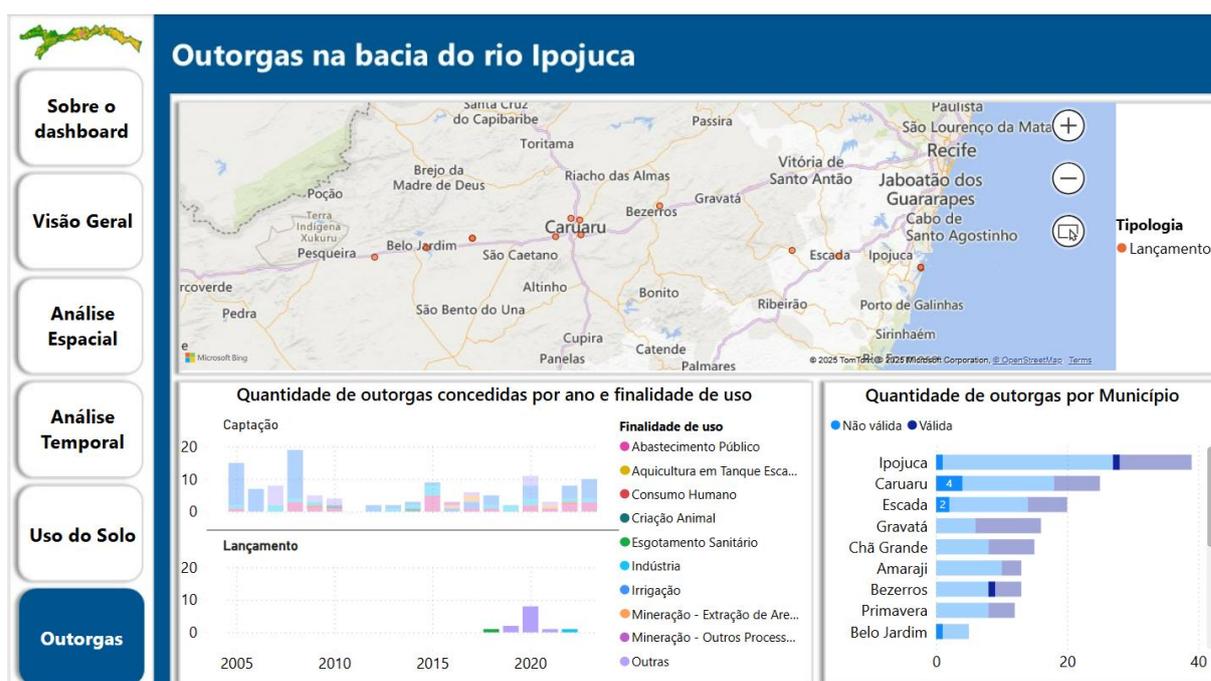
Figura 45 - Aba de análise das outorgas concedidas na bacia do rio Ipojuca



Fonte: Elaborado pela autora (2025)

A Figura 46 aprofunda essa análise ao destacar exclusivamente a tipologia de lançamento de efluentes, permitindo uma visualização mais clara dos locais onde essa atividade está concentrada. No mapa, os pontos de lançamento são destacados em laranja, enquanto os gráficos ilustram a evolução temporal da concessão dessas outorgas por finalidade de uso e sua distribuição por município. Caruaru, Ipojuca e Escada lideram em número de outorgas dessa tipologia, evidenciando áreas de grande pressão sobre a qualidade da água na bacia. No entanto, dentre os municípios com esse tipo de outorga, apenas Ipojuca e Bezerros ainda apresentam outorgas vigentes. A concessão de outorgas de lançamento teve início apenas em 2018, ano posterior à apresentação da Proposta de Implantação de Outorga para Lançamento de Efluentes na bacia do rio Ipojuca, conforme apresentado na linha do tempo da gestão de recursos hídricos da bacia.

Figura 46 - Aba de análise das outorgas concedidas na bacia do rio Ipojuca, com destaque para a tipologia lançamento de efluentes



Fonte: Elaborado pela autora (2025)

Informações extraídas por meio de visualizações eficientes e integradas contribuem significativamente para o entendimento da distribuição espacial e temporal da qualidade da água em bacias hidrográficas. Nesse contexto, a aplicação de tecnologias de *Business Intelligence* tem se mostrado uma abordagem promissora em diversos estudos.

Sinha et al. (2018), por exemplo, utilizaram técnicas de BI para organizar dados de águas subterrâneas e analisar a variação espacial e temporal da salinidade na região de Barmer, Rajasthan (Índia). A análise revelou que, entre 30 localidades avaliadas, apenas sete

apresentaram água potável, enquanto as demais excederam os limites permitidos de sólidos totais dissolvidos, conforme as normas usadas como referência. Considerando o caráter árido da região, os autores recomendaram a adoção de tecnologias como dessalinização e captação de água da chuva para ampliar a disponibilidade hídrica. O estudo também enfatizou a importância de promover o uso da Ciência de Dados no setor hídrico, especialmente para apoiar a gestão de águas subterrâneas, que são menos visíveis e mais difíceis de monitorar.

De forma complementar, Uronen (2018) avaliou o potencial do BI e do *Business Analytics* (BA) no setor hídrico da Finlândia, destacando entraves institucionais e técnicos. Os principais desafios identificados incluíram a infraestrutura obsoleta, a baixa integração entre bases de dados, a ausência de informações em tempo real e a má qualidade dos dados disponíveis. O autor apontou ainda a necessidade de uma transformação cultural nas instituições, com a incorporação de profissionais com perfil analítico e a implementação de sistemas básicos de BI, capazes de integrar diferentes fontes de informação e fomentar uma cultura de decisões orientadas por dados.

No Brasil, Almeida et al. (2019) propuseram um Sistema de Apoio à Decisão (SAD) voltado à fiscalização do uso da água, baseado em tecnologias de BI e dados geoespaciais, com interfaces adaptadas para *Web* e dispositivos móveis. O sistema foi implementado na Agência Executiva de Gestão das Águas da Paraíba (AESA). Os autores propuseram como investigação futura um estudo do impacto que suas funcionalidades trazem na fiscalização, e a integração do SAD de fiscalização do uso da água com o SAD de Outorga, resultando numa visão sistêmica destas duas áreas da gestão dos Recursos Hídricos.

Essas experiências reforçam a relevância de sistemas interativos, baseados em BI, como instrumentos para fortalecer a gestão hídrica. A proposta de painel de qualidade da água desenvolvida nesta pesquisa pode ser integrada a sistemas semelhantes, como os relacionados a licenciamento ambiental, vistorias, denúncias e multas. Isso potencializa o suporte ao processo decisório, à fiscalização ambiental e à identificação de áreas críticas de degradação.

Além disso, Ponce Romero et al. (2017) ressaltam que a incorporação de dados externos, como informações ambientais, demográficas e meteorológicas, pode ampliar a compreensão das dinâmicas ambientais e revelar relações não percebidas inicialmente. A análise integrada desses dados, aliada a ferramentas de visualização acessíveis aos diferentes públicos-alvo, amplia o potencial analítico dos painéis e contribui para uma comunicação mais eficaz dos resultados.

5 CONCLUSÕES E RECOMENDAÇÕES

Nesta seção, são apresentadas as conclusões obtidas após a análise e discussão dos resultados da pesquisa, assim como as recomendações para a gestão da qualidade da água na área de estudo.

5.1 CONCLUSÕES

Esta pesquisa confirmou a hipótese de que a aplicação de ferramentas da Ciência de Dados contribuiu efetivamente para o desenvolvimento de uma metodologia de análise integrada da qualidade da água em bacias hidrográficas. Utilizando a bacia do rio Ipojuca como estudo de caso, a abordagem proposta, estruturada em quatro etapas, gerou resultados significativos tanto para a compreensão da dinâmica espaço-temporal da qualidade da água quanto para o suporte à sua gestão.

Na primeira etapa, voltada à avaliação da evolução espaço-temporal e das tendências de alteração da qualidade da água, identificou-se que os parâmetros com maior número de violações à Resolução CONAMA 357/2005 em todas as estações foram oxigênio dissolvido, DBO, fósforo total e coliformes termotolerantes, todos fortemente associados à poluição. A estação IP-49, localizada em Caruaru, apresentou a maior frequência de violações aos limites recomendados, seguida pelas estações IP-55 (em Gravatá) e IP-90 (em Ipojuca, à jusante da Usina Ipojuca). Observou-se também que as estações IP-70 (em Primavera, à jusante da Usina União Indústria) e IP-90, situadas na UA4, apresentaram variações sazonais significativas para quase todos os parâmetros analisados, evidenciando a influência da dinâmica hidrológica local sobre a qualidade da água. A análise de componentes principais reforçou a importância de variáveis ligadas à poluição antrópica, como fósforo total, DBO e amônia, além de parâmetros relacionados à composição mineral, como condutividade, salinidade e sólidos dissolvidos totais. Já a análise de ponto de mudança revelou alterações significativas nas séries temporais de qualidade da água em quase todas as estações de amostragem, enquanto o Teste de Mann-Kendall identificou tendências de aumento de coliformes termotolerantes, fósforo total e DBO em determinadas estações, refletindo possíveis efeitos de mudanças no uso do solo, clima ou práticas de manejo ao longo do tempo.

A segunda etapa focou na análise do uso e ocupação do solo e sua relação com a qualidade da água. A principal mudança observada foi a redução de cerca de 5% nas áreas destinadas à agropecuária dentro da bacia. No entorno das estações, entre 2005 e 2019, houve

aumento expressivo da Formação Savânica em IP-12 e IP-14, e da área urbanizada em IP-55 e IP-90. Já a Formação Florestal cresceu em IP-64 e IP-13. A correlação entre o uso do solo e os parâmetros de qualidade revelou que áreas florestais estão associadas à melhoria na qualidade da água, com correlações negativas com DBO, amônia, salinidade e temperatura, enquanto áreas agropecuárias apresentaram correlação positiva com o fósforo total. A análise de redundância indicou que o uso e ocupação do solo explicam parcela significativa da variação da qualidade da água, com destaque para o período seco, quando os efeitos do uso do solo se mostraram mais pronunciados.

Na terceira etapa, foi elaborado um modelo conceitual de dados composto por tabelas relacionais, capaz de integrar informações relevantes para a gestão da qualidade da água na bacia do rio Ipojuca. Este modelo estabelece as bases para a construção de um painel analítico, favorecendo a organização, consulta e exploração das informações de maneira estruturada e flexível.

A quarta e última etapa envolveu o desenvolvimento de um painel interativo de visualização e integração dos dados. Foram implementadas visualizações variadas, como *boxplots*, gráficos de linha e barras, além de mapas georreferenciados, distribuídos em seis abas temáticas e acompanhadas de filtros dinâmicos, que permitem análises personalizadas conforme datas, estações de monitoramento e municípios. O painel evidenciou a abrangência do monitoramento na bacia: seis corpos hídricos monitorados, 25 parâmetros analisados, 17 estações de amostragem e mais de 9 mil registros. Também trouxe informações relevantes sobre a população e infraestrutura de saneamento, como a presença de três estações de tratamento de esgoto em Caruaru e o histórico das outorgas de captação e lançamento de efluentes, revelando os municípios mais pressionados em termos de número de outorgas para lançamento de efluentes, como Caruaru, Ipojuca e Escada.

Dessa forma, conclui-se que a metodologia proposta oferece uma alternativa eficaz para a análise integrada da qualidade da água, associando dados de diversas fontes por meio de um modelo consistente e ferramentas interativas de visualização. Os resultados demonstram o potencial de replicabilidade da abordagem em outras bacias hidrográficas, contribuindo para o melhoramento da gestão da qualidade da água com base em evidências e inteligência analítica.

5.2 RECOMENDAÇÕES

Com base nos resultados e observações feitas no andamento desta pesquisa, foi possível propor algumas recomendações voltadas ao aprimoramento da gestão da qualidade da água na

bacia do rio Ipojuca. A primeira recomendação consiste na ampliação e integração das bases de dados atualmente utilizadas no modelo conceitual. A incorporação de informações de processos de licenciamento e fiscalização ambiental, bem como de dados de qualidade da água dos reservatórios existentes na bacia, permitiria uma compreensão ainda mais abrangente das pressões exercidas sobre os corpos hídricos, viabilizando análises mais precisas sobre as origens e intensidades das cargas poluidoras.

Também se recomenda a inserção dos dados do Índice de Qualidade da Água e do Índice de Estado Trófico ao longo da série histórica de monitoramento no modelo conceitual, sempre que possível. A inclusão desses índices pode contribuir para uma avaliação mais objetiva e padronizada da condição ambiental dos corpos hídricos, facilitando comparações temporais e espaciais.

Adicionalmente, embora a CPRH disponibilize regularmente boletins mensais e relatórios anuais de monitoramento da qualidade da água, a análise das causas por trás das alterações observadas ao longo do tempo requer a criação/atualização de um inventário histórico das fontes poluidoras locais. Esse inventário deveria incluir a geolocalização das fontes, suas respectivas cargas poluidoras e o histórico de operação, o que permitiria cruzamentos com as séries temporais de dados de qualidade da água. Recomenda-se, nesse sentido, a atualização do diagrama unifilar das fontes poluidoras, consolidando informações em uma estrutura integrada, passível de ser incorporada a sistemas de apoio à decisão.

Ainda nesse contexto, destaca-se a importância da sistematização das principais atividades econômicas desenvolvidas na bacia, com ênfase naquelas que potencialmente impactam a qualidade da água, como a agropecuária, a indústria têxtil e a disposição inadequada de resíduos. A integração dessas informações ao modelo conceitual e ao painel pode aprimorar a identificação de áreas críticas e subsidiar ações mais direcionadas de fiscalização e mitigação de impactos. Isso seria uma contribuição importante, uma vez que os resultados desta pesquisa demonstraram que fontes de poluição pontual respondem por parcela relevante da variação da qualidade da água na bacia do rio Ipojuca.

Considerando que a região de Caruaru apresenta relevância no contexto da contaminação por cianotoxinas, com potenciais implicações para a saúde pública, sugere-se incluir no painel interativo uma camada temática que relacione dados de qualidade da água com indicadores de saúde pública, especialmente em áreas suscetíveis a eventos de floração de cianobactérias. Tal abordagem pode fortalecer o uso da ferramenta como apoio à gestão intersetorial entre meio ambiente e saúde.

Além disso, é importante expandir a rede de monitoramento, incluindo pontos adicionais que representem trechos do rio Ipojuca situados fora das zonas já reconhecidas por concentrarem elevadas pressões antrópicas, e contemplar trechos à jusante das novas estações de tratamento de esgoto implantadas pelo PSA Ipojuca. Essa abordagem possibilitaria avaliar a qualidade da água em regiões ainda pouco estudadas, além de permitir determinar o impacto do tratamento dos esgotos no estado do rio. De forma complementar, recomenda-se o aumento do número de estações fluviométricas na bacia, pois existem apenas quatro, com especial atenção à coincidência geográfica com os pontos de monitoramento da qualidade da água. Isso favoreceria análises integradas entre os regimes de vazão e os parâmetros físico-químicos e biológicos monitorados.

Os resultados também reforçaram a necessidade de se promover práticas de uso do solo mais sustentáveis, especialmente diante da constatação de que a presença de vegetação nativa, notadamente da Formação Florestal, está associada a melhores condições da qualidade da água. A proteção e recuperação da vegetação nativa em áreas ribeirinhas, como a própria Área de Preservação Permanente (APP) já estabelecida no Código Florestal brasileiro, é recomendada como estratégia para manter as características físicas dos canais fluviais, reduzir a erosão das margens e a sedimentação dos cursos d'água na bacia do rio Ipojuca.

Para estudos futuros, recomenda-se a ampliação da metodologia proposta para outras bacias hidrográficas com diferentes características climáticas, hidrológicas e de uso do solo, a fim de avaliar sua aplicabilidade e assertividade em contextos variados. Além disso, a incorporação de séries temporais de dados provenientes de sensoriamento remoto pode aprimorar a análise espaço-temporal da qualidade da água. Sugere-se também o uso de técnicas avançadas de Aprendizado de Máquina para elaborar modelos preditivos e fortalecer o suporte à tomada de decisão na gestão dos recursos hídricos.

REFERÊNCIAS

- ABBOTT, B. W. *et al.* Human domination of the global water cycle absent from depictions and perceptions. *Nature Geoscience*, v. 12, p. 533–540, 2019. <https://doi.org/10.1038/s41561-019-0374-y>
- ABDALLAH, A. M.; ROSENBERG, D. E. A data model to manage data for water resources systems modeling. *Environmental Modelling & Software*, v. 115, p. 113–127, 2019. <https://doi.org/10.1016/j.envsoft.2019.02.005>
- ADAMOVICH, B. V. *et al.* Relations between variations in the lake bacterioplankton abundance and the lake trophic state: Evidence from the 20-year monitoring. *Ecological Indicators*, v. 97, p. 120–129, 2019.
- AGRIZZI, D. V. *et al.* Qualidade da água de nascentes do Assentamento Paraíso. *Engenharia Sanitária e Ambiental*, v. 23, n. 3, p. 557–568, 2018. <https://doi.org/10.1590/S1413-41522018150701>
- ALMEIDA, D. *et al.* Uso de Business Intelligence na Gestão de Recursos Hídricos: o caso da Fiscalização do Uso da Água. In: WORKSHOP DE COMPUTAÇÃO APLICADA À GESTÃO DO MEIO AMBIENTE E RECURSOS NATURAIS (WCAMA), 10., 2019, Belém. *Anais...* Porto Alegre: Sociedade Brasileira de Computação, 2019. p. 11–20. ISSN 2595-6124. <https://doi.org/10.5753/wcama.2019.6415>.
- ALMEIDA, T. F. S. *et al.* Zoneamento pluviométrico na bacia hidrográfica do rio Ipojuca (PE) e sua relação com o escoamento dos cursos hídricos. In: EDITORA POISSON (Org.). *Recursos Hídricos em Foco – Volume 1*. Belo Horizonte: Poisson, 2019. p. 126–133.
- ANA – AGÊNCIA NACIONAL DE ÁGUAS. *Conjuntura dos recursos hídricos no Brasil 2017: relatório pleno*. Brasília: ANA, 2017.
- ANA – AGÊNCIA NACIONAL DE ÁGUAS. *Conjuntura dos recursos hídricos no Brasil 2019: informe anual*. Brasília: ANA, 2019.
- ANA – AGÊNCIA NACIONAL DE ÁGUAS E SANEAMENTO BÁSICO. *Conjuntura dos recursos hídricos no Brasil 2021: relatório pleno*. Brasília: ANA, 2022.
- ANDEM, A. B. *et al.* Seasonal variation in limnological characteristics of the upper, middle and lower reaches of Calabar River, Southern Nigeria. *International Journal of Biological Sciences*, v. 1, n. 2, p. 52–60, 2014.
- ANDRADE, V. S. *et al.* Influence of rainfall and seasonal crop practices on nutrient and pesticide runoff from soybean dominated agricultural areas in Pampean streams, Argentina. *Science of the Total Environment*, v. 788, p. 147676, 2021.
- APAC – AGÊNCIA PERNAMBUCANA DE ÁGUAS E CLIMA. *Bacias hidrográficas*. 2023. Disponível em: <https://www.apac.pe.gov.br/bacias-hidrograficas>. Acesso em: 10 dez. 2023.

APAC – AGÊNCIA PERNAMBUCANA DE ÁGUAS E CLIMA. *Plano hidroambiental da bacia hidrográfica do Rio Una e Grupos de Bacias Litorâneas GL4 e GL5*. Recife: APAC, 2019.

APAC – AGÊNCIA PERNAMBUCANA DE ÁGUAS E CLIMA. *Relatório de situação de recursos hídricos do Estado de Pernambuco 2011/2012*. Recife: APAC, 2013.

APAC – AGÊNCIA PERNAMBUCANA DE ÁGUAS E CLIMA. *Sobre APAC*. 2017. Disponível em: <https://www.apac.pe.gov.br/institucional>. Acesso em: 10 dez. 2023.

ASHERI ARNON, T.; EZRA, S.; FISHBAIN, B. Water characterization and early contamination detection in highly varying stochastic background water, based on Machine Learning methodology for processing real-time UV-Spectrophotometry. *Water Research*, v. 155, p. 333–342, 2019. <https://doi.org/10.1016/j.watres.2019.02.027>

ATHANASIADIS, I. N.; MITKAS, P. A. Knowledge discovery for operational decision support in air quality management. *Journal of Environmental Informatics*, v. 9, n. 2, p. 100–107, 2007. <https://doi.org/10.3808/jei.200700091>

AWS – AMAZON WEB SERVICES. *O que é OLAP (processamento analítico online)?* 2024. Disponível em: <https://aws.amazon.com/pt/what-is/olap/#:~:text=O%20processamento%20anal%C3%ADtico%20on%2Dline>. Acesso em: 15 dez. 2023.

AYED, L. *et al.* Hybrid coagulation-flocculation and anaerobic-aerobic biological treatment for industrial textile wastewater: pilot case study. *The Journal of The Textile Institute*, v. 112, n. 2, p. 200–206, 2021.

BARROS, A. M. L. *Modelagem da poluição pontual e difusa: aplicação do modelo Moneris à Bacia hidrográfica do rio Ipojuca, Pernambuco*. 2008. Dissertação (Mestrado em Engenharia Civil) – Universidade Federal de Pernambuco, Recife, 2008.

BERNIER, E.; BÉDARD, Y.; BADARD, T.; HUBERT, F. UMapIT© (Unrestricted Mapping Interactive Tool): Merging the datacube paradigm with an occurrence-based approach to support on-demand web mapping. In: PETERSON, M. P. (ed.). *International Perspectives on Maps and the Internet*. Berlin; Heidelberg: Springer, 2008. (Lecture Notes in Geoinformation and Cartography). https://doi.org/10.1007/978-3-540-72029-4_13

BID – BANCO INTERAMERICANO DE DESENVOLVIMENTO. *Projeto de Saneamento Ambiental da Bacia do rio Ipojuca, PSA-IPOJUCA*. Brasil, 2012.

BRAGA, B. *et al.* *Introdução à Engenharia Ambiental: o desafio do desenvolvimento sustentável*. São Paulo: Prentice Hall, 2005. 318 p.

BRASIL. *Lei nº 14.026, de 15 de julho de 2020*. Novo Marco Legal do Saneamento Básico. 2020. Disponível em: http://www.planalto.gov.br/ccivil_03/_ato2019-2022/2020/lei/L14026.htm. Acesso em: 25 nov. 2023.

BRASIL. *Lei nº 9.433, de 08 de janeiro de 1997*. Institui a Política Nacional de Recursos Hídricos, cria o Sistema Nacional de Gerenciamento de Recursos Hídricos, regulamenta o inciso XIX do art. 21 da Constituição Federal, e altera o art. 1º da Lei nº 8.001, de 13 de

março de 1990, que modificou a Lei nº 7.990, de 28 de dezembro de 1989. Brasília, 09 jan. 1997.

BU, H. *et al.* Relationships between land use patterns and water quality in the Taizi River basin, China. *Ecological Indicators*, v. 41, p. 187–197, 2014.

BURAS, M. P.; DONADO, F. S. Identifying and estimating the location of sources of industrial pollution in the sewage network. *Sensors*, v. 21, 3426, 2021.
<https://doi.org/10.3390/s21103426>

CAI, M. *et al.* Assessment of soil and water conservation practices in the loess hilly region using a coupled rainfall-runoff-erosion model. *Sustainability*, v. 12, n. 3, p. 934, 2020.

CALAZANS, G. M. *et al.* The use of multivariate statistical methods for optimization of the surface water quality network monitoring in the Paraopeba river basin, Brazil. *Environmental Monitoring and Assessment*, v. 190, p. 491, 2018.

CAO, J. *et al.* Spatiotemporal dynamics of phytoplankton biomass and community succession for driving factors in a meso-eutrophic lake. *Journal of Environmental Management*, v. 345, p. 118693, 2023.

CAPP, E.; NIENOV, O. H. *Bioestatística quantitativa aplicada*. Porto Alegre: UFRGS, 2020.
CETESB – COMPANHIA AMBIENTAL DO ESTADO DE SÃO PAULO. Apêndice D: Índices de Qualidade das Águas. In: CETESB. *Qualidade das Águas Interiores no Estado de São Paulo 2018*. São Paulo: CETESB, 2019. p. 1–32.

CETESB – COMPANHIA AMBIENTAL DO ESTADO DE SÃO PAULO. Apêndice C: Significado ambiental e sanitário das variáveis de qualidade das águas e dos sedimentos e metodologias analíticas e de amostragem. In: CETESB. *Qualidade das Águas Interiores no Estado de São Paulo 2022*. São Paulo: CETESB, 2023. p. 1–68.

CHAPMAN, D. V.; SULLIVAN, T. The role of water quality monitoring in the sustainable use of ambient waters. *One Earth*, v. 5, n. 2, p. 132–137, 2022.
<https://doi.org/10.1016/j.oneear.2022.01.008>

CHEN, Z. *et al.* Spatiotemporal analysis of land use pattern and stream water quality in southern Alberta, Canada. *Journal of Contaminant Hydrology*, v. 242, p. 103852, 2021.

CNRH – CONSELHO NACIONAL DE RECURSOS HÍDRICOS. *Resolução nº 91, de 5 de novembro de 2008*. Dispõe sobre procedimentos gerais para o enquadramento dos corpos de água superficiais e subterrâneos. *Diário Oficial da União*, Brasília, 06 fev. 2009.

COMPESA – COMPANHIA PERNAMBUCANA DE SANEAMENTO. *Plano de Comunicação do Programa de Saneamento Ambiental da Bacia Hidrográfica do Rio Ipojuca. Volume I: Mapeamento, Análise das áreas e dos públicos*. Recife, 2016.

COMPESA – COMPANHIA PERNAMBUCANA DE SANEAMENTO. *Programa de Saneamento Ambiental da Bacia do rio Ipojuca. Relatório de progresso do PSA Ipojuca, 1º semestre de 2023*. Recife: COMPESA, 2023.

CONAMA – CONSELHO NACIONAL DO MEIO AMBIENTE. *Resolução nº 357, de 17 de março de 2005*. Dispõe sobre a classificação dos corpos de água e diretrizes ambientais para o seu enquadramento, bem como estabelece as condições e padrões de lançamento de efluentes, e dá outras providências. *Diário Oficial da União*, Brasília, n. 53, p. 58-63, 18 mar. 2005.

CONAMA – CONSELHO NACIONAL DO MEIO AMBIENTE. *Resolução nº 396, de 3 de abril de 2008*. Dispõe sobre a classificação e diretrizes ambientais para o enquadramento das águas subterrâneas e dá outras providências. *Diário Oficial da União*, Brasília, n. 66, p. 64-68, 7 abr. 2008.

CONDEPE/FIDEM – AGÊNCIA ESTADUAL DE PLANEJAMENTO E PESQUISAS DE PERNAMBUCO. *Série Bacias Hidrográficas de Pernambuco nº 1 – Bacia Hidrográfica do Rio Ipojuca*. Recife: CONDEPE/FIDEM, 2005. 64 p.

CONWAY, D. *The Data Science Venn Diagram*, 2013. Disponível em: <http://drewconway.com/zia/2013/3/26/the-data-science-venn-diagram>. Acesso em: 12 dez. 2023.

CPRH – AGÊNCIA ESTADUAL DE MEIO AMBIENTE DE PERNAMBUCO. *Bacias Hidrográficas*, 2023. Disponível em: <https://www2.cprh.pe.gov.br/monitoramento-ambiental/qualidade-da-agua/bacias-hidrograficas/>. Acesso em: 10 dez. 2023.

CPRH – AGÊNCIA ESTADUAL DE MEIO AMBIENTE. *Relatório de monitoramento de bacias hidrográficas do Estado de Pernambuco – 2019*. Recife: CPRH, 2020. 200 p.

CYSNE, J. R. *Modelo conceitual para gerenciamento de recursos hídricos em ambiente SIG*. 2004. 94 f. Dissertação (Mestrado em Ciência da Computação) – Universidade Federal do Ceará, Fortaleza, 2004.

DING, J. et al. Influences of the land use pattern on water quality in low-order streams of the Dongjiang River basin, China: A multi-scale analysis. *Science of the Total Environment*, v. 551–552, p. 205–216, 2016.

ELARAB, M. et al. Estimating chlorophyll with thermal and broadband multispectral high resolution imagery from an unmanned aerial system using relevance vector machines for precision agriculture. *International Journal of Applied Earth Observation and Geoinformation*, v. 43, p. 32–42, 2015. <https://doi.org/10.1016/j.jag.2015.03.017>

ESTEVES, F. A. *Fundamentos de limnologia*. 3. ed. Rio de Janeiro: Interciência, 2011. 826 p.

FREIRE, L. L.; COSTA, A. C.; LIMA NETO, I. E. Spatio-temporal patterns of river water quality in the semiarid northeastern Brazil. *Water, Air, & Soil Pollution*, v. 232, n. 11, p. 452, 2021.

GAUCI, A. et al. A Machine Learning approach for automatic land cover mapping from DSLR images over the Maltese Islands. *Environmental Modelling & Software*, v. 99, p. 1–10, 2018. <https://doi.org/10.1016/j.envsoft.2017.09.014>

GIBERT, K. et al. Environmental Data Science. *Environmental Modelling & Software*, v. 106, p. 4–12, 2018. <https://doi.org/10.1016/j.envsoft.2018.04.005>

GIBERT, K.; SÀNCHEZ-MARRÈ, M.; IZQUIERDO, J. A Survey on pre-processing techniques in the context of environmental data mining. *Artificial Intelligence in Communications*, v. 29, n. 6, p. 627–663, 2016. <https://doi.org/10.3233/AIC-160710>

GIRI, S. et al. Evaluating the impact of land uses on stream integrity using machine learning algorithms. *Science of the Total Environment*, v. 696, p. 133858, 2019.

GIRI, S. Water quality prospective in Twenty First Century: Status of water quality in major river basins, contemporary strategies and impediments: A review. *Environmental Pollution*, v. 271, p. 116332, 2021.

GIRI, S.; QIU, Z. Understanding the relationship of land uses and water quality in Twenty First Century: A review. *Journal of Environmental Management*, v. 173, p. 41–48, 2016.

GONZÁLEZ, P. S.; STEHR, A.; BARRA, R. O. Assessment of water quality trends through the application of an aggregated water quality index with historical monitored data in a Mediterranean Andean basin. *Ecological Indicators*, v. 166, p. 112373, 2024.

GUNKEL, G. et al. Sugar cane industry as a source of water pollution – Case study on the situation in Ipojuca River, Pernambuco, Brazil. *Water, Air, and Soil Pollution*, v. 180, n. 1–4, p. 261–269, 2007. <https://doi.org/10.1007/s11270-006-9268-x>

HAIR, J. F. et al. *Análise multivariada de dados*. 6. ed. Porto Alegre: Bookman, 2009. 687 p.

HAN, H. et al. Incorporating a new landscape intensity indicator into landscape metrics to better understand controls of water quality and optimal width of riparian buffer zone. *Journal of Hydrology*, v. 625, p. 130088, 2023.

HANSFORD, M. R.; PLINK-BJÖRKLUND, P.; JONES, E. R. Global quantitative analyses of river discharge variability and hydrograph shape with respect to climate types. *Earth-Science Reviews*, v. 200, p. 102977, 2020.

HILL, D. J. et al. A virtual sensor system for user-generated, real-time environmental data products. *Environmental Modelling & Software*, v. 26, n. 12, p. 1710–1724, 2011. <https://doi.org/10.1016/j.envsoft.2011.09.001>

HORSBURGH, J. S. et al. Measuring water use, conservation, and differences by gender using an inexpensive, high frequency metering system. *Environmental Modelling & Software*, v. 96, p. 83–94, 2017. <https://doi.org/10.1016/j.envsoft.2017.06.035>

IBGE – INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA. *Brasil/Pernambuco*, 2022. Disponível em: <https://cidades.ibge.gov.br/brasil/pe/panorama>. Acesso em: 10 dez. 2023.

IBM – INTERNATIONAL BUSINESS MACHINES CORPORATION. *What is business intelligence?*, 2023. Disponível em: <https://www.ibm.com/topics/business-intelligence>. Acesso em: 15 dez. 2023.

JARDIM, F. A.; VON SPERLING, E.; JARDIM, B.; ALMEIDA, K. F. Fatores determinantes das florações de cianobactérias na água do Rio Doce, Minas Gerais, Brasil. *Engenharia Sanitária e Ambiental*, v. 19, p. 207-218, 2014.

JOLLIFFE, I. T.; CADIMA, J. Principal component analysis: a review and recent developments. *Philosophical Transactions of the Royal Society A*, v. 374, n. 2065, p. 20150202, 2016.

JONGMAN, E.; JONGMAN, S. R. R. *Data analysis in community and landscape ecology*. Cambridge: Cambridge University Press, 1995.

KENDALL, M. *Rank correlation methods*. 4. ed. London: Charles Griffin, 1975.

KOROGLU, E. et al. Scale-up and commercialization issues of the MFCs: challenges and implications. In: MOHAN, S. V.; VARJANI, S.; PANDEY, A. (ed.). *Microbial electrochemical technology*. Amsterdam: Elsevier, 2019. p. 565–583.

KRANSTAUBER, B. et al. The Movebank data model for animal tracking. *Environmental Modelling & Software*, v. 26, n. 6, p. 834-835, 2011.
<https://doi.org/10.1016/j.envsoft.2010.12.005>

KUMAR, E. S. et al. Comprehensive analysis of cloud based databases. *IOP Conference Series: Materials Science and Engineering*, v. 1131, p. 012021, 2021.
<https://doi.org/10.1088/1757-899X/1131/1/012021>

LE, T. V.; DO, D. D.; NGUYEN, B. T. Spatiotemporal assessment and pollution-source identification and quantification of the surface water system in a coastal region of Vietnam. *Hydrological Sciences Journal*, v. 68, n. 6, p. 782-793, 2023.

LEE, J. et al. Relationships between water quality parameters in rivers and lakes: BOD₅, COD, NBOPS, and TOC. *Environmental Monitoring and Assessment*, v. 188, p. 252, 2016.

LE MOAL, M. et al. Eutrophication: A new wine in an old bottle? *Science of the Total Environment*, v. 651, p. 1-11, 2019.

LI, B. et al. Combining multivariate statistical techniques and random forests model to assess and diagnose the trophic status of Poyang Lake in China. *Ecological Indicators*, v. 83, p. 74–83, 2017.

LI, X.; LI, Y.; LI, G. A scientometric review of the research on the impacts of climate change on water quality during 1998–2018. *Environmental Science and Pollution Research*, v. 27, p. 14322–14341, 2020.: <https://doi.org/10.1007/s11356-020-08176-7>

LIU, M.; LU, J. Support vector machine—An alternative to artificial neuron network for water quality forecasting in an agricultural nonpoint source polluted river? *Environmental Science and Pollution Research*, v. 21, p. 11036–11053, 2014.
<https://doi.org/10.1007/s11356-014-3046-x>

MA, J. et al. Soft detection of 5-day BOD with sparse matrix in city harbor water using deep learning techniques. *Water Research*, v. 170, p. 115350, 2020.
<https://doi.org/10.1016/j.watres.2019.115350>

MANN, H. B. Nonparametric tests against trend. *Econometrica*, v. 13, p. 245–259, 1945.

MANN, H. B.; WHITNEY, D. R. On a test of whether one of two random variables is stochastically larger than the other. *Annals of Mathematical Statistics*, v. 18, n. 1, p. 50–60, 1947. 10.1214/aoms/1177730491

MAPBIOMAS. Visão geral da metodologia. 2024. Disponível em: <https://mapbiomas.org/visao-geral-da-metodologia>. Acesso em: 25 abr. 2024.

MELLO, K. et al. Effects of land use and land cover on water quality of low-order streams in Southeastern Brazil: Watershed versus riparian zone. *Catena*, v. 167, p. 130-138, 2018.

MELLO, K. et al. Multiscale land use impacts on water quality: Assessment, planning, and future perspectives in Brazil. *Journal of Environmental Management*, v. 270, p. 110879, 2020.

MENDONÇA, V. S. Aplicabilidade de testes de toxicidade aguda com *Daphnia magna* e *Vibrio fischeri*, no monitoramento da qualidade das águas de bacias hidrográficas: o caso do rio Ipojuca em Pernambuco. 2005. Dissertação (Mestrado em Gestão e Políticas Ambientais) – Universidade Federal de Pernambuco, Recife, 2005.

MICROSOFT. Get started building with Power BI. 2024. Disponível em: <https://learn.microsoft.com/en-us/training/modules/get-started-with-power-bi/1-introduction>. Acesso em: 17 fev. 2024.

MILAN, B. F. Clean water and sanitation for all: interactions with other sustainable development goals. *Sustainable Water Resources Management*, v. 3, p. 479–489, 2017.

MOHAMMED, H.; LONGVA, A.; SEIDU, R. Predictive analysis of microbial water quality using machine-learning algorithms. *Environmental Research, Engineering and Management*, v. 74, n. 1, p. 7–20, 2018.

MOHAN, S.; KUMAR, B.; NEJADHASHEMI, A. P. Integration of machine learning and remote sensing for water quality monitoring and prediction: a review. *Sustainability*, v. 17, n. 3, p. 998, 2025.

MOHANAVELU, A.; NAGANNA, S. R.; AL-ANSARI, N. Irrigation induced salinity and sodicity hazards on soil and groundwater: an overview of its causes, impacts and mitigation strategies. *Agriculture*, v. 11, p. 983, 2021. <https://doi.org/10.3390/agriculture11100983>

MOLINA-NAVARRO, E. et al. Hydrological and water quality impact assessment of a Mediterranean limno-reservoir under climate change and land use management scenarios. *Journal of Hydrology*, v. 509, p. 354–366, 2014.

MUNTEAN, M. Business intelligence issues for sustainability projects. *Sustainability*, v. 10, p. 335, 2018. <https://doi.org/10.3390/su10020335>

NACAR, S. et al. Estimation of daily dissolved oxygen concentration for river water quality using conventional regression analysis, multivariate adaptive regression splines, and TreeNet techniques. *Environmental Monitoring and Assessment*, v. 192, p. 1-21, 2020.

NATIVI, S. et al. Big data challenges in building the global Earth observation system of systems. *Environmental Modelling & Software*, v. 68, p. 1–26, 2015.
<https://doi.org/10.1016/j.envsoft.2015.01.017>

NGAMILE, S.; MADONSELA, S.; KGANYAGO, M. Trends in remote sensing of water quality parameters in inland water bodies: a systematic review. *Frontiers in Environmental Science*, v. 13, p. 1549301, 2025.

NORIEGA, C. et al. Long-term water quality conditions and trends in 12 tropical coastal rivers in Northeast Brazil. *Environmental Monitoring and Assessment*, v. 193, n. 5, p. 308, 2021.

OLAVSRUD, T.; FRUHLINGER, J. What is business intelligence? Transforming data into business insights. 2023. Disponível em: <https://www.cio.com/article/272364/business-intelligence-definition-and-solutions.html>. Acesso em: 15 dez. 2023.

ONU – ORGANIZAÇÃO DAS NAÇÕES UNIDAS. Agenda 2030. 2015. Disponível em: <https://nacoesunidas.org/pos2015/agenda2030/>. Acesso em: 25 nov. 2023.

OSIANLIS, A.; ARNOTT, D. R. The application of business intelligence and data warehousing in an Australian water business. In: BREBBIA, C. A.; PASCOLO, P. (org.). *Management Information Systems 2002: Incorporating GIS and Remote Sensing*. Southampton: WITPress, 2002. p. 369–378.

PAK, H. Y. et al. Effects of land use configuration, seasonality and point source on water quality in a tropical watershed: A case study of the Johor River Basin. *Science of the Total Environment*, v. 780, p. 1–18, 2021.

PALMA, P. et al. Assessment of anthropogenic sources of water pollution using multivariate statistical techniques: a case study of the Alqueva’s reservoir, Portugal. *Environmental Monitoring and Assessment*, v. 165, p. 539–552, 2010.

PARSAIE, A.; NASROLAHI, A. H.; HAGHIABI, A. H. Water quality prediction using machine learning methods. *Water Quality Research Journal*, v. 53, p. 3–13, 2018.
<https://doi.org/10.2166/wqrj.2018.025>

PEREIRA, M. C. S. et al. Melhoria da qualidade da água de rios urbanos: novos paradigmas a explorar – Bacia hidrográfica do rio Pinheiros em São Paulo. *Engenharia Sanitária e Ambiental*, v. 26, n. 3, p. 577–590, 2021.

PÉREZ-BELTRÁN, C. H. et al. Artificial intelligence and water quality: From drinking water to wastewater. *TrAC Trends in Analytical Chemistry*, v. 172, p. 117597, 2024.

PERNAMBUCO. Lei nº 11.426, de 17 de janeiro de 1997. (Revogada pelo art.71 da Lei nº 12.984, de 30 de dezembro de 2005). Dispõe sobre a Política Estadual de Recursos Hídricos, e o Plano Estadual de Recursos Hídricos, institui o Sistema Integrado de Gerenciamento de Recursos Hídricos e dá outras providências. Recife, 17 jan. 1997.

PERNAMBUCO. Lei nº 12.984, de 30 de dezembro de 2005. Dispõe sobre a Política Estadual de Recursos Hídricos e o Sistema Integrado de Gerenciamento de Recursos Hídricos. Recife, 2005.

PERNAMBUCO. Lei nº 14.028, de 26 de março de 2010. Cria a Agência Pernambucana de Águas e Clima – APAC, e dá outras providências. Recife, 26 mar. 2010.

PERNAMBUCO. *Plano Hidroambiental da Bacia Hidrográfica do Rio Capibaribe*. Recife, PE: Secretaria de Recursos Hídricos, 2010a.

PERNAMBUCO. *Plano Hidroambiental da Bacia Hidrográfica do Rio Ipojuca*. Recife, PE: Secretaria de Recursos Hídricos, 2010b.

PERNAMBUCO. *Plano Estadual de Recursos Hídricos de Pernambuco: resumo executivo*. Recife: SEINFRA, 2022.

PERNAMBUCO. *Plano Estadual de Recursos Hídricos de Pernambuco: síntese*. Recife: SECTMA, 1998. 212 p.

PETTITT, A. N. A Non-Parametric Approach to the Change-Point Problem. *Applied Statistics*, v. 28, p. 126, 1976. <https://doi.org/10.2307/2346729>

PONCE ROMERO, J.; HALLETT, S.; JUDE, S. Leveraging Big Data Tools and Technologies: Addressing the Challenges of the Water Quality Sector. *Sustainability*, v. 9, n. 12, p. 2160, 2017. <http://dx.doi.org/10.3390/su9122160>

PRATT, B.; CHANG, H. Effects of land cover, topography, and built structure on seasonal water quality at multiple spatial scales. *Journal of Hazardous Materials*, v. 209–210, p. 48–58, 2012.

REIS, S. et al. Integrating modelling and smart sensors for environmental and human health. *Environmental Modelling & Software*, v. 74, p. 238–246, 2015. <https://doi.org/10.1016/j.envsoft.2015.06.003>

ROELOFSEN, H. D. et al. Mapping a priori defined plant associations using remotely sensed vegetation characteristics. *Remote Sensing of Environment*, v. 140, p. 639–651, 2014.

ROY, D. C. et al. Biodegradation of crystal violet dye by bacterial isolated from textile industry effluents. *PeerJ*, v. 6, p. e5015, 2018. <https://doi.org/10.7717/peerj.5015>

SAGAN, V. et al. Monitoring inland water quality using remote sensing: potential and limitations of spectral indices, bio-optical simulations, machine learning, and cloud computing. *Earth Science Reviews*, v. 205, p. 103187, 2020. <https://doi.org/10.1016/j.earscirev.2020.103187>

SAJEDI HOSSEINI, F. et al. A novel machine learning-based approach for the risk assessment of nitrate groundwater contamination. *Science of the Total Environment*, v. 644, p. 954–962, 2018. <https://doi.org/10.1016/j.scitotenv.2018.07.054>

SHAPIRO, S. S.; WILK, M. B. An analysis of variance test for normality (complete samples). *Biometrika*, v. 52, n. 3/4, p. 591–611, 1965.

SHEN, Z. et al. Impact of landscape pattern at multiple spatial scales on water quality: A case study in a typical urbanised watershed in China. *Ecological Indicators*, v. 48, p. 417–427, 2015.

SHI, B. et al. Understanding spatiotemporal variability of in-stream water quality in urban environments – a case study of Melbourne, Australia. *Journal of Environmental Management*, v. 246, p. 203–213, 2019.

SILVA, E. M. de S. et al. Evaluation of water quality in streams under strong influence of sugarcane cultivation using the *Allium cepa* test. *Ecotoxicology and Environmental Contamination*, v. 18, n. 1, p. 97–103, 2023. <https://doi.org/10.5132/eec.2023.01.11>

SILVA, J. et al. Multivariate statistical analysis of surface water quality in the Capibaribe River (Pernambuco state, Northeast Brazil): contributions to water management. *Marine Environmental Research*, v. 204, p. 106876, 2025.

SILVA, M. C. et al. Tendência pluviométrica na bacia hidrográfica do rio Ipojuca-PE, Brasil. *Observatório de la Economía Latinoamericana*, v. 21, n. 10, p. 17065–17086, 2023.

SILVA, R. R. et al. Avaliação da qualidade microbiológica e físico-química da água de um trecho do Rio Capibaribe, no município de São Lourenço da Mata, Pernambuco, Brasil. *Brazilian Journal of Development*, v. 7, n. 6, p. 62847–62866, 2021.

SINGH, P.; BANSAL, U. An investigation for river water quality via linear regression analysis of variance. *Water Resources*, v. 52, n. 1, p. 205–216, 2025.

SINHA, K.; SRIVASTAVA, D. K.; BHATNAGAR, R. Water quality management through data driven intelligence system in Barmer Region, Rajasthan. *Procedia Computer Science*, v. 132, p. 314–322, 2018. <http://dx.doi.org/10.1016/j.procs.2018.05.183>

SNIS – SISTEMA NACIONAL DE INFORMAÇÕES SOBRE SANEAMENTO. Diagnóstico Temático: Serviços de Água e Esgoto – ano de referência 2022. Brasília: SNS/MDR, 2024.

SOARES, A. E. P. Potencial de reúso de efluentes sanitários e sua interface com a gestão de recursos hídricos na bacia hidrográfica do Rio Ipojuca – Pernambuco. 2023. 115 p. Tese (Doutorado) – Universidade Federal de Pernambuco, Centro de Tecnologia e Geociências, Programa de Pós-Graduação em Engenharia Civil, Recife, 2023.

SOBRAL, M. C.; CARVALHO, R. M. C.; FIGUEIREDO, R. C. Environmental risk management of multipurpose use of reservoirs in semiarid area of São Francisco River, Brazil. In: GUNKEL, G.; SOBRAL, M. C. (org.). *Reservoir and River Basin Management: Exchange of Experiences from Brazil, Portugal and Germany*. Berlin: Universitätsverlag TUBerlin, 2007. p. 14–26. ISBN 978-3-7983-2056-7.

SONG, P. et al. A biophysiological perspective on enhanced nitrate removal from decentralized domestic sewage using gravitational-flow multi-soil-layering systems. *Chemosphere*, v. 240, p. 124868, 2020.

SPEARMAN, C. The proof and measurement of association between two things. *American Journal of Psychology*, v. 15, p. 72–101, 1904.

SUN, X. et al. Assessment of seasonal and spatial water quality variation in a cascading lake system in Chennai, India. *Science of the Total Environment*, v. 858, p. 159924, 2023.

SWERSKY, K. et al. Machine learning for water monitoring, hydrology, and sustainability. 2022. Disponível em: <http://www.cs.toronto.edu/~kswersky/wp-content/uploads/WorkshopPresentation.pdf>. Acesso em: 12 dez. 2023.

TSAI, C. W. et al. Big data analytics: a survey. *Journal of Big Data*, v. 2, p. 21, 2015. <https://doi.org/10.1186/s40537-015-0030-3>

TUNDISI, J. G.; TUNDISI, T. M. *Limnologia*. São Paulo: Oficina de Textos, 2008.

UNEP – UNITED NATIONS ENVIRONMENT PROGRAMME. Progress on Integrated Water Resources Management. Tracking SDG 6 series: global indicator 6.5.1 updates and acceleration needs. 2021. Disponível em: <https://wedocs.unep.org/bitstream/handle/20.500.11822/36690/PIWRS6.5.1.pdf>. Acesso em: 25 nov. 2023.

UN – UNITED NATIONS. Transforming our world: the 2030 Agenda for Sustainable Development. 2015. Disponível em: <https://sdgs.un.org/2030agenda>. Acesso em: 25 nov. 2023.

URONEN, S. Towards enhanced decision making in Finnish water sector with the help of business intelligence and business analytics. 2018. 67 p. Master's Thesis (Industrial Engineering and Management) – Lappeenranta University of Technology, Mikkeli, 2018.

VAN DER AALST, W. Data science in action. In: *Process Mining*. Berlin, Heidelberg: Springer, 2016. https://doi.org/10.1007/978-3-662-49851-4_1

VEGA, M. et al. Assessment of seasonal and polluting effects on the quality of river water by exploratory data analysis. *Water Research*, v. 32, p. 3581–3592, 1998.

VITOLO, C. et al. Web technologies for environmental big data. *Environmental Modelling & Software*, v. 63, p. 185–198, 2015.

VON SPERLING, M. *Introdução à qualidade das águas e ao tratamento de esgotos: volume 1*. Belo Horizonte: UFMG, 2005.

WANG, Y.; LIU, C.; LEE, J. Differentiating the spatiotemporal distribution of natural and anthropogenic processes on river water–quality variation using a self-organizing map with factor analysis. *Archives of Environmental Contamination and Toxicology*, v. 69, n. 2, p. 254–263, 2015. <https://doi.org/10.1007/s00244-015-0167-2>

WEBB, B. W.; CLACK, P. D.; WALLING, D. E. Water-air temperature relationships in a Devon river system and the role of flow. *Hydrological Processes*, v. 17, n. 15, p. 3069–3084, 2003.

WEBB, B. W. Trends in stream and river temperature. *Hydrological Processes*, v. 10, n. 2, p. 205–226, 1996.

WEISSTEINER, C.; BOURAOUI, F.; ALOE, A. Reduction of nitrogen and phosphorus loads to European rivers by riparian buffer zones. *Knowledge and Management of Aquatic Ecosystems*, n. 408, p. 08, 2013.

WEI, W. et al. Exploring the effect of basin land degradation on lake and reservoir water quality in China. *Journal of Cleaner Production*, v. 268, p. 122249, 2020.

WIEMANN, S. et al. Design and prototype of an interoperable online air quality information system. *Environmental Modelling & Software*, v. 79, p. 354-366, 2016. <https://doi.org/10.1016/j.envsoft.2015.10.028>

WILCOXON, F. Individual comparisons by ranking methods. *Biometrics Bulletin*, v. 1, n. 6, p. 80–83, 1945. <https://doi.org/10.2307/3001968>

WONG, B. P.; KERKEZ, B. Real-time environmental sensor data: an application to water quality using web services. *Environmental Modelling & Software*, v. 84, p. 505-517, 2016. <https://doi.org/10.1016/j.envsoft.2016.07.020>

WU, J. et al. Spatiotemporal characteristics of the water quality and its multiscale relationship with land use in the Yangtze River Basin. *Remote Sensing*, v. 13, n. 16, p. 3309, 2021.

WU, Y. et al. Attention neural network for water image classification under IoT environment. *Applied Sciences*, v. 10, p. 030909, 2020. <https://doi.org/10.3390/app10030909>

WWAP - UNESCO WORLD WATER ASSESSMENT PROGRAMME. *The United Nations World Water Development Report 2019: Leaving No One Behind*. Paris: UNESCO, 2019. Disponível em: <https://www.unesco.org/en/wwap/wwdr/2019>. Acesso em: 25 nov. 2023.

XAVIER, M. V. S.; QUADROS, H. C.; SILVA, M. S. S. da. Water potability parameters for human consumption: an integrative review. *Research, Society and Development*, v. 11, n. 1, p. e42511125118, 2022. <http://dx.doi.org/10.33448/rsd-v11i1.25118>

XIONG, F. et al. Land use, hydrology, and climate influence water quality of China's largest river. *Journal of Environmental Management*, v. 318, p. 115581, 2022.

XU, H. et al. Responses of water quality to land use in riparian buffers: a case study of Huangpu River, China. *GeoJournal*, v. 86, n. 4, p. 1657–1669, 2021.

ZHANG, J. et al. Influences of land use metrics at multi-spatial scales on seasonal water quality: A case study of river systems in the Three Gorges Reservoir Area, China. *Journal of Cleaner Production*, v. 206, p. 76–85, 2019.

ZHANG, Y. et al. Integrating water quality and operation into prediction of water production in drinking water treatment plants by genetic algorithm enhanced artificial neural network. *Water Research*, v. 164, p. 114888, 2019. <https://doi.org/10.1016/j.watres.2019.114888>

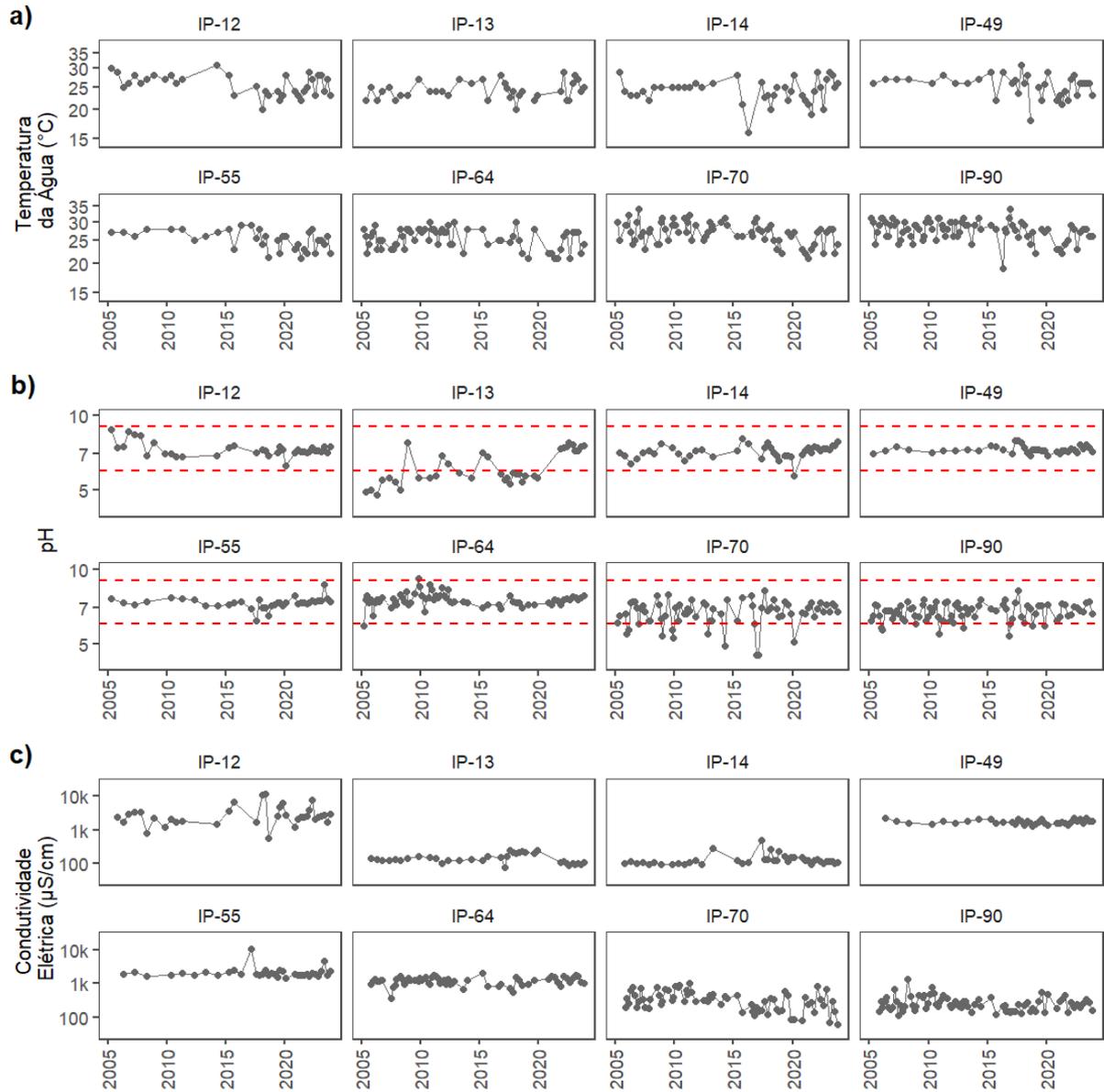
ZHAO, C. S. et al. Quantitative assessment of the effects of human activities on phytoplankton communities in lakes and reservoirs. *Science of the Total Environment*, v. 665, p. 213-225, 2019.

ZHOU, P. et al. New insight into the correlations between land use and water quality in a coastal watershed of China: Does point source pollution weaken it? *Science of the Total Environment*, v. 543, p. 591–600, 2016.

ZORZAL-ALMEIDA, S. et al. Effects of land use and spatial processes in water and surface sediment of tropical reservoirs at local and regional scales. *Science of the Total Environment*, v. 644, p. 237–246, 2018.

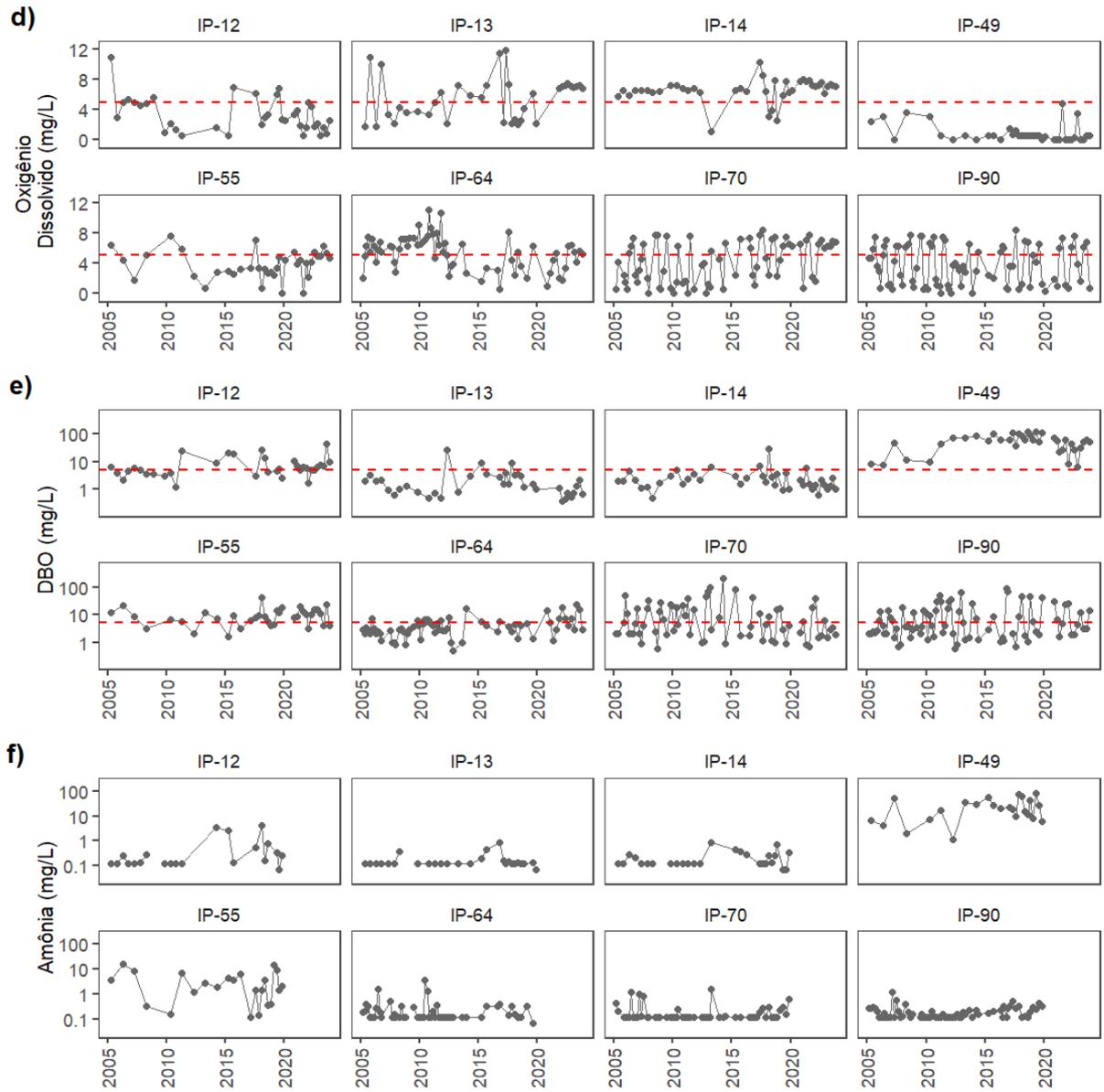
APÊNDICE A - GRÁFICOS DA SÉRIE TEMPORAL DOS PARÂMETROS DE QUALIDADE DA ÁGUA

Figura 47 - Variação dos parâmetros de qualidade da água nas estações de amostragem da BHRI ao longo do tempo



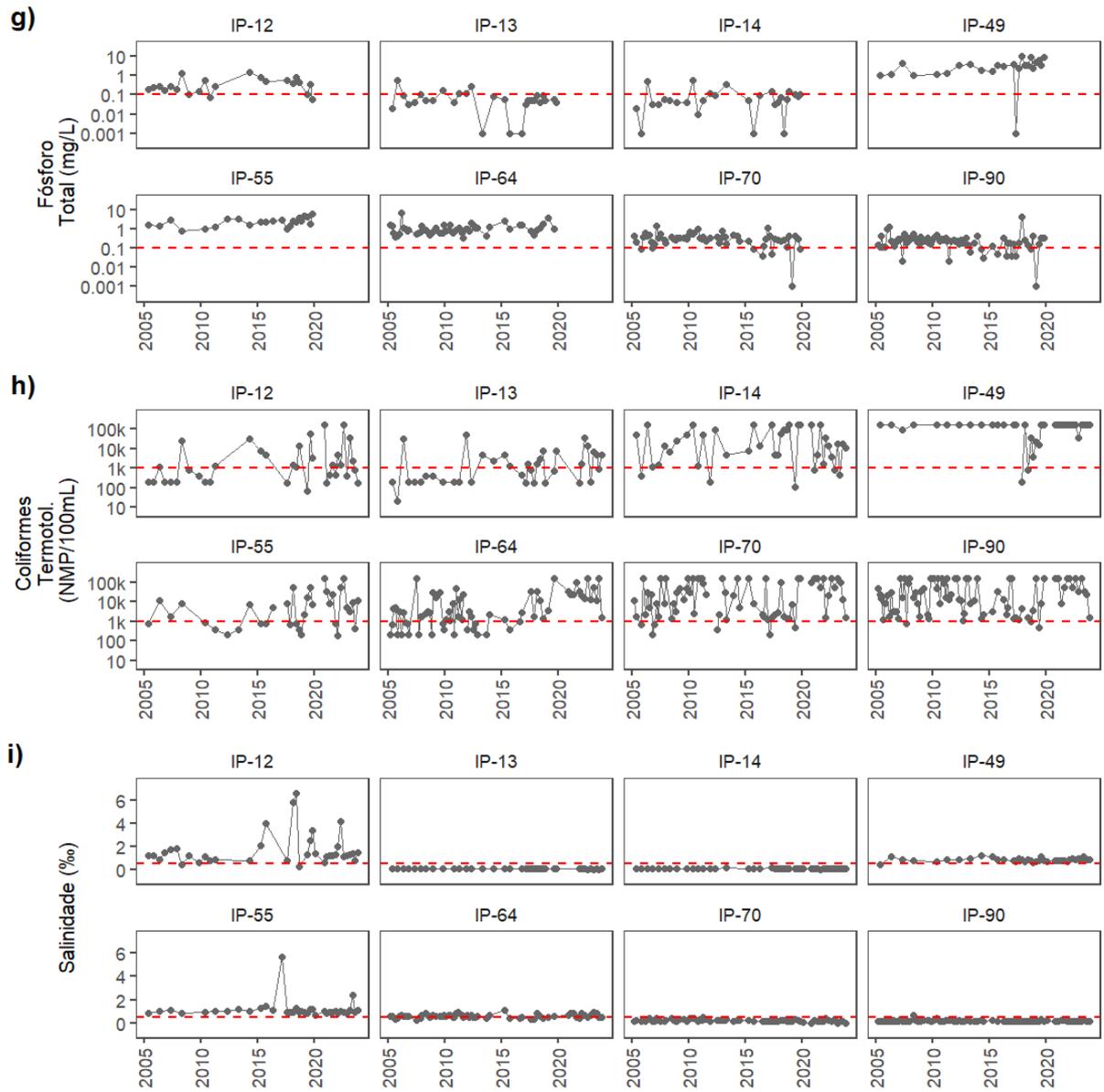
Fonte: Elaborado pela autora (2025)

Figura 48 - Continuação



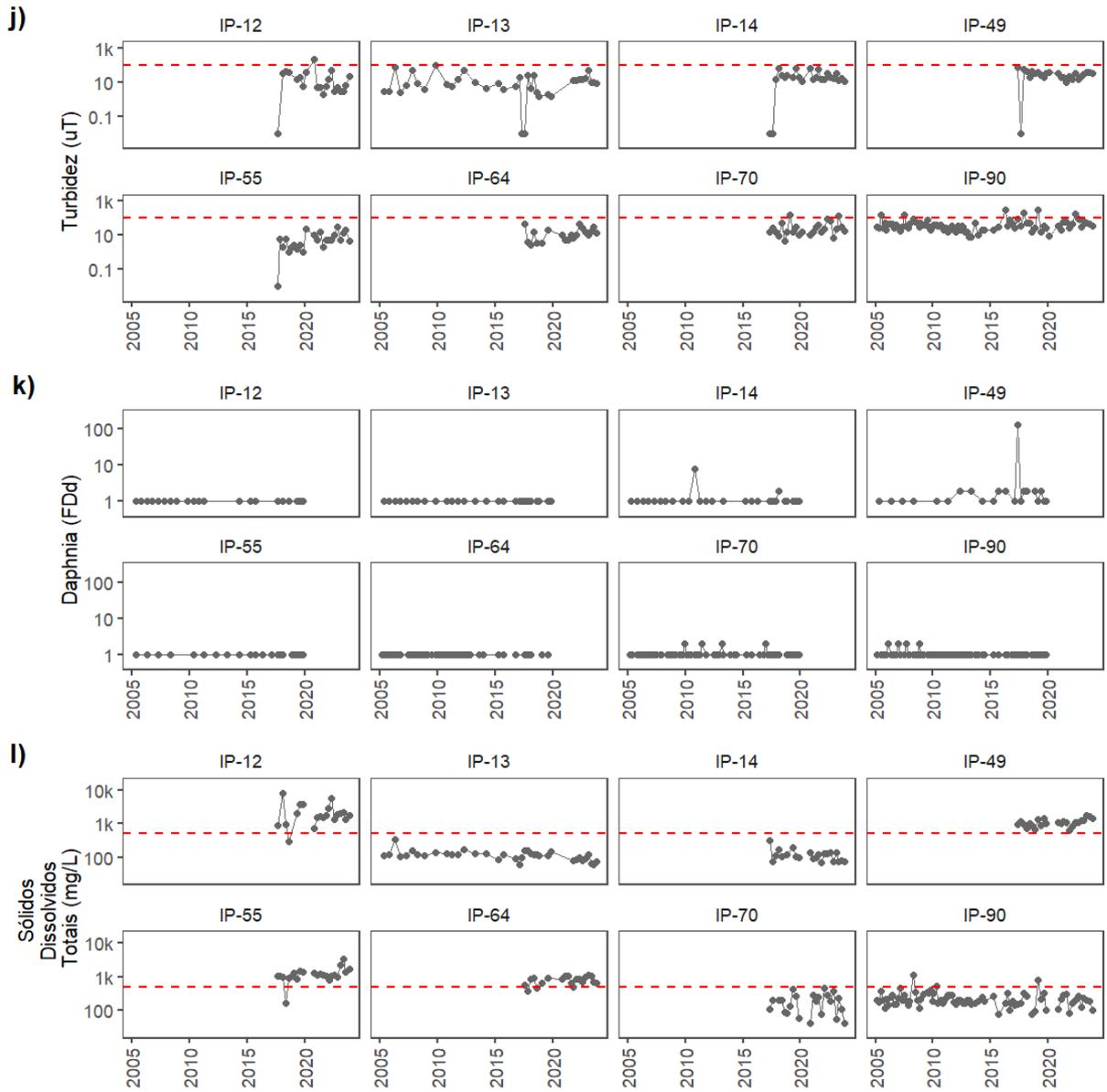
Fonte: Elaborado pela autora (2025)

Figura 49 - Continuação



Fonte: Elaborado pela autora (2025)

Figura 50 - Continuação



Fonte: Elaborado pela autora (2025)