



Universidade Federal de Ouro Preto
Escola de Minas
Departamento de Engenharia Ambiental
Graduação em Engenharia Ambiental



Gabriela Souza Peret Paulino

PLATAFORMA DE MONITORAMENTO AMBIENTAL DO RIO DAS VELHAS: TRANSPARÊNCIA, PARTICIPAÇÃO PÚBLICA E PRESERVAÇÃO EM CÓDIGO ABERTO

Ouro Preto

2024

Gabriela Souza Peret Paulino

**Plataforma de monitoramento ambiental do rio das Velhas: transparência,
participação pública e preservação em código aberto**

Monografia apresentada ao Curso de Engenharia Ambiental da Universidade Federal de Ouro Preto como requisito parcial para obtenção do título de Engenheira Ambiental.

Orientadora: Dra. Lívia Cristina Pinto Dias.

Ouro Preto

2024

SISBIN - SISTEMA DE BIBLIOTECAS E INFORMAÇÃO

P328p Paulino, Gabriela Souza Peret.
Plataforma de monitoramento ambiental do Rio Das Velhas
[manuscrito]: transparência, participação pública e preservação em
código aberto. / Gabriela Souza Peret Paulino. - 2024.
55 f.: il.: color., gráf., tab., mapa.

Orientadora: Profa. Dra. Livia Cristina Pinto Dias.
Monografia (Bacharelado). Universidade Federal de Ouro Preto.
Escola de Minas. Graduação em Engenharia Ambiental .

1. Monitoramento ambiental. 2. Estatísticas Ambientais. 3. Software
livre. I. Dias, Livia Cristina Pinto. II. Universidade Federal de Ouro Preto.
III. Título.

CDU 504:004

Bibliotecário(a) Responsável: Soraya Fernanda Ferreira e Souza - CRB6/2322



FOLHA DE APROVAÇÃO

Gabriela Souza Peret Paulino

Plataforma de monitoramento ambiental do rio Das Velhas: Transparência, participação pública e preservação em código aberto

Monografia apresentada ao Curso de Engenharia Ambiental da Universidade Federal de Ouro Preto como requisito parcial para obtenção do título de Engenheira Ambiental

Aprovada em 30 de setembro de 2024

Membros da banca

Profa. Dra. Lívia Cristina Pinto Dias (Orientadora) - Departamento de Engenharia Ambiental - Universidade Federal de Ouro Preto
Profa. Dra. Aline de Araújo Nunes - Departamento de Engenharia Urbana - Universidade Federal de Ouro Preto
Dra. Grazielle Rocha dos Santos - Programa de Pós-graduação em Engenharia Ambiental - Universidade Federal de Ouro Preto

Lívia Cristina Pinto Dias, orientadora do trabalho, aprovou a versão final e autorizou seu depósito na Biblioteca Digital de Trabalhos de Conclusão de Curso da UFOP em 13/05/2026.



Documento assinado eletronicamente por **Livia Cristina Pinto Dias, PROFESSOR DE MAGISTERIO SUPERIOR**, em 13/05/2026, às 13:19, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



A autenticidade deste documento pode ser conferida no site http://sei.ufop.br/sei/controlador_externo.php?acao=documento_conferir&id_orgao_acesso_externo=0, informando o código verificador **1106933** e o código CRC **7236261C**.

AGRADECIMENTOS

A Deus, por me guardar nos momentos difíceis.

A UFOP e ao PET Ambiental pelo ensino superior de qualidade e gratuito.

Ao FNDE pela bolsa concedida.

A professora Lívia, pela orientação e sensibilidade.

A Grazielle dos Santos pela brilhante tutoria durante a minha Iniciação Científica.

A Anglo American pela oportunidade de estágio na área de dados, em especial, a diretoria de IM e Felipe Valadão, por ser um exemplo de profissional e pessoa.

A Arcadis e as relações felizes de amizade que construí, em especial ao Digital Saneamento e Squad Projetos. Vocês são brilhantes e tenho orgulho de ter pertencido a esses times de peso!

Aos meus pais, Carlos e Gláucia, pela vida e por não pouparem esforços por seus filhos.

A minha avó, por todo amor e ensinamentos de vida.

Aos meus irmãos, Flá e Gustavo, e primas, Lari e Carol, por serem meus primeiros amigos e por me apoiarem, mesmo que de longe.

Aos pequenos Matheus, Mirela e Antônio, que amamos tanto.

Aos meus tios, Miro, Deyse, Najila e Flávia pelos bons momentos e memórias de afeto.

Aos amigos de sempre, Alan, Leticya, Dara, Wesley, Henrique, Nickolas, Maia e Dedeco.

Aos amigos de Ouro Preto, Beatriz, Debora, Laura, Luan e Grazielle.

Ao meu companheiro, Arthur, por me lembrar de nunca desistir.

Aos meus professores, em especial Aníbal, Leandro e Tatiana.

“Pegue para você o que lhe pertence, e o que lhe pertence é tudo aquilo que sua vida exige. Parece uma moral amoral. Mas o que é verdadeiramente imoral é ter desistido de si mesma.”

- Clarice Lispector

RESUMO

A Política Nacional de Recursos Hídricos (PNRH), instituída pela Lei nº 9.433 de 1997, promove um modelo de gestão que busca a participação pública, descentralização e integração como princípios fundamentais para a conservação e uso sustentável dos recursos hídricos. Um dos pilares dessa política é o monitoramento participativo, que incentiva a sociedade civil, o poder público e os usuários a interagirem ativamente na gestão e fiscalização da qualidade e quantidade das águas. Essa participação é crucial para garantir a transparência e efetividade das decisões e ações voltadas à preservação dos mananciais. No entanto, embora o Brasil possua um vasto conjunto de repositórios de dados ambientais, como os sistemas do Instituto Mineiro de Gestão das Águas (IGAM) e da Agência Nacional de Águas (ANA), esses dados frequentemente são subutilizados devido à complexidade na sua extração e interpretação. Para que essas informações possam gerar impactos práticos, é necessário que ferramentas tecnológicas e de ciência de dados sejam aplicadas, facilitando a coleta, armazenamento, análise e visualização desses dados. Ferramentas como APIs, bancos de dados abertos e plataformas de visualização interativa, como o Power BI e GIS, possibilitam a criação de soluções que atribuem valor aos dados disponíveis, tornando-os acessíveis e compreensíveis para a sociedade. Isso promove não apenas o acompanhamento mais próximo da população sobre as condições dos recursos hídricos, mas também fortalece a tomada de decisões informada e colaborativa, conforme propõe a PNRH. Sendo assim, este trabalho apresenta o desenvolvimento de uma plataforma digital voltada ao monitoramento ambiental da Bacia Hidrográfica do Rio das Velhas, em Minas Gerais, Brasil. Utilizando ferramentas de código aberto e ciência de dados, o projeto visa otimizar a análise e a comunicação de dados ambientais, contribuindo para uma gestão mais eficaz e sustentável da bacia hidrográfica.

PALAVRAS-CHAVE: Monitoramento ambiental participativo; Transparência; Código Aberto; Repositórios Ambientais; Dados Ambientais.

ABSTRACT

The National Water Resources Policy (PNRH), established by Law No. 9,433 of 1997, promotes a management model that seeks public participation, decentralization, and integration as fundamental principles for the conservation and sustainable use of water resources. One of the pillars of this policy is participatory monitoring, which encourages civil society, government, and users to actively interact in the management and monitoring of water quality and quantity. This participation is crucial to ensure the transparency and effectiveness of decisions and actions aimed at preserving water sources. However, although Brazil has a vast set of environmental data repositories, such as the systems of the Minas Gerais Water Management Institute (IGAM) and the National Water Agency (ANA), these data are often underutilized due to the complexity of their extraction and interpretation. In order for this information to generate practical impacts, technological and data science tools must be applied, facilitating the collection, storage, analysis, and visualization of these data. Tools such as APIs, open databases, and interactive visualization platforms, such as Power BI and GIS, enable the creation of solutions that add value to available data, making them accessible and understandable to society. This not only promotes closer monitoring of the population's water resource conditions, but also strengthens informed and collaborative decision-making, as proposed by the PNRH. Therefore, this work presents the development of a digital platform focused on environmental monitoring of the Rio das Velhas Hydrographic Basin, in Minas Gerais, Brazil. Using open source tools and data science, the project aims to optimize the analysis and communication of environmental data, contributing to more effective and sustainable management of the river basin.

KEYWORDS: Participatory environmental monitoring; Transparency; Open Source; Environmental Repositories; Environmental Data.

LISTA DE FIGURAS

Figura 1 - Divisões fisiográficas da Bacia Hidrográfica do Rio São Francisco (PALMA, 2017).	15
Figura 2 - Municípios que compõem a Bacia do Rio das Velhas e antiga segmentação em Alto, Médio e Baixo Velhas (COSTA, 2008).	17
Figura 3 - Unidades Territoriais Estratégicas (UTES) na BH do Velhas (AGÊNCIA PEIXE VIVO e CBH VELHAS, 2015).	19
Figura 4 – Cinco estágios do ciclo de vida da ciência de dados (Adaptado de BERKELEY SCHOOL OF INFORMATION, 2023).	29
Figura 5 – Exemplo de Data Model no Power BI Desktop (SQLBI, 2018).	31
Figura 6 – Fragmento dos dados tabulados da CONAMA 357/2005 (Autora, 2024).	36
Figura 7- Código desenvolvido para funcionamento das requisições de forma simultânea para as 58 estações disponíveis na BHRV (Autora, 2024).	36
Figura 8 – Output das requisições realizadas via API da ANA (Autora, 2024).	39
Figura 9 – Modelo de Dados Integrado para o Monitoramento Ambiental Participativo no Power BI (Autora, 2024).	40
Figura 10 - Medida DAX criada para classificar os meses entre período de chuva e estiagem (Autora, 2024).	42
Figura 11 – Interfaces construídas em POWER BI (Autora, 2024).	42
Figura 12 – Interfaces construídas em POWER BI; tela de apresentação da Bacia Hidrográfica (Autora, 2024).	43
Figura 13 – Interfaces construídas em POWER BI, tela de análise de resultados (Autora, 2024).	44
Figura 14 – Interfaces construídas em POWER BI, tela de análise de não conformidades (Autora, 2024).	45
Figura 15 – Interfaces construídas em POWER BI, tela de análise pluviométrica (Autora, 2024).	46

Figura 16 – Localização da estação PV220 na página “Estações” e detalhamento do Uso e Ocupação do Solo na município (Autora, 2024).	48
Figura 17 – Não conformidades para a estação PV220, seguindo os limites estabelecidos pela resolução CONAMA 357/2005 (Autora, 2024).	48
Figura 18 – Série histórica para o parâmetro de Clorofila-a na estação PV220 (Autora, 2024).	49
Figura 19 – Estatísticas de FQ para a estação PV220 (Autora, 2024).	50
Figura 20 – BoxPlot e estatísticas sazonais para a estação PV220 (Autora, 2024). ..	51

LISTA DE TABELAS E QUADROS

Quadro 1 - Unidades territoriais da bacia do rio das Velhas, localização e principais características.	18
Tabela 1 – Estações de monitoramento do IGAM localizadas na Bacia Hidrográfica do Rio das Velhas e sua distribuição por municípios.	220
Tabela 2 - Arquivos obtidos no repositório do IGAM, período de referência dos dados de monitoramento e extensão dos registros.....	34
Tabela 3 - <i>Dataframe</i> extraído a partir da tabela de atributos no software QGIS com o auxílio do plugin MAP BIOMAS para o município de Araçuaí.	35

LISTA DE SIGLAS

API – Application Programming Interface

CBH – Comitê de Bacia Hidrográfica

CBHSF – Comitê da Bacia Hidrográfica do Rio São Francisco

IBAMA - Instituto Brasileiro do Meio Ambiente e dos Recursos Naturais Renováveis

IGAM – Instituto Mineiro de Gestão de Águas

PDA – Plano de Dados Abertos

PNRH – Política Nacional de Recursos Hídricos

SEMAD - Secretaria de Estado de Meio Ambiente e Desenvolvimento Sustentável

SF5 - Subunidade de Gerenciamento Hidrográfico 5

Sisema – Sistema Estadual de Meio Ambiente e Recursos Hídricos

SNIRH – Sistema Nacional de Recursos Hídricos

UPGRH – Unidade de Planejamento e Gestão dos Recursos Hídricos

UTE - Unidades Territoriais Estratégicas

VMP – Valor Máximo Permissível

SUMÁRIO

1. INTRODUÇÃO	12
2. OBJETIVOS	13
2.1 Objetivo geral	13
2.2 Objetivos específicos	13
3. REVISÃO BIBLIOGRÁFICA.....	14
3.1 A bacia hidrográfica como unidade de planejamento	14
3.2 A bacia hidrográfica do rio São Francisco	15
3.3 A bacia hidrográfica do rio das Velhas.....	16
3.4 Monitoramento ambiental na bacia hidrográfica do rio das Velhas	19
3.5 Participação pública na gestão de recursos hídricos	23
3.6 Os repositórios de código aberto ambientais	24
3.7 Ciência de dados aplicada às questões ambientais	25
3.8 Ferramentas na ciência de dados	26
4. MÉTODOS.....	28
4.1 Caracterização da Área de Estudo	28
4.2 Ciclo de vida da Ciência de Dados	29
5. RESULTADOS.....	33
5.1 Integrar dados provenientes de diferentes fontes para uma compreensão maior da dinâmica da bacia hidrográfica do Rio das Velhas	33

5.2 Implementar uma interface de visualização amigável.....	42
6. CONCLUSÃO	53
7. REFERÊNCIAS BIBLIOGRÁFICAS.....	54

1. Introdução

A Política Nacional de Recursos Hídricos (PNRH), instituída pela Lei nº 9.433 de 1997, estabelece uma gestão descentralizada, participativa e integrada dos recursos hídricos, envolvendo usuários, poder público e sociedade civil. Um dos principais instrumentos dessa política é a criação dos Comitês de Bacias Hidrográficas (CBH), que funcionam como fóruns de discussão sobre os usos múltiplos da água, promovendo uma gestão compartilhada e orientada pelas particularidades de cada região (PNRH, 1997).

O Instituto Mineiro de Gestão de Águas (IGAM) é responsável por monitorar a qualidade e a quantidade das águas em Minas Gerais, gerenciando dados por meio do Sistema Estadual de Informações de Recursos Hídricos (Sisema) (SEMAD, 2023) e dando apoio técnico aos Comitês. Trindade et al. (2013) descrevem o banco de dados brasileiros de recursos hídricos como extenso e complexo, o que torna a extração de informações precisas e compreensíveis uma tarefa desafiadora. Embora a era do Big Data tenha trazido avanços significativos para as ciências ambientais, a interpretação e integração desses dados ainda são difíceis para a sociedade civil, o que limita a gestão participativa proposta pela PNRH.

Diante desse cenário, destaca-se a importância deste trabalho, que visa o desenvolvimento de produtos digitais para analisar a situação atual das Unidades Territoriais Estratégicas (UTE). O propósito principal é elucidar, de maneira clara e objetiva, a evolução das atividades antrópicas e suas consequências no uso do solo, disponibilizando informações acessíveis à sociedade civil. O objetivo final é criar um recurso de monitoramento ambiental participativo, valorizando os dados armazenados nos repositórios brasileiros. Observa-se um impasse significativo relacionado à utilização e integração desses dados, o que ressalta a necessidade de superar desafios para otimizar sua aplicação eficaz em iniciativas como a proposta neste estudo.

2. Objetivos

2.1 Objetivo geral

Desenvolver uma plataforma para integração e visualização de dados de monitoramento ambiental para a bacia hidrográfica do Rio das Velhas – MG.

2.2 Objetivos específicos

- Integrar dados provenientes de diferentes fontes para uma compreensão maior da dinâmica da bacia hidrográfica do Rio das Velhas;
- Implementar uma interface de visualização amigável.

3. Revisão Bibliográfica

3.1 A bacia hidrográfica como unidade de planejamento

O estudo da bacia hidrográfica é um pilar central para a gestão e o planejamento ambiental eficaz. De acordo com Ross e Del Prette (1998), a bacia hidrográfica, independentemente de sua ordem ou tamanho é, por natureza, uma unidade física e natural. Seu elemento fundamental e integrador é a rede de drenagem, composta pelos rios e canais naturais. Os autores alertam que o planejamento exige uma visão que vai além dos aspectos puramente hídricos, devendo integrar o sistema em sua totalidade, o que inclui as componentes naturais (como o relevo, os solos, a flora e a fauna) e as dimensões sociais (englobando as atividades econômicas e as estruturas político-administrativas).

Em consonância com essa visão, o Brasil adotou um modelo de gestão que institucionaliza a bacia hidrográfica como a unidade territorial prioritária. A Política Nacional de Recursos Hídricos (Lei nº 9.433, de 1997) estabeleceu um sistema de gerenciamento que busca a descentralização e a participação. É nesse contexto que surgem as Unidades de Planejamento e Gerenciamento de Recursos Hídricos (UPGRH).

A UPGRH pode ser entendida como a subdivisão geográfica ou um agrupamento de bacias hidrográficas dentro de um estado, criada com o objetivo específico de aplicar os instrumentos da Política Hídrica (cobrança, outorga, plano de recursos hídricos) de maneira mais organizada e eficiente (ANA, 2020). Essa unidade é fundamentalmente ligada aos Comitês de Bacia Hidrográfica (CBHs), órgãos colegiados de natureza consultiva e deliberativa que atuam como o principal fórum de debate e tomada de decisão, reunindo representantes do poder público, dos usuários da água e da sociedade civil. Dessa forma, as UPGRHs concretizam o modelo de gestão compartilhada, transformando o conceito geográfico da bacia em uma unidade de intervenção política e institucional.

3.2 A bacia hidrográfica do rio São Francisco

A Bacia Hidrográfica do rio São Francisco, que abrange aproximadamente 8% do território nacional, é uma região de grande relevância geográfica e ambiental. Com uma extensão de 2.863 quilômetros, sua nascente está localizada na Serra da Canastra, em Minas Gerais, e segue até a foz no Oceano Atlântico, na divisa dos estados de Alagoas e Sergipe. Essa bacia é uma das 12 regiões hidrográficas do Brasil e, para fins de planejamento e gestão, ela foi subdividida em quatro zonas: Alto, Médio, Submédio e Baixo São Francisco, como apresentado na Figura 1 (CBHSF, 2023).

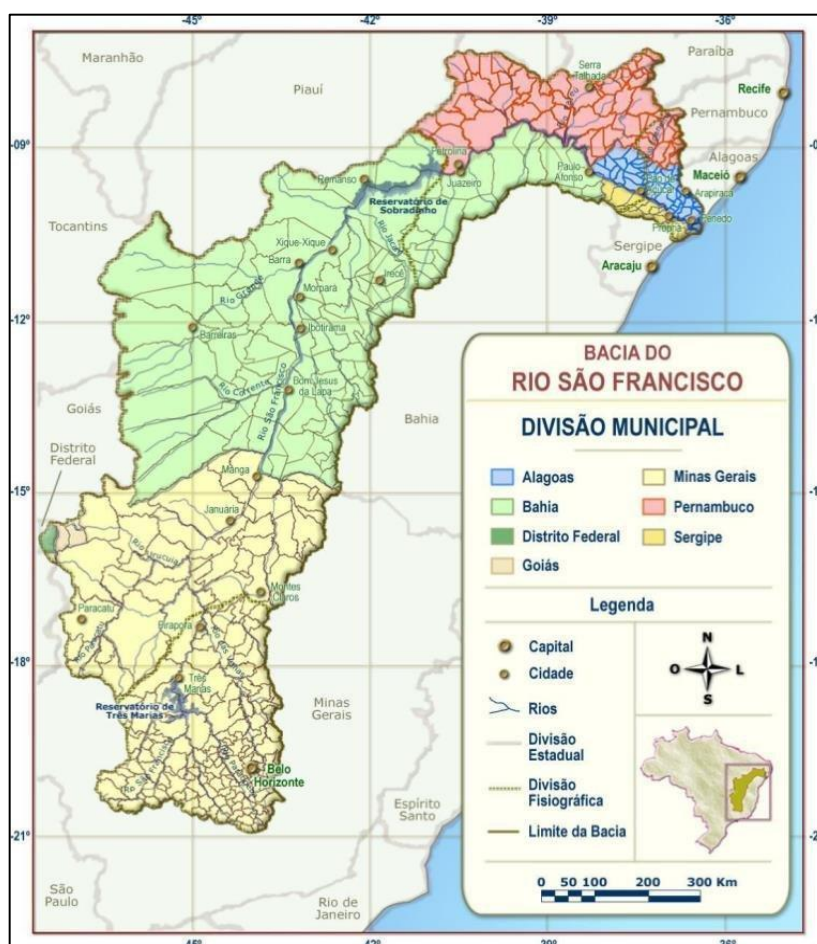


Figura 1 - Divisões fisiográficas da Bacia Hidrográfica do Rio São Francisco (PALMA, 2017).

De acordo com Castro & Pereira (2019), a BHSF é a maior bacia localizada exclusivamente em território brasileiro. O seu uso e cobertura do solo é diversificado devido ao aproveitamento intensivo do rio São Francisco em múltiplas atividades - como navegação, irrigação e pesca - e ocupação indiscriminada de suas margens, ocasionando diversos impactos ambientais que resultaram em um processo de degradação contínua.

Em um cenário de crescimento da demanda hídrica, a preocupação com a garantia de água para os usos múltiplos da bacia repercute-se nas metas, na estratégia e nas diretrizes do Comitê da Bacia Hidrográfica do Rio São Francisco – CBHSF e, a divisão hidrográfica facilita essa administração para que aconteça a conservação dos recursos hídricos, garantindo uma abordagem mais assertiva na sua preservação (CBHSF, 2023).

3.3 A bacia hidrográfica do rio das Velhas

Localizado no Alto São Francisco, com 801 km, o rio das Velhas é o maior afluente em extensão da BHSF. Nasce no município de Ouro Preto, no Parque Municipal das Andorinhas, e deságua no Velho Chico no distrito de Barra do Guaicuy, município de Várzea da Palma (CBH Velhas, 2023).

A população, estimada em 4.406.190 milhões de habitantes (IBGE, 2010), está distribuída nos 51 municípios (Figura 2) cortados pelo rio e seus afluentes e, apesar da região metropolitana de Belo Horizonte ocupar apenas 10% da área territorial da bacia, ela possui 70% de toda a sua população. A região caracterizada pelas atividades industriais e processos de urbanização avançados, o que contribui de forma significativa para a sua degradação. (CBH Velhas, 2023).

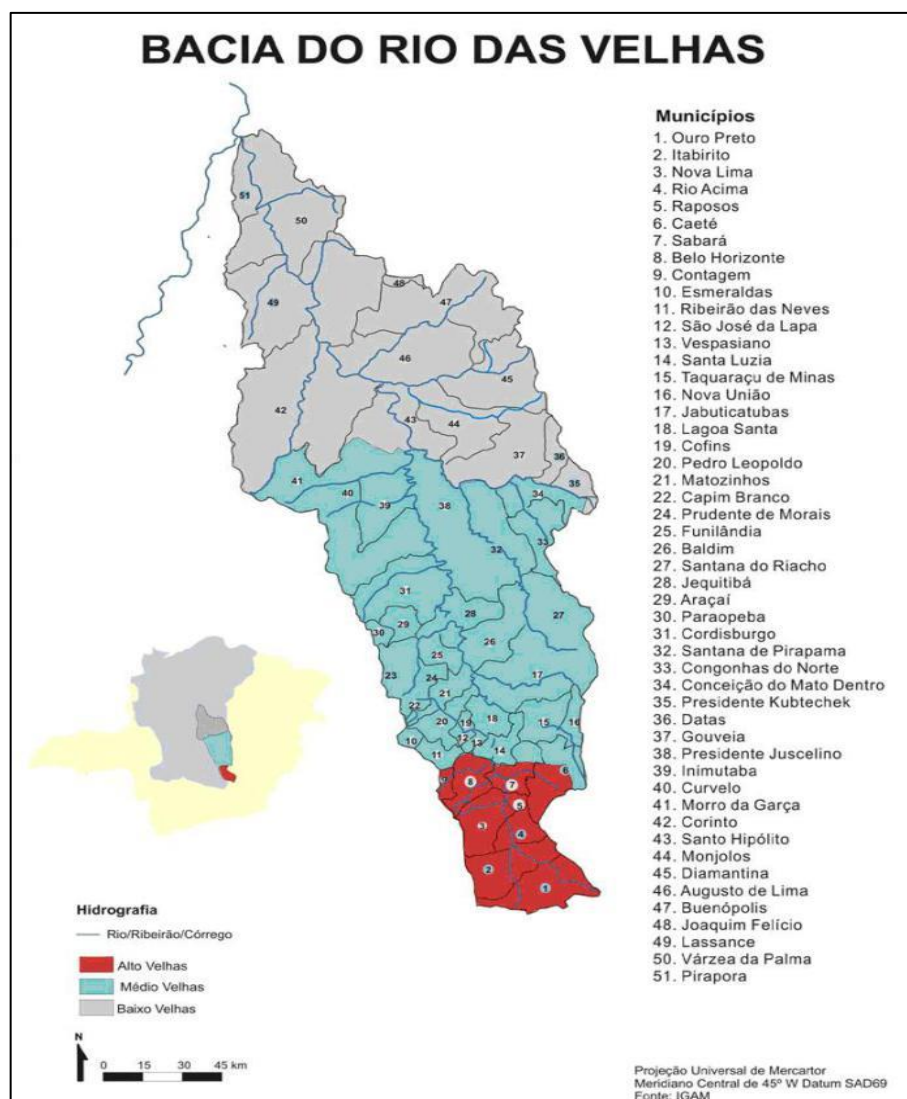


Figura 2 - Municípios que compõem a Bacia do Rio das Velhas e antiga segmentação em Alto, Médio e Baixo Velhas (COSTA, 2008).

Historicamente, a bacia foi segmentada em três grandes divisões, e, com a atualização do plano diretor, foram definidas quatro macrorregiões de planejamento: Alto, Médio (grupo alto), Médio (grupo baixo) e Baixo. A segmentação na bacia do rio das Velhas também foi readaptada em função das características do uso e ocupação do solo na região, o que resulta em diferentes fontes de poluição para a qualidade das águas superficiais em sua extensão como apresentado no Quadro 1 (CALAZANS, 2015).

Quadro 1 - Unidades territoriais da bacia do rio das Velhas, localização e principais características.

Unidade	Localização	Principais características
Alto	Com área total de 2.740 km ² (9,8% da Bacia), o Alto rio das Velhas compreende toda a região denominada Quadrilátero Ferrífero, tendo o Município de Ouro Preto como o limite ao sul e os municípios de Belo Horizonte, Contagem e Sabará como limite ao norte.	Apresenta maior contingente populacional, com uma expressiva atividade econômica, concentrada, principalmente, na RMBH. Nessa região encontra-se o sistema de abastecimento integrado Rio das Velhas (abastece 74% da cidade de Belo Horizonte, além das cidades de Raposos, Nova Lima, Sabará e Santa Luzia). Os principais agentes poluidores são os esgotos industriais e domésticos não tratados e os efluentes gerados pelas atividades minerárias.
Médio – Grupo Alto	A área de drenagem corresponde a 15,4% da Bacia, e compreende 20 municípios.	Apresenta menor concentração populacional que o Alto rio das Velhas, com o predomínio das atividades agrícolas e pecuárias.
Médio – Grupo Baixo	Representa a maior porção dentro da bacia do rio das Velhas (43,8% - 12.204 km ²), e 23 municípios inseridos total ou parcialmente.	
Baixo	Composta por oito municípios e representa o segundo maior grupo (31%).	O aspecto diferenciado desta região é que nenhum dos municípios tem 100% do território inserido na bacia. Baixa concentração populacional com predomínio das atividades agrícolas e pecuárias.

Fonte: CALAZANS, 2015 & CBH Velhas, 2014.

Na bacia hidrográfica do rio das Velhas (BHRV), uma outra divisão fisiográfica adotada é a de Unidades Territoriais Estratégicas – UTE que, segundo a deliberação normativa nº 01, de 09 de fevereiro de 2012, se refere à “área hidrográfica, bacia, grupo de bacias ou sub-bacias contíguas, com características naturais, sociais e econômicas similares a fim de auxiliar o planejamento e gestão da bacia” (Deliberação Normativa 01/CBH Velhas, 2012).

Essa divisão fisiográfica resultou no reconhecimento de 23 regiões de planejamento e gestão (Figura 3), onde cada UTE prevê a implantação de um subcomitê. Os subcomitês são entidades consultivas e propositivas que funcionam obrigatoriamente com a participação dos três segmentos da sociedade: poder público, usuários da água e sociedade civil organizada (CBH Velhas, 2023). Atualmente, há 18 subcomitês instituídos na BHRV.

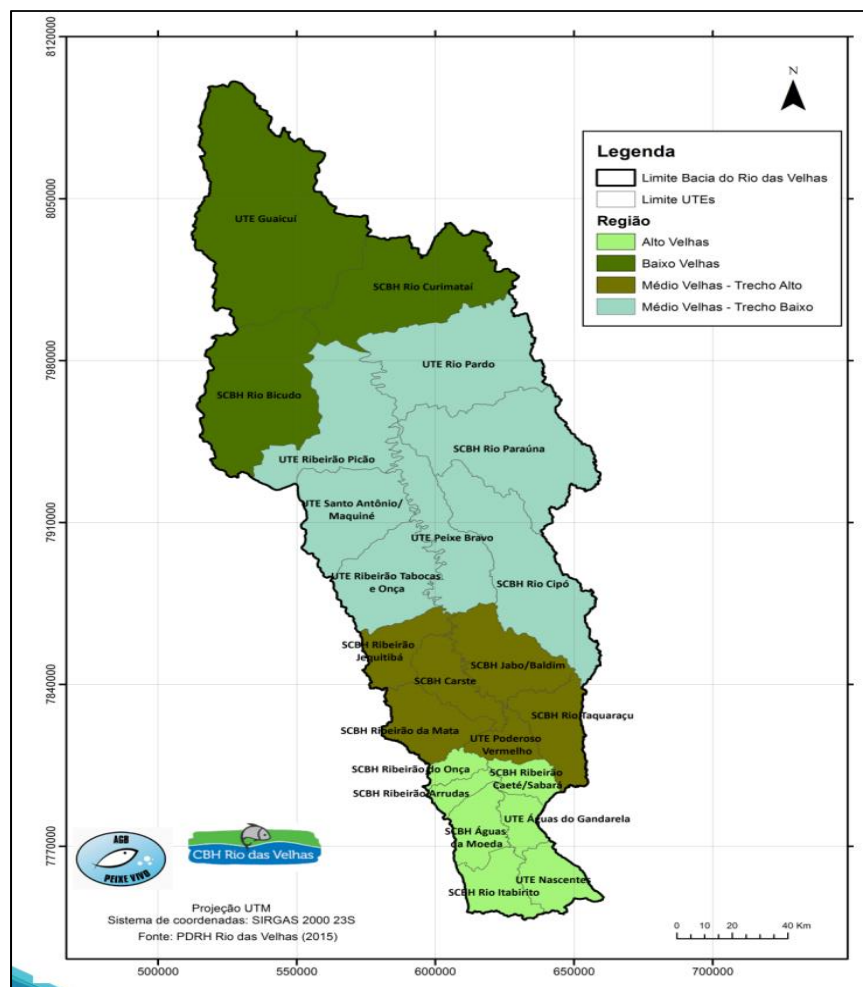


Figura 3 - Unidades Territoriais Estratégicas (UTES) na BH do Velhas (AGÊNCIA PEIXE VIVO e CBH VELHAS, 2015).

3.4 Monitoramento ambiental na bacia hidrográfica do rio das Velhas

Para garantir uma gestão apropriada dos recursos hídricos, é essencial monitorar e avaliar constantemente a qualidade das águas superficiais e subterrâneas, permitindo assim a caracterização e análise temporal das bacias hidrográficas e gerando informações cruciais que podem auxiliar no planejamento, outorga e enquadramento dos recursos hídricos (TRINIDADE *et al.*, 2013).

Na BHRV, com propósito de prestar apoio técnico ao funcionamento dos comitês, o Instituto Mineiro de Gestão de Águas (IGAM) cumpre exigências legais de medição e monitoramento da qualidade e da quantidade de águas (SEMAD, 2023), consolidando estas informações e gerindo o Sistema Estadual de Informações de

Recursos Hídricos – Sisema. Com programas de monitoramento robustos em funcionamento desde 1997, a entidade fornece um histórico de dados sobre a qualidade da água para 100 estações em diferentes pontos da bacia, como mostrado na Tabela 1, e que são disponibilizados em repositórios abertos (IGAM, 2011).

Além dos monitoramentos de qualidade e quantidade cumpridos pelo IGAM, existe a atuação de órgãos ambientais estaduais e municipais em Minas Gerais, que coletam informações abrangentes sobre áreas de preservação, vegetação, fauna e levantamentos sobre uso e ocupação das áreas ao entorno da bacia. Além disso, instituições acadêmicas e de pesquisa (como as universidades federais) desempenham um papel significativo na coleta de dados adicionais, conduzindo estudos científicos que aprofundam o entendimento dos ecossistemas, biodiversidade e impactos ambientais na região (CBH Velhas, 2023).

Como parte dos esforços conjuntos para melhor compreensão e preservação da BHRV, agências do governo federal, como o Instituto Brasileiro do Meio Ambiente e dos Recursos Naturais Renováveis (IBAMA), também participam do monitoramento ambiental em áreas específicas da bacia, com foco especial em áreas protegidas e conservação ambiental (IBAMA e Governo Federal, 2023).

Tabela 1 – Estações de monitoramento do IGAM localizadas na Bacia Hidrográfica do Rio das Velhas e sua distribuição por municípios.

Estação	Classe de Enquadramento	Curso D'água	Latitude	Longitude	Municípios
AV007	Classe 1	Ribeirão Funil	-20,37	-43,63	OURO PRETO (MG)
AV010	Classe 1	Rio das Velhas	-20,24	-43,68	OURO PRETO (MG)
AV020	Classe 2	Rio Maracujá	-20,27	-43,71	ITABIRITO (MG)
AV050	Classe 2	Ribeirão do Silva ou Ribeirão Mata Porcos	-20,36	-43,9	ITABIRITO (MG)
AV060	Classe 2	Ribeirão Carioca	-20,29	-43,81	ITABIRITO (MG)
AV070	Classe 2	Ribeirão Mata Porcos	-20,32	-43,79	ITABIRITO (MG)
AV080	Classe 2	Rio Itabirito	-20,28	-43,8	ITABIRITO (MG)
AV120	Classe 2	Córrego Moleque	-20,18	-43,81	ITABIRITO (MG)
AV160E	Classe 2	Lagoa dos Ingleses ou Represa Lagoa Grande	-20,17	-43,89	NOVA LIMA (MG)
AV180E	Classe 2	Represa da Codorna	-20,18	-43,94	NOVA LIMA (MG)
AV200	Classe 2	Rio do Peixe	-20,12	-43,87	NOVA LIMA (MG)
AV210	Classe 2	Rio das Velhas	-20,09	-43,79	RIO ACIMA (MG)
AV250	Classe 1	Ribeirão dos Macacos	-20,03	-43,83	NOVA LIMA (MG)

AV300	Classe 2	Córrego do Cardoso	-19,98	-43,85	NOVA LIMA (MG)
AV320	Classe 2	Córrego da Mina	-19,98	-43,82	RAPOSOS (MG)
AV340	Classe 1	Ribeirão da Prata	-19,97	-43,8	RAPOSOS (MG)
BV001	Classe Especial	Rio das Velhas	-20,3	-43,55	OURO PRETO (MG)
BV010	Classe Especial	Rio Cipó	-19,34	-43,61	SANTANA DO RIACHO (MG)
BV013	Classe 2	Rio das Velhas	-20,21	-43,74	ITABIRITO (MG)
BV035	Classe 2	Rio Itabirito	-20,23	-43,8	ITABIRITO (MG)
BV037	Classe 2	Rio das Velhas	-20,13	-43,8	RIO ACIMA (MG)
BV041	Classe 1	Ribeirão Cortesia	-20,1	-44	RIO ACIMA (MG)
BV062	Classe 2	Ribeirão Água Suja	-19,98	-43,83	NOVA LIMA (MG)
BV063	Classe 2	Rio das Velhas	-19,98	-43,81	NOVA LIMA (MG), RAPOSOS (MG)
BV067	Classe 2	Rio das Velhas	-19,94	-43,83	SABARÁ (MG)
BV070	Classe 2	Córrego do Galinha	-19,88	-43,78	SABARÁ (MG)
BV076	Classe 3	Ribeirão Sabará	-19,89	-43,8	SABARÁ (MG)
BV080	Classe 3	Rio das Velhas	-19,89	-43,83	SABARÁ (MG)
BV081	Classe Especial	Córrego Clemente ou Córrego do Barreiro	-20	-44	BELO HORIZONTE (MG)
BV083	Classe 3	Rio das Velhas	-19,85	-43,87	SABARÁ (MG)
BV085	Classe 3	Ribeirão Isidoro	-19,82	-43,93	BELO HORIZONTE (MG)
BV105	Classe 3	Rio das Velhas	-19,8	-43,88	SANTA LUZIA (MG)
BV130	Classe 2	Ribeirão da Mata	-19,7	-43,88	VESPASIANO (MG)
BV133	Classe 1	Rio Vermelho	-19,69	-43,6	NOVA UNIÃO (MG)
BV135	Classe 1	Rio Taquaraçu	-19,61	-43,78	JABOTICATUBAS (MG), SANTA LUZIA (MG)
BV136	Classe 1	Rio Jaboticatubas	-19,46	-43,9	JABOTICATUBAS (MG)
BV137	Classe 3	Rio das Velhas	-19,54	-43,9	LAGOA SANTA (MG)
BV138	Classe 3	Rio das Velhas	-19,53	-43,93	LAGOA SANTA (MG)
BV139	Classe 2	Rio das Velhas	-20,06	-43,82	RIO ACIMA (MG)
BV140	Classe 2	Ribeirão Jequitibá	-19,25	-44,05	JEQUITIBÁ (MG)
BV141	Classe 2	Rio das Velhas	-19,02	-44,04	SANTANA DE PIRAPAMA (MG)
BV142	Classe 2	Rio das Velhas	-18,54	-44,18	INIMUTABA (MG), PRESIDENTE JUSCELINO (MG)
BV143	Classe 1	Rio Paraúna	-18,65	-44,04	PRESIDENTE JUSCELINO (MG)
BV144	Classe 2	Ribeirão da Onça	-19,11	-44,32	CORDISBURGO (MG)
BV145	Classe 1	Rio Pardo Pequeno	-18,3	-44,16	MONJOLOS (MG)
BV146	Classe 2	Rio das Velhas	-18,21	-44,35	AUGUSTO DE LIMA (MG), CORINTO (MG)
BV147	Classe 1	Rio Bicudo	-18,13	-44,54	CORINTO (MG)
BV148	Classe 2	Rio das Velhas	-17,61	-44,7	VÁRZEA DA PALMA (MG)
BV149	Classe 2	Rio das Velhas	-17,2	-44,82	VÁRZEA DA PALMA (MG)
BV150	Classe 2	Rio das Velhas	-18,48	-44,2	SANTO HIPÓLITO (MG)
BV151	Classe 2	Rio das Velhas	-17,87	-44,55	LASSANCE (MG)
BV152	Classe 2	Rio das Velhas	-18,31	-44,24	SANTO HIPÓLITO (MG)
BV154	Classe 3	Ribeirão do Onça	-19,82	-43,88	SANTA LUZIA (MG)

BV155	Classe 3	Ribeirão Arrudas	-19,88	-43,86	SABARÁ (MG)
BV156	Classe 2	Rio das Velhas	-19,28	-44,01	BALDIM (MG)
BV158	Classe 2	Ribeirão Cotovelo	-17,75	-44,67	LASSANCE (MG)
BV159	Classe 2	Ribeirão da Corrente	-17,46	-44,68	VÁRZEA DA PALMA (MG)
BV160	Classe 2	Ribeirão das Neves	-19,63	-44,04	PEDRO LEOPOLDO (MG)
BV161	Classe 2	Ribeirão Santo Antônio	-18,72	-44,23	INIMUTABA (MG)
BV162	Classe 1	Rio Cipó	-18,69	-43,99	PRESIDENTE JUSCELINO (MG)
CA002S	Classe 2	Lagoa menor	-19,78	-43,95	BELO HORIZONTE (MG)
CA006S	Classe 2	Lagoa maior	-19,78	-43,95	BELO HORIZONTE (MG)
CA007	Classe 2	Córrego Manoel Pereira ou Córrego Floresta	-19,79	-43,96	BELO HORIZONTE (MG)
CA009	Classe 2	Córrego Manoel Pereira ou Córrego Floresta	-19,79	-43,95	BELO HORIZONTE (MG)
PV005	Classe 2	Córrego do bairro Cinco	-19,94	-44,01	CONTAGEM (MG)
PV037	Classe 2	Córrego Sarandi	-19,89	-44,04	CONTAGEM (MG)
PV065	Classe 2	Córrego Cabral	-19,88	-44,04	CONTAGEM (MG)
PV070	Classe 2	Córrego Sarandi	-19,88	-44,04	CONTAGEM (MG)
PV090	Classe 2	Córrego Ressaca	-19,89	-44	BELO HORIZONTE (MG)
PV105	Classe 2	Córrego da Avenida Tancredo Neves	-19,87	-44	BELO HORIZONTE (MG)
PV110	Classe 2	Córrego Sarandi	-19,86	-44	BELO HORIZONTE (MG)
PV115	Classe 2	Córrego Ressaca	-19,86	-44	BELO HORIZONTE (MG)
PV167	Classe 2	Córrego Bom Jesus	-19,85	-44,01	BELO HORIZONTE (MG)
PV175	Classe 2	Córrego Braúnas	-19,85	-44,01	BELO HORIZONTE (MG)
PV180	Classe 2	Córrego AABB	-19,84	-44	BELO HORIZONTE (MG)
PV190	Classe 2	Córrego Olhos D'água	-19,84	-43,99	BELO HORIZONTE (MG)
PV210	Classe 2	Córrego Mergulhão	-19,86	-43,98	BELO HORIZONTE (MG)
PV220	Classe 3	Ribeirão Pampulha	-19,84	-43,96	BELO HORIZONTE (MG)
PV230	Classe 2	Lagoa da Pampulha	-19,85	-43,99	BELO HORIZONTE (MG)
PV235	Classe 2	Lagoa da Pampulha	-19,86	-43,98	BELO HORIZONTE (MG)
PV240	Classe 2	Lagoa da Pampulha	-19,85	-43,97	BELO HORIZONTE (MG)
SC03	Classe 2	Córrego Caeté	-19,88	-43,67	CAETÉ (MG)
SC10	Classe 3	Ribeirão do Onça	-19,82	-43,9	SANTA LUZIA (MG)
SC12	Classe 2	Ribeirão Areias ou Ribeirão das Areias	-19,81	-44,04	RIBEIRÃO DAS NEVES (MG)
SC13	Classe 2	Ribeirão Areias ou Ribeirão das Areias	-19,81	-44,03	RIBEIRÃO DAS NEVES (MG)
SC14	Classe 2	Ribeirão Poderoso	-19,78	-43,91	SANTA LUZIA (MG)
SC16	Classe 3	Rio das Velhas	-19,72	-43,82	SANTA LUZIA (MG)
SC17	Classe 2	Ribeirão da Mata	-19,69	-43,91	VESPASIANO (MG)
SC19	Classe 2	Ribeirão das Neves	-19,67	-44,05	PEDRO LEOPOLDO (MG)
SC21	Classe 2	Ribeirão da Mata	-19,64	-44,01	PEDRO LEOPOLDO (MG)
SC22	Classe 2	Ribeirão da Mata	-19,58	-44,08	MATOZINHOS (MG)
SC23	Classe 2	Ribeirão da Mata	-19,6	-44,06	PEDRO LEOPOLDO (MG)
SC24	Classe 2	Ribeirão Jequitibá	-19,47	-44,17	PRUDENTE DE MORAIS (MG)

SC25	Classe 2	Córrego do Diogo	-19,47	-44,24	SETE LAGOAS (MG)
SC26	Classe 2	Ribeirão do Matadouro	-19,45	-44,21	SETE LAGOAS (MG)
SC27	Classe 1	Ribeirão do Chiqueiro	-18,44	-43,75	GOUVEIA (MG)
SC28	Classe 1	Ribeirão do Chiqueiro	-18,51	-43,75	GOUVEIA (MG)
SC30	Classe 1	Rio Paraúna	-18,63	-44,06	PRESIDENTE JUSCELINO (MG)
SC33	Classe 2	Rio Curumataí	-18,1	-44,31	AUGUSTO DE LIMA (MG)
SC39	Classe 2	Córrego Matadouro	-18,34	-44,41	CORINTO (MG)

Fonte: Adaptação da autora, dados de monitoramento do IGAM (1997 – 2023).

3.5 Participação pública na gestão de recursos hídricos

De acordo com Coelho et al. (2005), o planejamento do uso dos recursos hídricos deve ser um processo participativo, no qual todas as entidades encarregadas de implementar os planos específicos têm a oportunidade de contribuir ativamente. Isso implica que cada entidade pode influenciar o processo de elaboração, assegurando que suas necessidades e interesses relacionados ao uso da água sejam considerados.

A PNRH reconhece a importância da participação da sociedade civil na gestão dos recursos hídricos e isso se reflete em vários dispositivos da lei por meio de uma gestão sistemática, que envolve diferentes atores, incluindo o poder público, os usuários de água e a sociedade em geral (Governo Federal Brasileiro - Lei nº 9.433, 1997).

Para promover a participação pública destes atores, a PNRH estabelece uma série de instrumentos, sendo o Comitê de Bacia Hidrográfica (CBH) um dos mais conhecidos, e normalizada pela PNRH em seu artigo 36, parágrafo único, que diz: "A composição dos Comitês deve assegurar a participação de todos os segmentos usuários, em igualdade de condições, e garantir a representação dos poderes públicos, dos usuários e das entidades civis de recursos hídricos."

Além disso, a PNRH prevê a criação de sistemas de informação que devem ser acessíveis ao público e fornecer informações atualizadas sobre a situação dos recursos hídricos em cada bacia hidrográfica. Isso permite que a sociedade acompanhe de perto a gestão dos recursos hídricos e contribua com informações relevantes para a tomada de decisões (SNIRH, 2023).

3.6 Os repositórios de código aberto ambientais

Entre as estruturas de compartilhamento e acesso de dados, destacam-se os repositórios de dados de código aberto, os quais não apenas fornecem uma infraestrutura tecnológica para o armazenamento e a disponibilização de dados, mas também desempenham uma função formal de validação das informações contidas nesse ambiente. Para Kindiling e Schirmbacher (2013), “os dados não têm valor sem seus metadados e documentação apropriada que descrevem seus contextos e as ferramentas usadas para criá-los, armazená-los, adaptá-los e analisá-los”.

Na mesma lógica, segundo Sayão e Sales (2016), tornar os dados acessíveis na internet sem uma devida estruturação contextual e semântica, impede a interpretação e a reutilização desses recursos. Essa abordagem também inviabiliza a transferência efetiva do conhecimento que esses dados carregam, bem como a sua capacidade de serem reinterpretados em diferentes contextos. Como consequência direta, isso acarreta a diminuição do valor desses dados para fins de pesquisa.

Sendo assim, além de fornecer suporte a PNRH, os repositórios de código aberto de dados ambientais promovem a transparência, garantia, detalhamento e a colaboração. O recurso atua auxiliando na participação pública efetiva, disponibilizando informações confiáveis que permitam a compreensão dos desafios e oportunidades no que diz respeito ao uso da água. No contexto contemporâneo em que a geração e o uso intensivo de dados têm-se tornado cada vez mais robustos, esta ação abre um leque de perspectivas inéditas para descobertas em todas as áreas do conhecimento (SAYÃO & SALES, 2016), inclusive na área ambiental.

O Portal Brasileiro de Dados Abertos, instituído pelo Decreto n.º 8.777, de 11 de maio de 2016, disponibiliza os dados produzidos e sistematizados em formato aberto, processáveis por máquina e acessíveis a qualquer público, um dos instrumentos contemplados dessa política é o Plano de Dados Abertos (PDA) da ANA, que disponibiliza mais de 100 conjuntos de dados em diferentes formatos, visando ao acesso livre a esses dados por toda a sociedade.

Ainda no cenário institucional do PDA, o Sistema Nacional de Informações sobre Recursos Hídricos (SNIRH) é a base para disponibilização das informações sobre águas no Brasil, agrupadas nas categorias de: divisão hidrográfica, quantidade e

qualidade das águas, usos da água, balanço hídrico, eventos hidrológicos críticos, institucional, planejamento de recursos hídricos, regulação e fiscalização dos recursos hídricos e programas voltados à conservação e gestão (SNIRH, 2023). Instituído como um recurso da PNRH, a iniciativa tem como atribuição coletar, armazenar, integrar e disponibilizar informações, funcionando como um sistema de gestão de dados e informações que abrange todo o território nacional (PNRH, 2023).

Com políticas sólidas de transparência, essas entidades governamentais e ambientais cumprem requisitos essenciais em relação à documentação, disponibilizando metadados detalhados e documentação apropriada que descrevem o contexto e a origem dos dados ambientais armazenados. No entanto, um desafio persistente é a falta de integração entre esses sistemas. Embora forneçam informações valiosas, a falta de correlação entre os repositórios ambientais pode dificultar a busca, a recuperação e a utilização eficaz desses dados. Para promover uma gestão ambiental mais eficiente e colaborativa, é crucial avançar na integração desses sistemas, permitindo um acesso mais amplo e uma análise mais holística dos dados ambientais.

3.7 Ciência de dados aplicada às questões ambientais

No cenário do *Big Data*, onde os dados emergem como ativos valiosos, a ciência de dados desempenha um papel fundamental na busca por atribuir valor aos dados coletados, transformando-os em informações úteis e acionáveis. A ciência de dados deve ser capaz de identificar questões relevantes, coletar informações a partir de diversas fontes, organizá-las e traduzi-las de maneira que impacte positivamente um ecossistema (UNIVERSIDADE DE BERKLEY, 2023).

De acordo com Hal Varian (2023), economista-chefe do Google e professor de ciências da informação, negócios e economia na UC Berkeley, "A capacidade de obter dados – de ser capaz de compreendê-los, de processá-los, de extrair valor deles, de visualizá-los, de comunicá-los – será uma habilidade extremamente importante nas próximas décadas." Esse reconhecimento sublinha a importância crescente da ciência de dados não apenas na academia, mas também na esfera prática, onde a habilidade de gerir, analisar e comunicar informações a partir de dados se torna uma competência crucial. Portanto, a ciência de dados não apenas lida com a coleta e

organização de dados, mas também com a extração de insights e a capacidade de comunicar eficazmente esses insights, promovendo um impacto positivo em diversos setores e domínios.

A ciência de dados é um recurso amplamente utilizado por engenheiros ambientais e frequentemente aplicada na forma de "modelos". A modelagem, definida como um conjunto de hipóteses que buscam explicar ou prever a realidade de um sistema a partir de observações (MONTEIRO, 2000) pode ser aplicada, por exemplo, para prever níveis de poluentes com base em variáveis meteorológicas, para prever riscos geotécnicos, inundações, padrões de biodiversidade e o comportamento de escoamento de aquíferos.

Adicionalmente, a ciência de dados se manifesta em formas mais simplificadas, como na coleta, análise e interpretação de conjuntos de informações. Este processo busca descrever comportamentos a partir de variáveis mensuradas, como é evidenciado no monitoramento da qualidade da água e na descrição e acompanhamento da destinação de resíduos sólidos.

3.8 Ferramentas na ciência de dados

Na ciência de dados, um conjunto diversificado de ferramentas desempenha papéis fundamentais na coleta, análise e visualização de informações ambientais e geoespaciais. O Microsoft Excel, apesar de suas limitações em relação a ferramentas mais avançadas, continua a ser uma escolha popular devido à sua ampla acessibilidade e familiaridade. Neste contexto, o Excel também desempenha um papel vital no armazenamento tabular dos conjuntos de dados, facilitando a organização e a representação dos dados (MICROSOFT, 2023).

Para análises geoespaciais mais avançadas, os Sistemas de Informações Geográficas (SIG) oferecem soluções robustas. O QGIS, um SIG de código aberto, permite análises espaciais avançadas e, com a adição de plugins como o MAPBIOMAS, é possível analisar as mudanças no uso da terra ao longo do tempo. Essas ferramentas são cruciais para compreender as dinâmicas ambientais e geoespaciais, fornecendo insights valiosos para a tomada de decisões (PENTEADO *et al.*, 2018). Além disso, o *Google Earth*, uma outra ferramenta de visualização

geoespacial amplamente reconhecida, permite a exploração de imagens de satélite e a modelagem 3D da Terra, possibilitando sobrepor informações geográficas de maneira intuitiva e informativa.

Linguagens de programação, como Python e R, desempenham um papel essencial na análise de dados ambientais. Ambas oferecem recursos avançados de estatística, modelagem de dados e desenvolvimento de algoritmos personalizados. Python, conhecido por sua versatilidade, também é frequentemente usado na automação de tarefas relacionadas à ciência de dados (MCKINNEY, 2018).

Para acesso a dados em tempo real, as APIs (*Application Programming Interfaces*) desempenham um papel fundamental. A API disponibilizada pela Agência Nacional de Águas (ANA), por exemplo, fornece dados atualizados sobre recursos hídricos e outras variáveis ambientais, tornando-se uma fonte valiosa para análises em tempo real (ANA, 2023).

Por fim, ferramentas como o Power BI, uma ferramenta de *Business Intelligence* da Microsoft®, e a linguagem DAX (*Data Analysis Expressions*), são essenciais para a criação de painéis de controle interativos e relatórios de dados ambientais. Essas ferramentas possibilitam a criação de visualizações dinâmicas que facilitam a comunicação de insights derivados de dados ambientais de maneira clara e eficaz (MICROSOFT, 2023).

4. Métodos

4.1 Caracterização da Área de Estudo

A área de estudo delimitada corresponde à Unidade de Planejamento e Gerenciamento de Recursos Hídricos (UPGRH) do Rio das Velhas (código SF-04). O Rio das Velhas é o maior afluente em extensão da Bacia Hidrográfica do São Francisco (BHSF), desempenhando um papel crucial no abastecimento e na dinâmica hídrica de uma das regiões mais populosas de Minas Gerais.

Com uma extensão de 801 km, o Rio das Velhas está integralmente inserido no estado de Minas Gerais. Sua nascente se localiza no município de Ouro Preto, especificamente no Parque Municipal das Andorinhas. O rio percorre o Alto São Francisco até desaguar no Rio São Francisco, o "Velho Chico", no distrito de Barra do Guaicuy, no município de Várzea da Palma (CBH Velhas, 2023).

A escolha da Bacia Hidrográfica do Rio das Velhas como área de estudo justifica-se pela sua importância socioeconômica e, principalmente, pela maturidade e consistência dos seus programas de monitoramento de recursos hídricos.

A gestão da bacia, coordenada pelo Comitê da Bacia Hidrográfica do Rio das Velhas (CBH Velhas), é sustentada por programas robustos de monitoramento que operam de forma contínua desde 1997. Essa longevidade garante um valioso histórico de dados, essenciais para análises temporais de longo prazo. O Instituto Mineiro de Gestão das Águas (IGAM) disponibiliza informações de qualidade da água para uma rede de monitoramento que abrange, atualmente, 100 estações em diferentes pontos da bacia (IGAM, 2011).

Dessa forma, a vasta e acessível base de dados de domínio público para as estações, detalhadas na Tabela 1, possibilita a execução deste trabalho.

4.2 Ciclo de vida da Ciência de Dados

Para o desenvolvimento da plataforma de monitoramento, foi utilizada a metodologia de cinco etapas do ciclo de vida da ciência de dados.

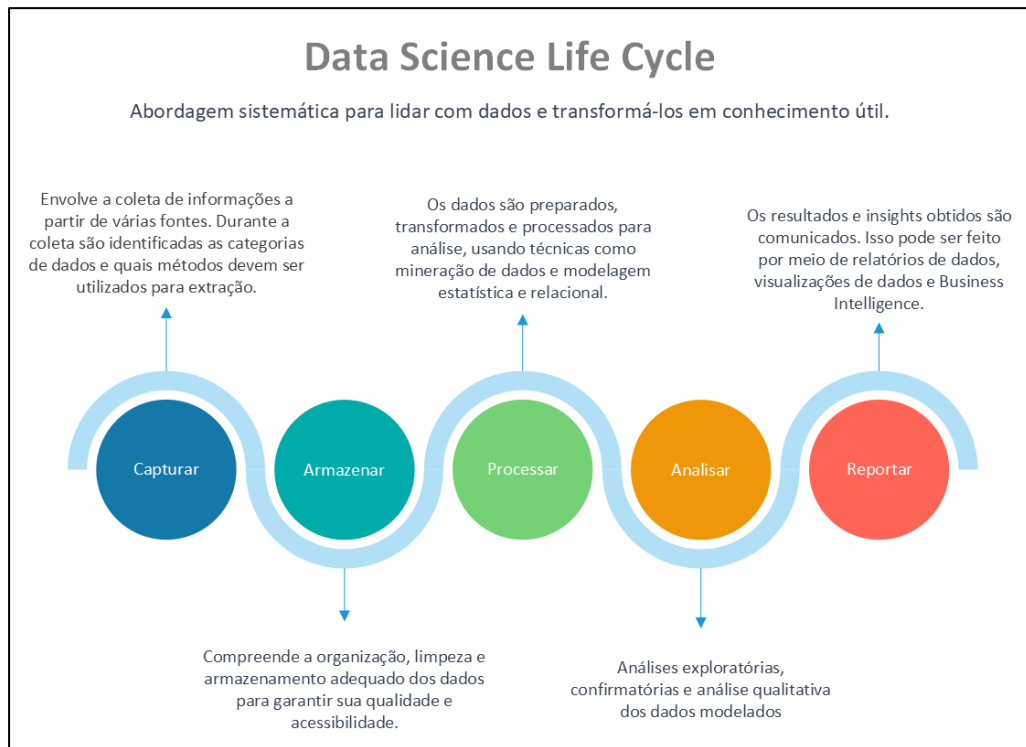


Figura 4 – Cinco estágios do ciclo de vida da ciência de dados (Adaptado de BERKELEY SCHOOL OF INFORMATION, 2023).

4.2.1 Capturar (aquisição e extração de dados)

Após a identificação do problema e delimitação da área de estudo, foi feita a seleção e aquisição de dados *open source* relevantes. Os conjuntos de dados utilizados para o desenvolvimento incluíram:

- Monitoramento da Qualidade das Águas Superficiais - IGAM (Instituto Mineiro de Gestão das Águas): Este conjunto de dados abrange informações sobre a qualidade das águas superficiais, detalhando os diversos parâmetros e métricas monitoradas.
- Pluviometria das Estações Automáticas da ANA (Agência Nacional de Águas): Incorporados os registros de precipitação obtidos das estações automáticas operadas pela ANA e disponibilizados por API. Esses dados são fundamentais para a análise de eventos climáticos e seu impacto nas águas superficiais. O

método é baseado nas requisições HTTP GET do site <https://www.snirh.gov.br/hidroweb/serieshistoricas>.

- c) Uso e Ocupação do Solo - MAP BIOMAS: Disponibilizado por meio de sistemas de informações geográficas, o plugin do MAP BIOMAS, acessado por meio do software de código aberto QGIS, forneceu uma descrição detalhada do uso da terra e padrões de ocupação nos municípios pertencentes à bacia, o que será fundamental para uma análise mais aprofundada das influências do uso do solo na qualidade da água.
- d) Regulamentação Ambiental - Artigos 14, 15 e 16 da Resolução CONAMA 357: Os artigos relevantes da Resolução CONAMA 357 foram coletados para fins de comparação com os resultados dos parâmetros monitorados pelo IGAM em relação aos respectivos Valores Máximos Permissíveis (VMPs). Esta comparação é essencial para avaliar a conformidade regulatória.

Essa abordagem estruturada na aquisição de dados estabeleceu uma base sólida para uma proposta de integração, análise e avaliação da qualidade da água e sua relação com os fatores ambientais, climáticos e socioeconômicos.

4.2.2 Armazenar (armazenamento, limpeza, preparação e arquitetura)

Trabalhando com uma extensão robusta de registros faz-se necessária a implementação de automatização de tarefas para tratamento dos dados. Para esta etapa, foi desenvolvido um script em Python. Este script desempenha uma etapa fundamental no processo de preparação, consistindo na consolidação das múltiplas fontes.

Este processo visa eliminar duplicatas, padronizar textos e filtrar apenas as informações necessárias para o contexto da BHRV. Nesta etapa, também foram modeladas as tabelas que formam a estrutura de um banco de dados relacional, definindo-se as chaves primárias e estrangeiras.

4.2.3 Processar (classificação, clusterização, modelagem e sumarização)

Nesta etapa, buscou-se extrair informações essenciais do conjunto de dados previamente tratado. Foram realizadas métricas que auxiliaram na simplificação e

compreensão dos dados. Aqui, trabalhou-se com uma estrutura relacional baseada em tabelas, e um recurso importante aplicado foi a definição de relacionamentos entre elas para garantir a eficiência na consulta e recuperação de dados.

No Power BI, a definição dos relacionamentos foi semelhante aos relacionamentos em bancos de dados relacionais tradicionais, onde se identificou uma chave primária e uma chave estrangeira. A chave primária foi utilizada para identificar exclusivamente cada linha em uma tabela, enquanto a chave estrangeira foi utilizada para estabelecer a relação com outra tabela.

Com os relacionamentos estabelecidos, é mais fácil criar visualizações e relatórios que agreguem informações de várias tabelas. Isso permite análises mais sofisticadas e insights mais profundos, facilitando a identificação de padrões e tendências nos dados, o que é crucial para o entendimento de um ambiente dinâmico como a Bacia Hidrográfica.

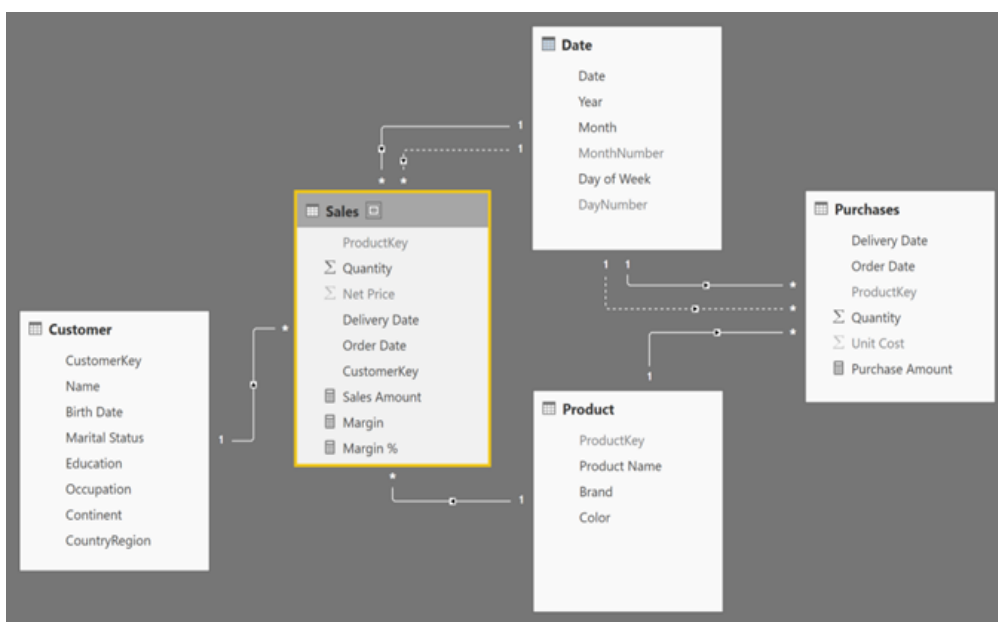


Figura 5 – Exemplo de Data Model no Power BI Desktop (SQLBI, 2018).

4.2.4 Analisar (análise exploratória, confirmatória e qualitativa)

Nesta etapa, os dados foram explorados de maneira mais ampla para identificar padrões, tendências e anomalias. Foram usados alguns recursos gráficos como box plots e histogramas para entender a distribuição dos dados e a qualidade deles.

Aqui, o intuito foi garantir que as conclusões derivadas da análise estejam alinhadas com o conhecimento ambiental no qual o projeto está inserido.

4.2.5 Reportar (visualização e *Business Intelligence*)

Finalmente, a etapa de reportar consiste na construção visual da plataforma, incluindo recursos gráficos quantitativos, qualitativos e georreferenciais que utilizam os dados modelados.

Foram utilizados os recursos do próprio Power BI e suas integrações, criando páginas de análise navegáveis separadas por contexto. Além disso, a experiência do usuário (*User Experience - UX*) foi também considerada durante o desenvolvimento da plataforma. Buscou-se criar uma interface amigável e intuitiva, que permitisse a utilização por qualquer usuário, independentemente do seu nível de conhecimento técnico. Elementos de design como a simplicidade na navegação, a clareza na apresentação dos dados e a interatividade das visualizações foram disponibilizados para garantir que os usuários pudessem acessar e interpretar facilmente as informações.

5. Resultados

5.1 Integrar dados provenientes de diferentes fontes para uma compreensão maior da dinâmica da bacia hidrográfica do Rio das Velhas

Os resultados foram obtidos a partir de um processo de integração de dados provenientes de fontes heterogêneas, visando uma compreensão aprofundada da dinâmica de qualidade hídrica na Bacia Hidrográfica do Rio das Velhas (BHRV).

O procedimento analítico é delineado em três subetapas principais, que espelham o fluxo de tratamento de dados realizado:

1. Capturar (Aquisição e Extração de Dados): Detalha a obtenção das séries históricas de monitoramento de qualidade da água (IGAM), dos dados de uso e ocupação do solo (MAP BIOMAS/QGIS), dos padrões regulatórios (CONAMA 357/2005) e das informações pluviométricas (SNIRH/ANA), fundamentando a base de dados do estudo.
2. Armazenamento, Limpeza, Preparação e Arquitetura: Descreve os processos de higienização, padronização e transformação aplicados aos *datasets* brutos, incluindo a consolidação via script Python e o tratamento detalhado no Power Query (despivotamento, tratamento de sinais de quantificação, filtragem de parâmetros). Esta etapa é crítica para garantir a integridade e a comparabilidade dos dados utilizados.
3. Processar (Classificação, Modelagem e Sumarização): Apresenta a construção do Modelo de Dados Integrado no Power BI, explicitando o relacionamento entre as diversas tabelas (Estações IGAM/ANA, Resultados, Uso do Solo e CONAMA 357) e as lógicas de classificação e cálculo (medidas DAX) empregadas para a análise final e a visualização dinâmica dos resultados.

5.1.1 Capturar (aquisição e extração de dados)

Para a obtenção dos resultados do monitoramento de águas superficiais da Bacia Hidrográfica do Rio das Velhas, foram extraídos 21 conjuntos de dados disponíveis no repositório mantido pelo IGAM.

Sendo o órgão a entidade responsável pelo monitoramento dos recursos hídricos no estado de Minas Gerais os dados são disponibilizados de forma compilada e exibem informações referentes a todas as bacias e sub-bacias do estado em formatos de arquivo como xlsx, xlx e csv. A relação dos documentos obtidos, bem como a extensão dos mesmos, é evidenciada pela Tabela 2.

Tabela 2 - Arquivos obtidos no repositório do IGAM, período de referência dos dados de monitoramento e extensão dos registros.

Período de Referência	Nome do Arquivo e Link de Acesso	Nº de linhas	Nº de Colunas
1997 a 2001	Séries Históricas de Monitoramentos da Qualidade das Águas Superficiais no Estado de Minas Gerais(1997 - 2001)	2786	205
2002 a 2006	Séries Históricas de Monitoramentos da Qualidade das Águas Superficiais no Estado de Minas Gerais (2002 -2006)	5780	205
2007 a 2011	Séries Históricas de Monitoramentos da Qualidade das Águas Superficiais no Estado de Minas Gerais (2007-2011)	9934	205
2012 a 2016	Séries Históricas de Monitoramentos da Qualidade das Águas Superficiais no Estado de Minas Gerais (2012-2016)	14666	205
2017	Séries Históricas de Monitoramentos da Qualidade das Águas Superficiais no Estado de Minas Gerais (2017)	1659	200
1º Trimestre de 2018	Séries Históricas de Monitoramentos da Qualidade das Águas Superficiais no Estado de Minas Gerais (1º TRIM - 2018)	318	206
2º Trimestre de 2018	Séries Históricas de Monitoramentos da Qualidade das Águas Superficiais no Estado de Minas Gerais (2º TRIM- 2018)	314	204
3º Trimestre de 2018	Séries Históricas de Monitoramento da Qualidade das Águas Superficiais no Estado de Minas Gerais (3º TRIM- 2018)	310	204
4º Trimestre de 2018	Séries Históricas de Monitoramento da Qualidade das Águas Superficiais no Estado de Minas Gerais (4º TRIM- 2018)	351	204
1º Trimestre de 2019	Séries Históricas de Monitoramentos da Qualidade das Águas Superficiais no Estado de Minas Gerais (1º TRIM - 2019)	315	204
2º Trimestre de 2019	Séries Históricas de Monitoramentos da Qualidade das Águas Superficiais no Estado de Minas Gerais (2º TRIM - 2019)	310	197
3º Trimestre de 2019	Séries Históricas de Monitoramentos da Qualidade das Águas Superficiais no Estado de Minas Gerais (3º TRIM - 2019)	340	197
4º Trimestre de 2019	Séries Históricas de Monitoramentos da Qualidade das Águas Superficiais no Estado de Minas Gerais (4º TRIM - 2019)	351	203
1º Semestre de 2020	Séries Históricas de Monitoramentos da Qualidade das Águas Superficiais no Estado de Minas Gerais (1º TRIM - 2020)	726	195
2º Semestre de 2020	Séries Históricas de Monitoramentos da Qualidade das Águas Superficiais no Estado de Minas Gerais (3º TRIM e 4º TRIM- 2020)	723	195
1º Semestre de 2021	Séries Históricas de Monitoramentos da Qualidade das Águas Superficiais no Estado de Minas Gerais (1º TRIM e 2º TRIM- 2021)	1488	195
2º Semestre de 2021	Séries Históricas de Monitoramentos da Qualidade das Águas Superficiais no Estado de Minas Gerais (3º TRIM e 4º TRIM- 2021)	1485	195
1º Semestre de 2022	Séries Históricas de Monitoramentos da Qualidade das Águas Superficiais no Estado de Minas Gerais (1º TRIM e 2º TRIM- 2022)	1503	195

Fonte: Autora, 2023.

Para as informações de uso e ocupação do solo, a extração desses dados foi realizada com o auxílio do plugin do MAP BIOMAS disponibilizado pelo software QGIS. A etapa envolveu o carregamento de um shapefile contendo informações geoespaciais e a subsequente extração das áreas de diferentes tipos de uso e ocupação do solo.

A tabela 3 exibe um fragmento exemplo dos atributos extraídos para todos os 51 municípios que compõem a Bacia Hidrográfica do Rio das Velhas.

Tabela 3 - *Dataframe* extraído a partir da tabela de atributos no software QGIS com o auxílio do plugin MAP BIOMAS para o município de Araçáí.

Município	Uso e Cobertura do Solo	Área	Unidade
Araçáí	Pastagem	8.525	km ²
Araçáí	Mosaico de Agricultura e Pastagem	4.417	km ²
Araçáí	Formação Savânica	2.532	km ²
Araçáí	Formação Florestal	2.126	km ²
Araçáí	Silvicultura (monocultura)	594	km ²
Araçáí	Formação Campestre	426	km ²
Araçáí	Área Urbanizada	76	km ²
Araçáí	Campo Alagado e Área Pantanosa	26	km ²
Araçáí	Soja	16	km ²
Araçáí	Rio, Lago e Oceano	15	km ²
Araçáí	Área não Vegetada	0	km ²

Fonte: Autora, 2023.

Para obter os valores estabelecidos pela CONAMA 357 – 2005 para águas superficiais, o documento oficial disponibilizado pelo órgão foi consultado em formato PDF. Utilizou-se uma ferramenta do Excel para converter o PDF em dados tabulares, no formato de linhas e colunas. Os artigos foram então organizados de acordo com as classes de referência a que pertencem, a fim de facilitar a conexão e a comparação dos dados.

Legislacao	Parametro	VMP	Unidade	Classe
CONAMA 357 (2005)-CLASSE 3	Cianobactérias	50000,0000	cel/mL	Classe 3
CONAMA 357 (2005)-CLASSE 2	Cianobactérias	50000,0000	cel/mL	Classe 2
CONAMA 357 (2005)-CLASSE 3	Cloreto	250,0000	mg/L	Classe 3
CONAMA 357 (2005)-CLASSE 3	Bário Total	1,0000	mg/L	Classe 3
CONAMA 357 (2005)-CLASSE 3	Alumínio dissolvido	0,2000	mg/L	Classe 3
CONAMA 357 (2005)-CLASSE 3	Arsênio Total	0,0330	mg/L	Classe 3
CONAMA 357 (2005)-CLASSE 3	Chumbo Total	0,0330	mg/L	Classe 3
CONAMA 357 (2005)-CLASSE 3	Cianeto (íon)	0,0220	mg/L	Classe 3
CONAMA 357 (2005)-CLASSE 3	Fenóis Totais	0,0100	mg/L	Classe 3
CONAMA 357 (2005)-CLASSE 3	Benzeno	0,0050	mg/L	Classe 3
CONAMA 357 (2005)-CLASSE 2	Cloreto	250,0000	mg/L	Classe 2
CONAMA 357 (2005)-CLASSE 2	Bário Total	0,7000	mg/L	Classe 2

Figura 6 – Fragmento dos dados tabulados da CONAMA 357/2005 (Autora, 2024).

Por fim, para obtenção dos dados das estações pluviométricas, foram selecionados 58 códigos de estações disponíveis na Bacia Hidrográfica do Rio das Velhas através do inventário do SNIRH.

O Power BI foi utilizado como ferramenta para realizar as requisições da API. Foi desenvolvido um código na linguagem #m para que a captura fosse feita de forma simultânea para as 58 estações. O script é mostrado na figura 7, com descrições da função executada em cada etapa.

```

let
    // URL base da API da ANA
    UriBase = "https://www.snirh.gov.br/hidroweb/rest/api/estacaotelemetrica?id=",

    // Função para obter dados de cada estação
    GetEstacaoData = (codigoEstacao as text) =>
        let
            // Monta a URL completa para a estação
            Uri = UriBase & codigoEstacao,
            // Obtém o conteúdo da URL
            Source = Web.Contents(Uri),
            // Converte o conteúdo para JSON
            JsonData = Json.Document(Source),
            // Ajuste conforme a estrutura da resposta da API
            Data = JsonData[content]
        in
            Data
in
    GetEstacaoData

```

Figura 7- Código desenvolvido para funcionamento das requisições de forma simultânea para as 58 estações disponíveis na BHRV (Autora, 2024).

5.1.2 Armazenamento, limpeza, preparação e arquitetura

Durante a etapa de processamento de dados para a tabela de resultados do IGAM, as seguintes transformações foram aplicadas para garantir a limpeza e a consistência dos dados.

Primeiramente, os dados foram carregados a partir de um arquivo Excel consolidado através do script python desenvolvido. Cada arquivo foi processado e realocado para um único *dataset*, implicando na unificação de todas as informações, independentemente do tamanho ou formato dos arquivos originais. Como resultado, o script gerou um novo arquivo de saída consolidado, no formato .xlsx.

Definido este arquivo de saída como fonte principal, foi realizada a ingestão do arquivo no Power Query e então iniciada suas manipulações para assegurar a integridade dos dados. Colunas contendo datas e horas foram definidas como *datetime* e *time*, respectivamente. As colunas não dinâmicas, como "Estação", "Data de Amostragem", "Hora de Amostragem", "Arquivo" e "Planilha", foram mantidas intactas, enquanto as outras colunas foram despivotadas. Este processo resultou na criação de duas novas colunas: "Parâmetro" e "Resultado". O "despivoteamento" ou "derretimento" de dados, transformou o *dataset* de formato largo em um formato longo. A redução das colunas resultou a um grande incremento de linhas.

Em seguida a coluna "Resultado" foi então dividida em duas partes utilizando o espaço como delimitador. Esta divisão gerou as colunas "Resultado.1" e "Resultado.2". A razão para essa divisão está no fato de que os resultados dos analitos podem ser reportados com um sinal de menor ou maior, dependendo dos limites de quantificação definidos pelas metodologias laboratoriais. Portanto, essa informação foi mantida e tratada em duas colunas distintas: a coluna "Sinal", que armazena o sinal de menor ou maior, e a coluna "Resultado", que armazena o valor numérico correspondente.

Valores "nan" foram substituídos por *strings* vazias na coluna "Sinal", garantindo que apenas valores válidos fossem considerados. Uma nova coluna, chamada "Resultado Reportável?", foi adicionada para indicar se o valor é reportável ou não, baseada na presença do valor "nan". Além disso, valores "nan" foram substituídos por

strings vazias e pontos foram substituídos por vírgulas na coluna "Valor" para padronizar o formato dos números.

As linhas foram filtradas para remover registros irrelevantes, como aqueles onde o parâmetro é "Condição de tempo", "Profundidade" ou "Ensaio ecotoxicológico", que não foram definidos como parâmetros de interesse. Apenas registros com "Resultado Reportável?" igual a "Y" foram mantidos, garantindo que somente dados válidos e relevantes fossem utilizados na etapa de modelagem.

Na etapa de limpeza, foram identificados os parâmetros que mais diretamente refletem a qualidade da água na bacia do Rio das Velhas. A seleção incluiu o Índice de Qualidade da Água (IQA), uma métrica abrangente que incorpora indicadores físico-químicos e biológicos, oferecendo uma visão holística da saúde dos corpos d'água.

Além do IQA, foram incluídos alguns metais específicos, como chumbo, mercúrio, arsênio, cádmio, ferro, manganês e zinco. A escolha desses metais foi baseada em sua relevância para indicar diferentes fontes de contaminação, relacionadas a atividades industriais, agrícolas, de mineração e urbanas. O critério para a escolha dos parâmetros considerou não apenas a representatividade ambiental, mas também a disponibilidade e consistência dos dados.

A remoção de dados duplicados foi outro passo fundamental. Dados duplicados podem surgir por diversos motivos, incluindo erros no processo de coleta ou na inserção dos dados nos sistemas de monitoramento. A presença de duplicatas compromete a precisão das análises, podendo levar a interpretações errôneas das tendências e padrões da qualidade da água.

Para as demais tabelas utilizadas no estudo, como a de uso e ocupação do solo, legislação e precipitação, foi necessário padronizar informações-chave em comum entre elas. Esta padronização teve como objetivo garantir que as relações entre as tabelas ocorressem de maneira correta e eficiente na próxima etapa.

Por exemplo, adotar o mesmo nome e formato de município tanto nas tabelas de localização das estações quanto nas tabelas de uso e ocupação do solo. Isso permitiu criar uma relação consistente entre as duas tabelas, facilitando a integração e a análise conjunta dos dados. Além disso, foi necessário harmonizar a nomenclatura

dos parâmetros químicos entre a tabela de resultados do IGAM e a tabela de legislação. Esse alinhamento de termos garantiu que as comparações e cruzamentos de dados pudessem ser realizados na etapa seguinte.

Para os dados das estações da ANA os dados capturados foram então expandidos, resultando em várias colunas individuais contendo detalhes como id, código da estação, nome, localização, entre outros. Para garantir a qualidade dos dados, foram removidas duplicatas e linhas em branco, mantendo apenas registros válidos e únicos. O output gerado para as estações é evidenciado na Figura 8.

codigoEstacao	Estação	codigoNomeBacia	codigoNomeSubBacia	altitude	Latitude	Longitude	municipio + estado	nomeRio	operadoraSigla	responsavelSigla	maiorDataPeriodo	menorDataPeriodo
01943084	1943084	4 - Rio São Francisco	41 - RIOS SAO FRANCISCO.DAS VELHAS		-20,00	-43,99	BELO HORIZONTE (MG)	-	CEMIG	CEMIG	07/07/2015	15/05/2015
01943146	1943146	4 - Rio São Francisco	41 - RIOS SAO FRANCISCO.DAS VELHAS	706	-19,90	-43,87	SABARÁ (MG)	-	FERTILIGAS	FERTILIGAS	28/07/2024	12/03/2024
18430000	1843000	1 - Rio Amazonas	18 - RIO AMAZONAS.XINGUI.RIRI.PARU	274	-10,78	-53,10	MARCELÂNDIA (MG)	RIO XINGU	SGB-CPRM	ANA	21/02/2018	08/04/2012
41140080	2043065	4 - Rio São Francisco	41 - RIOS SAO FRANCISCO.DAS VELHAS		-20,21	-43,73	ITABIRITO (MG)	RIO DAS VELHAS	MANG	MANG	29/07/2024	05/02/2010
41151000	2043056	4 - Rio São Francisco	41 - RIOS SAO FRANCISCO.DAS VELHAS	965	-20,30	-43,62	OURO PRETO (MG)	RIO DAS VELHAS	SGB-CPRM	ANA	17/04/2023	18/07/2016
41152000	2043076	4 - Rio São Francisco	41 - RIOS SAO FRANCISCO.DAS VELHAS		-20,24	-43,68	ITABIRITO (MG)	RIO DAS VELHAS	MANG	MANG	29/07/2024	05/02/2010
41180500	2043077	4 - Rio São Francisco	41 - RIOS SAO FRANCISCO.DAS VELHAS		-20,22	-43,80	ITABIRITO (MG)	RIO ITABIRITO	MANG	MANG	03/04/2024	13/03/2015
41192900	2043057	4 - Rio São Francisco	41 - RIOS SAO FRANCISCO.DAS VELHAS	1198	-20,17	-43,89	NOVA LIMA (MG)	RIO DO PEIXE	ANGLO GOLD	ANGLO GOLD	29/07/2024	13/05/2015
41193000	2043079	4 - Rio São Francisco	41 - RIOS SAO FRANCISCO.DAS VELHAS	781	-20,12	-43,87	NOVA LIMA (MG)	RIO DO PEIXE	ANGLO GOLD	ANGLO GOLD	29/07/2024	24/02/2015
41199998	2043128	4 - Rio São Francisco	41 - RIOS SAO FRANCISCO.DAS VELHAS	721	-20,02	-43,82	NOVA LIMA (MG)	RIO DAS VELHAS	SGB-CPRM	ANA	30/11/2021	03/06/2014
41200430	1943148	4 - Rio São Francisco	41 - RIOS SAO FRANCISCO.DAS VELHAS		-19,96	-43,81	RAPOSOS (MG)	RIO DAS VELHAS	SGB-CPRM	ANA	29/07/2024	10/10/2017
41650002	1844010	4 - Rio São Francisco	41 - RIOS SAO FRANCISCO.DAS VELHAS	560	-18,67	-44,19	PRESIDENTE JUSCELINO (MG)	RIO DAS VELHAS	SGB-CPRM	ANA	29/07/2024	16/07/2018
41780002	1844009	4 - Rio São Francisco	41 - RIOS SAO FRANCISCO.DAS VELHAS	560	-18,65	-44,05	PRESIDENTE JUSCELINO (MG)	RIO PARAUNA	SGB-CPRM	ANA	29/07/2024	16/11/2018
41818000	1844001	4 - Rio São Francisco	41 - RIOS SAO FRANCISCO.DAS VELHAS	499	-18,31	-44,23	SANTO HIPÓLITO (MG)	RIO DAS VELHAS	SGB-CPRM	ANA	29/07/2024	01/05/2015
41990000	1744009	4 - Rio São Francisco	41 - RIOS SAO FRANCISCO.DAS VELHAS	488	-17,60	-44,71	VÁRZEA DA PALMA (MG)	RIO DAS VELHAS	SGB-CPRM	ANA	29/07/2024	05/08/2015
41991000	1744047	4 - Rio São Francisco	41 - RIOS SAO FRANCISCO.DAS VELHAS		-17,59	-44,71	VÁRZEA DA PALMA (MG)	RIO DAS VELHAS	CEMIG	CEMIG	07/07/2015	15/05/2015

Figura 8 – Output das requisições realizadas via API da ANA (Autora, 2024).

- A tabela Estações_IGAM está relacionada à tabela Resultados_SF5 através do campo Estação, permitindo a análise dos dados de qualidade da água em relação às estações específicas.
- A tabela Resultados_SF5 está relacionada à tabela CONAMA_357 pelo campo chave_classe_parametro, que permite comparar os resultados monitorados com os parâmetros regulamentados.
- A tabela Estações_ANA está conectada à tabela municípios via o campo municipio + estado, integrando dados de precipitação com informações geográficas municipais.
- A tabela Uso do Solo está conectada à tabela municípios, integrando dados de uso da terra com informações geográficas municipais.

Além dos passos citados, foram criados cálculos usando a linguagem DAX que permitiram agregar e analisar dados de forma dinâmica e personalizada e são recalculadas automaticamente conforme os filtros e segmentações são aplicados no relatório. A figura 10 exemplifica uma medida DAX criada e que foi aplicada diversas vezes na confecção dessa plataforma, nas análises de chuva e estiagem.

```
1 periodo =  
2 VAR MES = 'Resultados SF5'[mês]  
3 RETURN  
4 SWITCH(TRUE(),  
5 MES >= 10, "Chuvoso",  
6 MES <= 3, "Chuvoso",  
7 "Estiagem")
```

Figura 10 – Medida DAX criada para classificar os meses entre período de chuva e estiagem (Autora, 2024).

5.2 Implementar uma interface de visualização amigável

Cada página do relatório Power BI foi projetada para proporcionar uma análise detalhada dos dados de monitoramento ambiental. As conexões entre diferentes fontes de dados e as ferramentas de visualização interativas buscam permitir uma exploração intuitiva, facilitando a análise da dinâmica da bacia hidrográfica do Rio das Velhas. As interfaces criadas são apresentadas na figura 11 e detalhadas adiante.

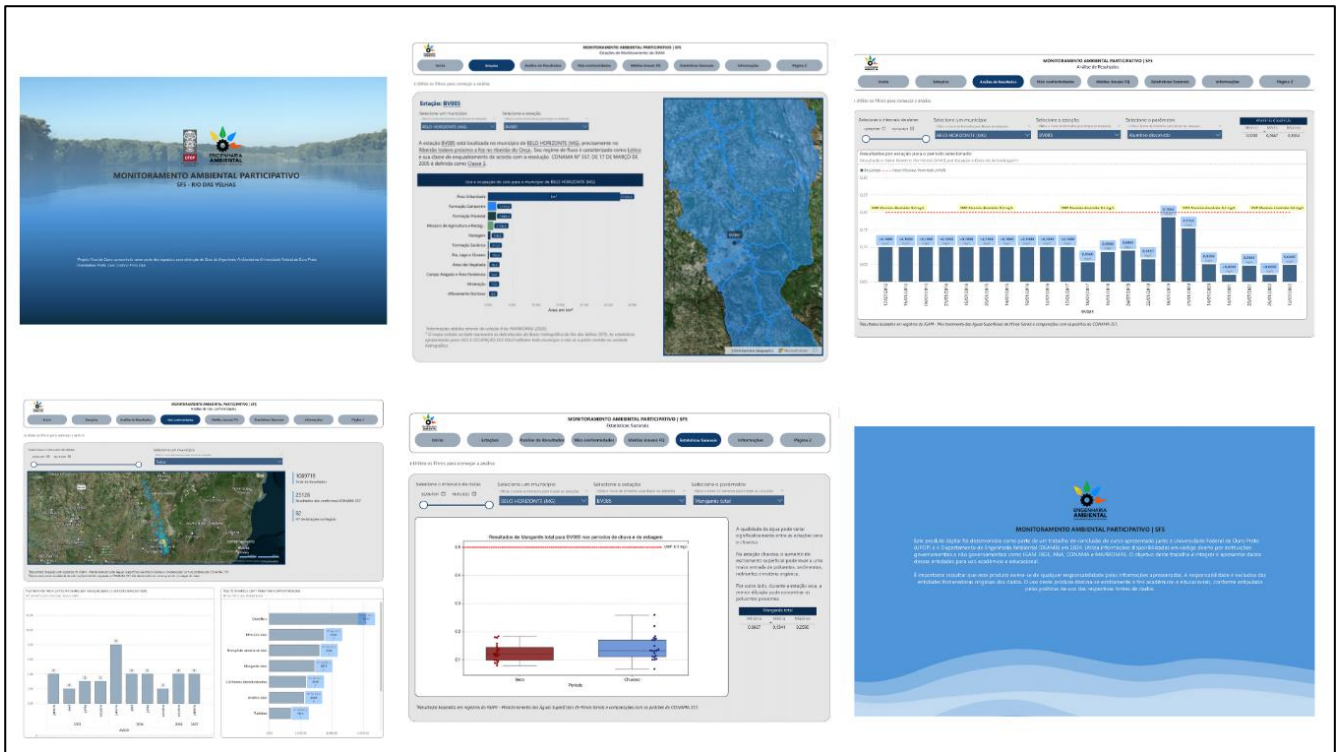


Figura 11 – Interfaces construídas em POWER BI (Autora, 2024).

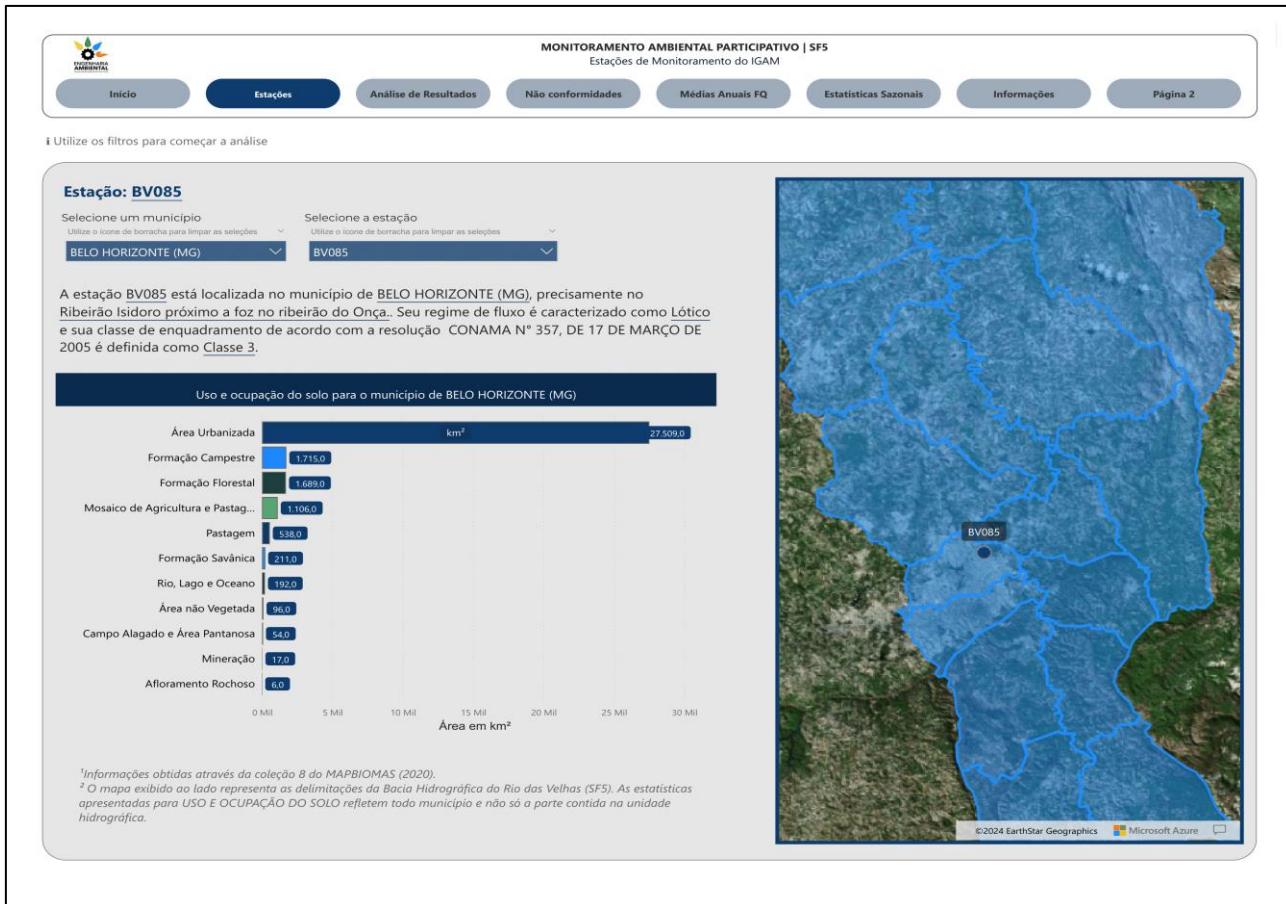


Figura 12 – Interfaces construídas em POWER BI; tela de apresentação da Bacia Hidrográfica (Autora, 2024).

Na página "Estações", evidenciada na Figura 12, há vários elementos que facilitam a interação e a compreensão dos dados. Primeiramente, há os *dropdowns* de seleção de município e estação, que permitem aos usuários escolher o município e a estação específica para visualizar os dados. Em seguida, há uma descrição da estação, que inclui um texto detalhado sobre a estação selecionada, abrangendo sua localização e características do regime de fluxo.

Além disso, um gráfico de uso e ocupação do solo é apresentado, visualizando como o solo é utilizado no município da estação selecionada. Por último, há um mapa geográfico que exibe a localização da estação, proporcionando uma visão espacial do entorno da estação e facilitando a compreensão de sua posição em relação ao município e às características geográficas locais

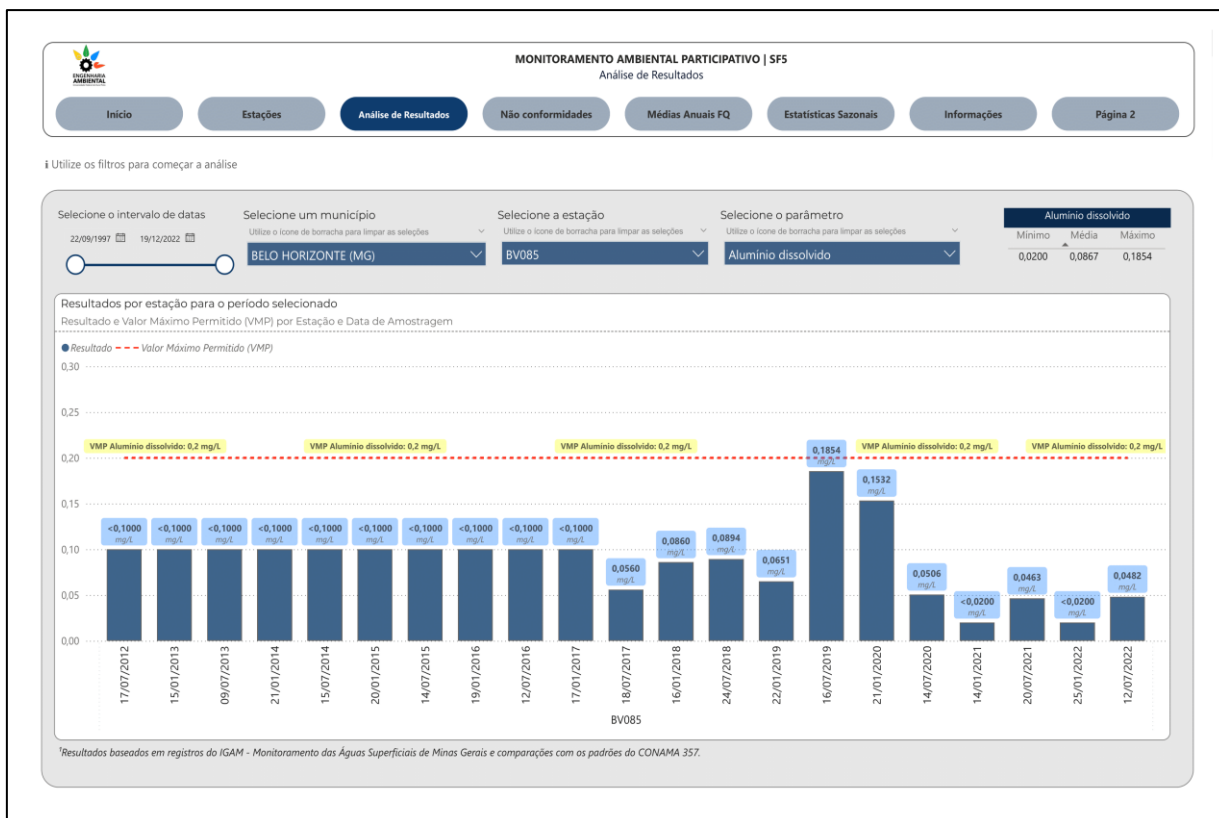


Figura 13 – Interfaces construídas em POWER BI, tela de análise de resultados (Autora, 2024).

A de Análise dos Resultados, mostrada na Figura 13, é dedicada à análise dos resultados das amostragens de qualidade da água, comparando-os com os Valores Máximos Permitidos (VMP) conforme a Resolução CONAMA 357 / 2005. Os filtros permitem selecionar o intervalo de datas, o município, a estação e o parâmetro de qualidade da água para análise.

Foi adicionado um gráfico de barras que exibe os resultados das amostragens ao longo do tempo para a estação e parâmetro selecionados, comparando os valores medidos com os VMPs, indicado por uma linha referência, facilitando a identificação de amostras que excedem os limites regulamentares.

As legislações variam dinamicamente de acordo com a classe de enquadramento da estação selecionada.

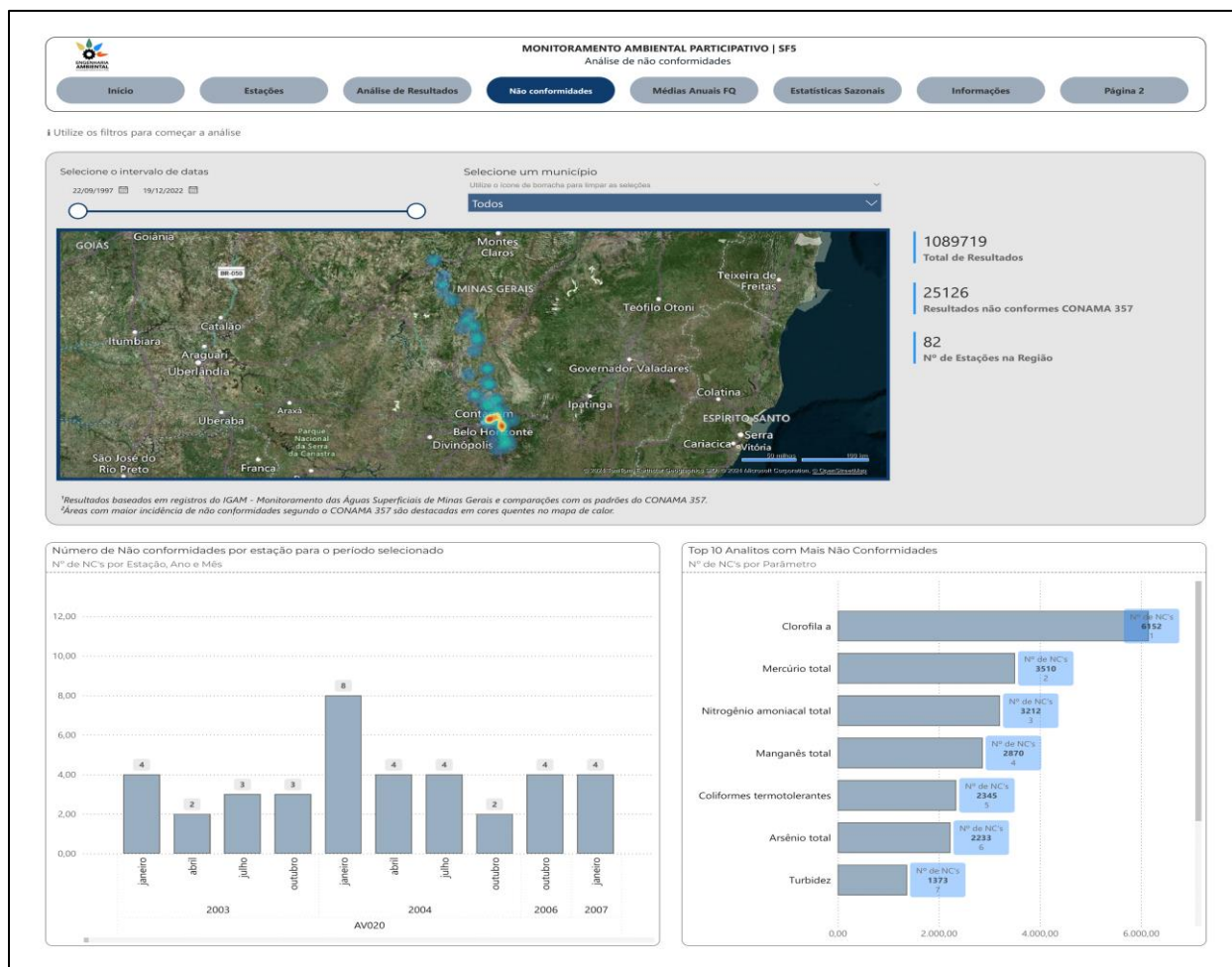


Figura 14 – Interfaces construídas em POWER BI, tela de análise de não conformidades (Autora, 2024).

A página de Não Conformidades, evidenciada na Figura 14, identifica apenas as inconformidades nos dados de qualidade da água em relação aos VMPs estabelecidos. O usuário pode usufruir de filtros de data e de município, sendo evidenciados KPIs nos cartões de dados com o número de resultados e número de estações.

Foi adicionado um mapa de calor onde as áreas com maior incidência de desvios segundo o CONAMA 357/ 2005 são destacadas em cores quentes. O gráfico de barras a esquerda quantifica o número de superações do valor estabelecido pela legislação temporalmente enquanto o gráfico a direita mostra um ranking dos analitos que mais apresentam superações do padrão de qualidade.

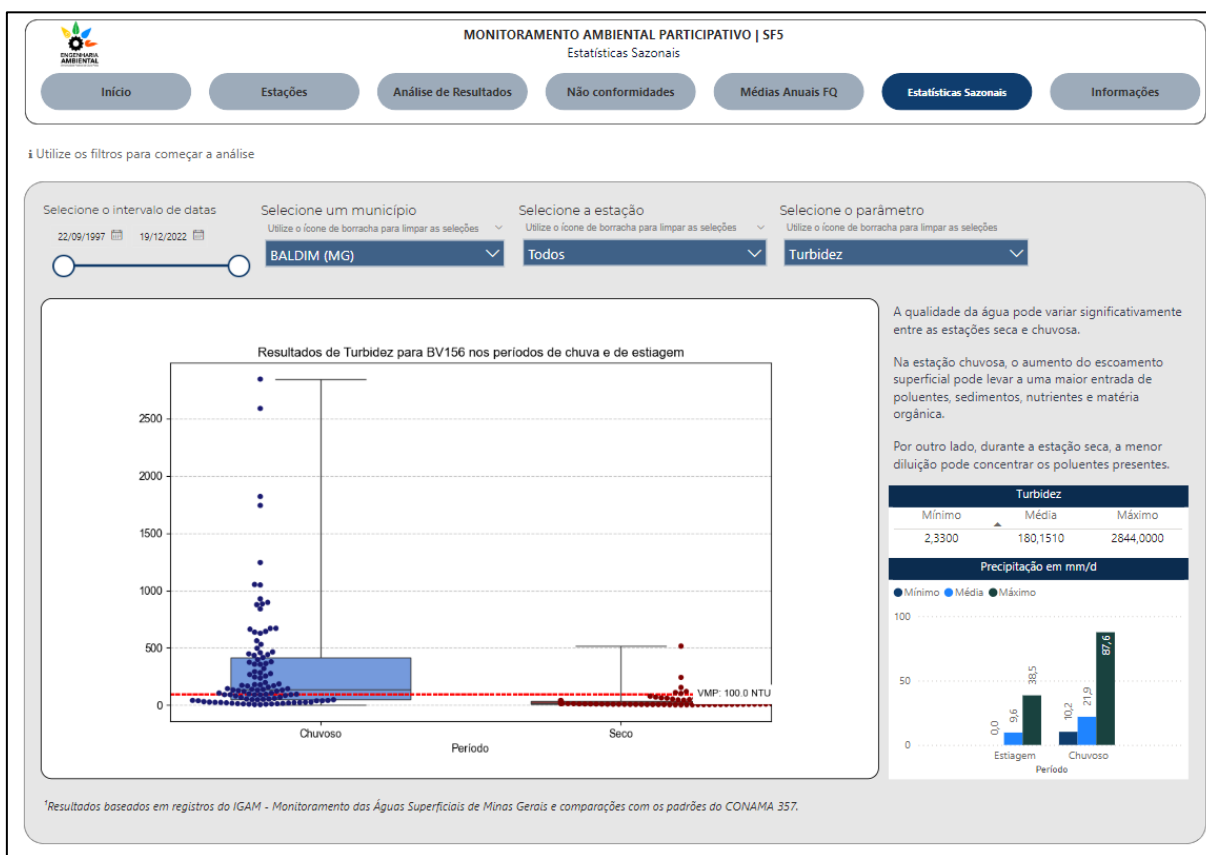


Figura 15 – Interfaces construídas em POWER BI, tela de análise pluviométrica (Autora, 2024).

A última página do projeto, mostrada na Figura 15, apresenta uma análise estatística descritiva dos dados de monitoramento das águas superficiais da bacia comparados com os padrões estabelecidos pela CONAMA 357/2005.

A análise está focada nos períodos de chuva e estiagem, destacando como as condições sazonais afetam a qualidade da água, especialmente, no exemplo, podemos verificar como essa influência é significativa para o parâmetro de Turbidez.

Assim como nas outras visualizações é disponibilizado um controle deslizante que permite selecionar o período de análise, um menu suspenso para escolher o município dentro da bacia, a estação de monitoramento e o analito. O Boxplot, construído pela linguagem Python, visualiza os resultados nos dois períodos distintos e a distribuição

dos dados, incluindo valores mínimos, máximos, medianos e possíveis outliers, também foi adicionado uma linha de referência para o Valor Máximo Permitido (VMP) conforme os padrões da CONAMA.

No canto direito se exibe um gráfico da precipitação média em milímetros por dia durante os períodos de estiagem e chuvoso para os parâmetros selecionados pelo usuário, auxiliando a entender a correlação entre a quantidade de chuva e a concentração dinâmica dos analitos e a variação dos parâmetros físico-químicos.

5.2.1. Performando uma Análise de Qualidade Hídrica na Estação PV220

A interface de visualização desenvolvida evidenciada na Figura 11 permitiu a análise pontual de estações estratégicas. Para o presente exemplo, foi selecionada a estação PV220, localizada em Belo Horizonte (MG), no trecho urbano da Bacia do Rio das Velhas, inserida em um contexto de intensa ocupação e uso do solo como mostra a Figura 16. O uso do solo no município é predominantemente caracterizado por Área Urbanizada, que soma aproximadamente 27.509,0 km² e seguido por Formação Campestre e Formação Florestal, indicando a forte pressão antrópica sobre o corpo hídrico neste ponto.

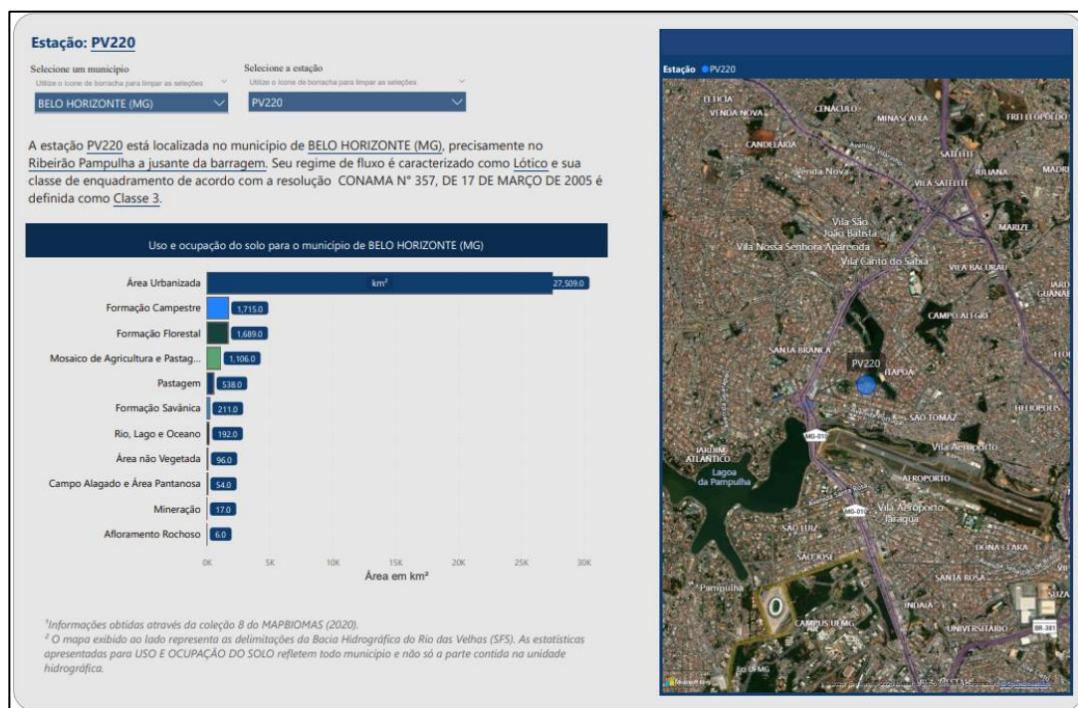


Figura 16 – Localização da estação PV220 na página “Estações” e detalhamento do Uso e Ocupação do Solo na município (Autora, 2024).

A análise na página de "Não Conformidades" na Figura 17 revela o estado de degradação da qualidade da água na estação PV220. De um total de 3.375 resultados registrados na base de dados desde a primeira campanha em 2006, 253 apresentaram violações dos limites estabelecidos pela legislação (Resolução CONAMA 357/2005) para a classe de enquadramento da estação.

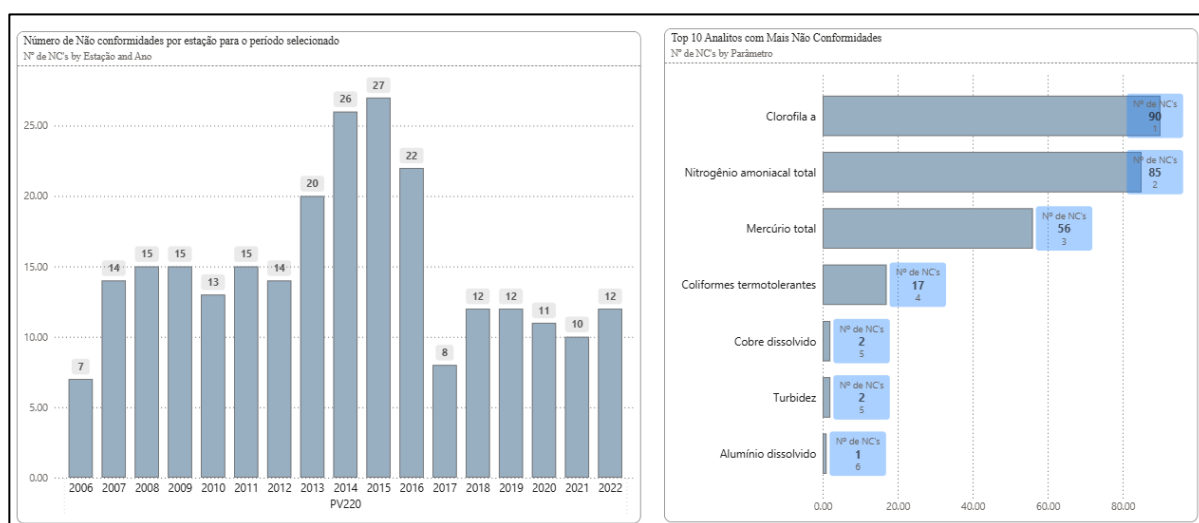


Figura 17 – Não conformidades para a estação PV220, seguindo os limites estabelecidos pela resolução CONAMA 357/2005 (Autora, 2024).

O ranking dos Top 10 Analitos com Mais Não Conformidades destaca a predominância de indicadores relacionados à poluição orgânica, eutrofização e contaminação por metais pesados. A Clorofila-a e Nitrogênio Amoniacal aparecem no topo da lista. Este cenário sugere um quadro de eutrofização, frequentemente causada pelo lançamento de esgotos domésticos não tratados e efluentes ricos em nutrientes que estimulam o crescimento excessivo de algas (SILVA, D. C. V. R. *et al*).

Para a presença de Mercúrio Total, Ferro Dissolvido, Manganês Total e Alumínio Dissolvido indica a provável influência de atividades industriais e/ou de mineração na área de contribuição da estação (COSTA, H. C. T. *et al*). Por fim, as violações de Coliformes Termotolerantes reforçam o diagnóstico de contaminação por esgoto.

Observou-se, ainda, que os picos de não conformidade tendem a se manifestar com maior frequência e intensidade no período chuvoso.

Na página de Análise de Resultados, mostrada na Figura 18, foi dado enfoque no parâmetro Clorofila-a. A série histórica da estação PV220 demonstra a persistência do problema ao longo dos anos (2006 a 2022).

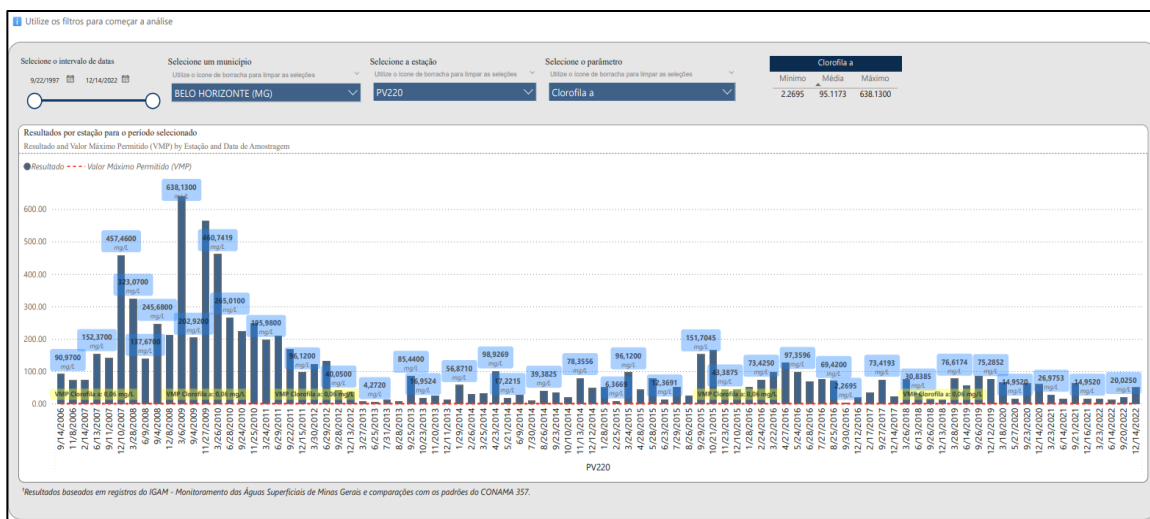


Figura 18 – Série histórica para o parâmetro de Clorofila-a na estação PV220 (Autora, 2024).

A visualização compara o resultado da amostragem com o Valor Máximo Permitido (VMP). Ao longo do período, observam-se diversos picos de concentração, muitos dos quais superam largamente o limite legal. Notavelmente, entre 2008 e 2011, a estação registrou algumas das maiores concentrações.

A alta concentração de Clorofila-a na PV220, localizada no meio do perímetro urbano, pode ser uma consequência direta do alto índice de Nitrogênio Amoniacal e da grande área de Uso Urbanizado no município. A presença constante de contaminantes e o histórico de superação do VMP indicam que o corpo hídrico na PV220 opera consistentemente em uma classe de qualidade inferior àquela que seria ideal ou estabelecida, exigindo ações de saneamento e controle de efluentes na área para reverter o Quadro.

Na Figura 19, apresentam-se as estatísticas de parâmetros físico-químicos, a análise da Turbidez na estação PV220 revelou um padrão de degradação intermitente,

com o registro de um pico máximo de 106,21 NTU em 2013, o que excede o Valor Máximo Permitido (VMP) para enquadramento de corpos hídricos Classe 3 (IGAM, 2024). Este evento de não conformidade pode estar ligado à intensa pressão urbana na área. A Turbidez é um indicador direto do carreamento de material sólido e sedimentar, sendo que picos abruptos em ambientes urbanos são frequentemente correlacionados com processos de erosão acelerada, obras de infraestrutura e o escoamento superficial descontrolado (CETESB, 2018).

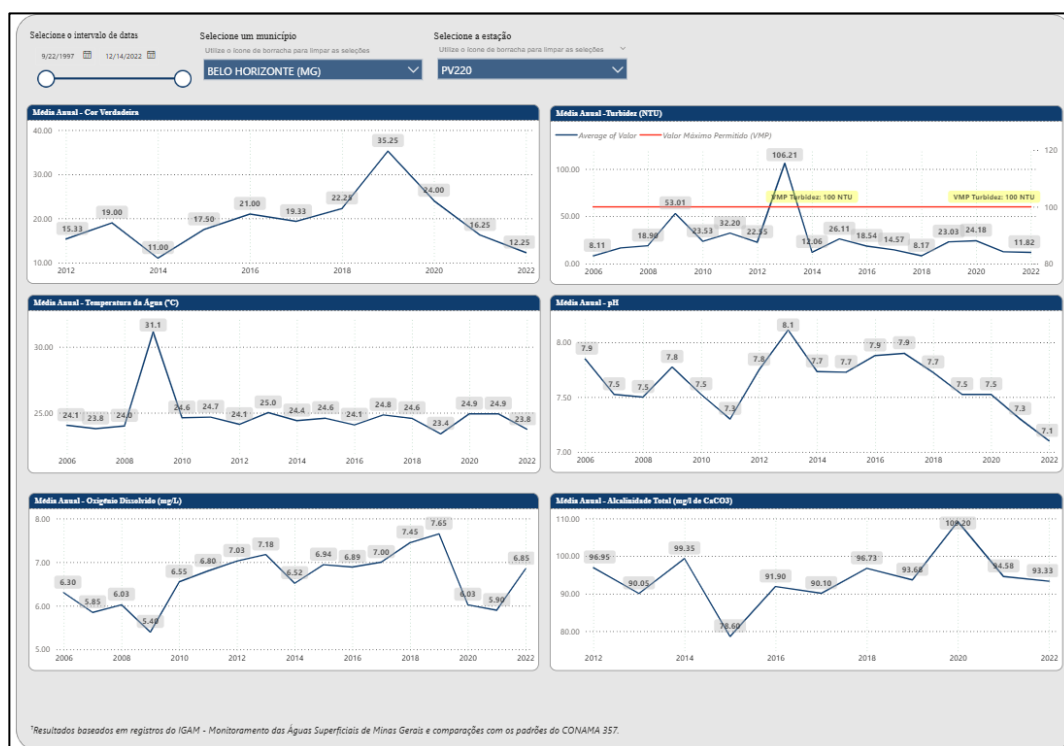


Figura 19 – Estatísticas de FQ para a estação PV220 (Autora, 2024).

O comportamento do Oxigênio Dissolvido (OD) na PV220 reflete provavelmente a luta do corpo hídrico contra a poluição orgânica. A série histórica demonstrou uma recuperação notável nos níveis médios de OD entre 2011 e 2018, atingindo valores próximos a 7,65 mg/L. O OD é o principal indicador da saúde do rio, sendo consumido nos processos de degradação da matéria orgânica; portanto, essa melhoria sugere um impacto positivo de intervenções na bacia.

Contudo, essa tendência positiva foi interrompida a partir de 2019, com uma queda nos níveis de OD para aproximadamente 6,0 mg/L em 2020 e 2021. Essa redução é um possível alerta para o aumento da carga orgânica e a consequente

demanda bioquímica de oxigênio no trecho, indicando que a autodepuração do rio está sendo novamente comprometida. Em ambientes lânticos e lóticos, o baixo OD é a evidência mais clara do excesso de matéria orgânica e da incapacidade do sistema de assimilar a poluição (VON SPERLING, 2005).

Os parâmetros químicos, como pH e Alcalinidade Total, apresentaram grande estabilidade ao longo da série histórica, oscilando em faixas que variam de neutras a ligeiramente alcalinas, entre 7,1 e 8,1. Em relação à Temperatura, apesar de um pico anômalo em 2010, o perfil médio anual se manteve estável na faixa de 24°C a 26°C.

Por fim, a Figura 20 exibe as estatísticas sazonais para o parâmetro de Nitrito, como já reportado, importante indicador de poluição orgânica.

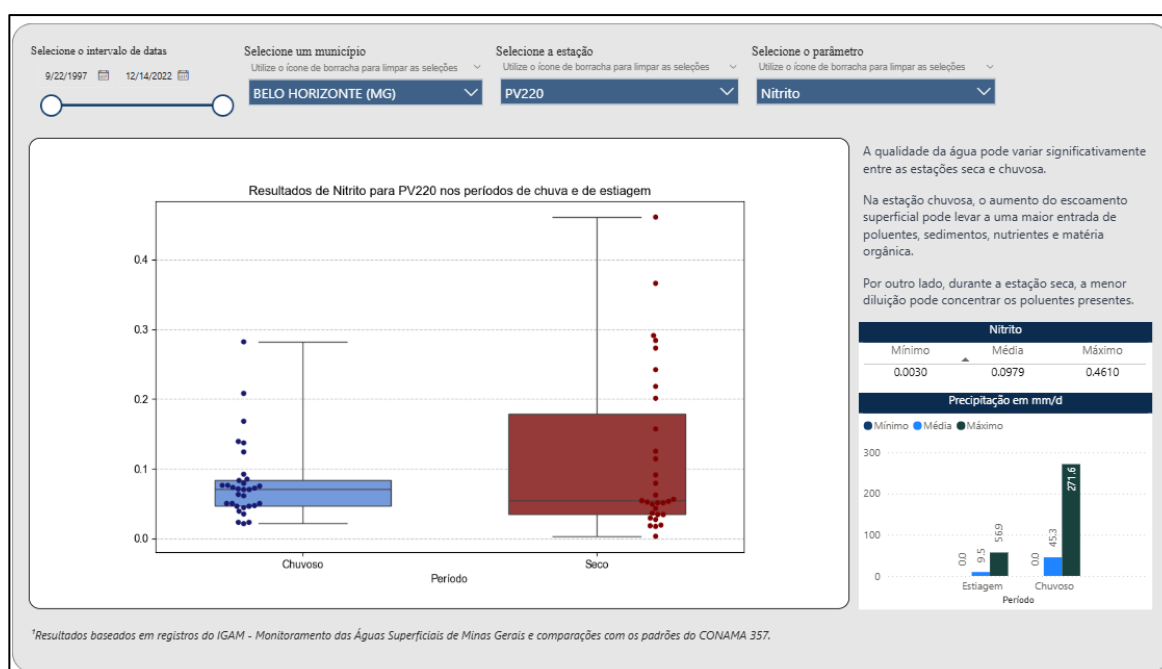


Figura 20 – BoxPlot e estatísticas sazonais para a estação PV220 (Autora, 2024).

O gráfico de barras no canto inferior direito ilustra a precipitação média no período chuvoso sendo significativamente maior do que no período de estio. Esta discrepância pode ser fator chave que modula a concentração de poluentes. Como já dito, a alta precipitação na estação chuvosa promove o arraste de matéria orgânica e poluentes da superfície, mas também proporciona um alto poder de diluição na maioria dos

casos. Por outro lado, a precipitação extremamente baixa no estio leva à baixa vazão, resultando em uma capacidade mínima de diluição e, conseqüentemente, nos picos de concentração de alguns analitos.

O período de seca apresenta uma dispersão muito maior de dados, com a caixa mais alongada e, principalmente, registra o valor máximo absoluto mais alto e diversos *outliers* de alta concentração. O período chuvoso, embora tenha *outliers*, possui um máximo significativamente menor. Ainda, é possível que no período de seca, por causa da provável baixa capacidade de diluição, se registra os picos de poluição por Nitrito, indicando que a menor vazão leva a concentrações potencialmente mais elevadas.

6. Conclusão

No cenário atual, o papel do engenheiro ambiental tem evoluído significativamente. Além de possuir entendimento das questões ambientais, é cada vez mais necessário que esses profissionais tenham habilidades em análise de dados e programação. Esse conhecimento se tornou um diferencial essencial, e no contexto deste trabalho, foi crucial para alcançar os objetivos propostos.

Apesar das dificuldades encontradas, é importante ressaltar que os sistemas e repositórios ambientais brasileiros são robustos e desempenham um papel crucial na democratização do acesso à informação. Pontua-se que apesar de existirem uma infinidade de dados disponíveis, estes não são mantidos em um repositório centralizado, o que dificulta sua coleta e integração.

A bacia hidrográfica é um ambiente extremamente dinâmico, com múltiplos parâmetros influenciando a qualidade e a quantidade de água, como a precipitação e o uso do solo, associados à degradação dos corpos hídricos. A ausência de um repositório centralizado impede uma visão abrangente e integrada dessas variáveis, essencial para uma gestão eficaz e informada dos recursos hídricos.

Além disso, os dados disponíveis muitas vezes carecem de contexto adequado. Um exemplo é que, nos dados publicados pelo IGAM, não foram incluídas as unidades de medida dos analitos, obrigando a busca por essas informações em fontes externas.

Um outro ponto percebido é a insuficiente distribuição das estações de monitoramento na bacia do Rio das Velhas, com uma concentração maior na cidade de Belo Horizonte, o que não reflete adequadamente a complexidade e a diversidade das áreas dentro da bacia. A cobertura limitada das estações impede uma análise detalhada e precisa de todas as regiões, dificultando a implementação de políticas de gestão participativa.

Ademais, este trabalho é um exemplo concreto de como a tecnologia pode ser aplicada as questões ambientais para contribuir para a sustentabilidade e a preservação ambiental, através de uma plataforma de monitoramento participativa.

7. Referências Bibliográficas

Agência Nacional de Águas (ANA). API de Dados Hidro. Disponível em: https://dados.gov.br/dataset/api-de-dados-hidro(https://dados.gov.br/dataset/api-de-dados-hidro).

Agência Peixe Vivo, CBH Velhas & K2. (2017). PLATAFORMA SIGA RIO DAS VELHAS; RELATÓRIO DO MÓDULO AGBMap. Disponível em: https://cdn.agenciapeixevivo.org.br/arquivos/uploads/2017/09/Relatorio_Modulo_Velhas-MAP.pdf(https://cdn.agenciapeixevivo.org.br/arquivos/uploads/2017/09/Relatorio_Modulo_Velhas-MAP.pdf).

Berkeley Edu. (2023). What is Data Science? Disponível em: https://ischoolonline.berkeley.edu/data-science/what-is-data-science/(https://ischoolonline.berkeley.edu/data-science/what-is-data-science/).

BRASIL. Lei nº 9.433, de 8 de janeiro de 1997. Institui a Política Nacional de Recursos Hídricos. Diário Oficial da União, Brasília, DF, 9 jan. 1997. Disponível em: [planalto.gov.br](http://www.planalto.gov.br/ccivil_03/LEIS/L9433.html).

Calazans, G.M. (2015). Avaliação e Proposta de Adequação da Rede de Monitoramento da Qualidade das Águas Superficiais das Sub-Bacias do Rio das Velhas e do Rio Paraopeba Utilizando Técnicas Estatísticas Multivariadas. Dissertação (Mestrado) - Escola de Engenharia, Universidade Federal de Minas Gerais - UFMG.

Castro, C.N.D. & Pereira, C.N. (2019). Revitalização da bacia hidrográfica do rio São Francisco: histórico, diagnóstico e desafios. Instituto de Pesquisa Econômica Aplicada (Ipea). Disponível em: [Repositório do Conhecimento do Ipea](https://repositorio.ipea.gov.br/handle/11058/9351).

CBH Velhas. (2023). CBH Afluentes do Alto São Francisco aprova adesão à Agência Peixe Vivo como sua Agência de Bacia. Disponível em: [agenciapeixevivo.org.br](https://agenciapeixevivo.org.br/noticias/noticias-internas/cbh-afluentes-do-alto-).

CBH Velhas. (2023). CBH Rio das Velhas | A Bacia Hidrográfica do Rio das Velhas. Disponível em: [cbhvelhas.org.br](https://2019.cbhvelhas.org.br/a-bacia-hidrografica-do-rio-das-velhas-).

CBHSF – Comitê da Bacia Hidrográfica do Rio São Francisco. (2023). A Bacia - CBHSF. Disponível em: [cbhsaofrancisco.org.br](http://www.cbhsaofrancisco.org.br).

CETESB. *Qualidade das Águas Interiores no Estado de São Paulo, 2018.* São Paulo: CETESB, 2018.

Coelho, A.C.; Gontijo, W.C.; Cardoso Neto, A. (2005). Unidades de planejamento e gestão de recursos hídricos: Uma proposta metodológica. In: Anais 7º Simpósio de Hidráulica e Recursos Hídricos dos Países de Língua Oficial Portuguesa, Silusba, Portugal.

COSTA, H. C. T.; GOMES, J. C.; MENDES, L. A. Análise da concentração e distribuição de metais pesados na água do Rio das Velhas... **Revista Brasileira de Meio Ambiente**, v. 1, n. 2, p. 70-87, 2018

Kindling, Maxi & Schirmbacher, Peter. (2013). Die digitale forschungswelt als gegenstand der forschung. Information: Wissenschaft & Praxis, v. 64, n. 2-3, p. 127–136.

LEMOS, R. S.; MAGALHÃES JUNIOR, A. P. Dinâmica territorial, transformações ambientais e implicações no manancial de abastecimento público da Região Metropolitana de Belo Horizonte - bacia hidrográfica do alto Rio das Velhas, Minas Gerais. GeoTextos, [S. l.], v. 15, n. 1, 2019. DOI: 10.9771/geo.v15i1.28766. Disponível em:

<https://periodicos.ufba.br/index.php/geotextos/article/view/28766>.

MANUELZÃO. Saúde, Ambiente e Cidadania na Bacia do Rio das Velhas. Revista Manuelzão, n. 80, agosto de 2017. Disponível em: [https://manuelzao.ufmg.br/wp-content/uploads/2018/08/Revista-80_digital.pdf](https://manuelzao.ufmg.br/wp-content/uploads/2018/08/Revista-80_digital.pdf). Acesso em: 10 de janeiro de 2024.

McKinney, W. Data Science for Environmental Scientists in Python. Environmental Science & Technology, v. 52, n. 5, 2018.

ROSS, J. L. S.; DEL PRETTE, M. E. Recursos hídricos e as bacias hidrográficas: âncoras do planejamento e gestão ambiental. **Revista do Departamento de Geografia**, n. 12, p. 89-121, 1998.

SILVA, D. C. V. R. et al. Avaliação da eficiência de um índice de estado trófico na determinação da qualidade da água de reservatórios. **Engenharia Sanitária e Ambiental**, v. 20, n. 4, p. 709-718, 2015

VON SPERLING, Marcos. Introdução à Qualidade das Águas e ao Tratamento de Esgotos (Princípios do Tratamento Biológico de Águas Residuárias; vol. 1). Belo Horizonte: DESA-UFMG, 2005.