

UNIVERSIDADE FEDERAL DE MINAS GERAIS
Escola de Engenharia
Programa de Pós-Graduação em Saneamento, Meio Ambiente e Recursos Hídricos

Fábio Monteiro Cruz

PROGRAMAS DE MONITORAMENTO EMERGENCIAL DE QUALIDADE DA ÁGUA:
redesign de rede através de métodos baseados na Entropia de Shannon

Belo Horizonte
2024

Fábio Monteiro Cruz

**PROGRAMAS DE MONITORAMENTO EMERGENCIAL DE QUALIDADE DA ÁGUA:
redesign de rede através de métodos baseados na Entropia de Shannon**

Tese apresentada ao Programa de Pós-graduação em Saneamento, Meio Ambiente e Recursos Hídricos da Universidade Federal de Minas Gerais, como requisito parcial à obtenção do título de Doutor em Saneamento, Meio Ambiente e Recursos Hídricos.

Área de concentração: Hidráulica e recursos hídricos

Linha de pesquisa: Sistemas de recursos hídricos

Orientadora: Talita Fernanda das Graças Silva

Belo Horizonte
2024

C957p Cruz, Fábio Monteiro.
Programas de monitoramento emergencial de qualidade da água
[recurso eletrônico] : *redesign* de rede através de métodos baseados na
Entropia de Shannon / Fábio Monteiro Cruz. – 2024.
1 recurso online (225 f. : il., color.) : pdf.

Orientadora: Talita Fernanda das Graças Silva.

Tese (doutorado) – Universidade Federal de Minas Gerais,
Escola de Engenharia.

Apêndices: f. 195-225.

Bibliografia: f. 160-194.

1. Engenharia sanitária – Teses. 2. Recursos hídricos – Teses.
3. Desastres ambientais – Teses. 4. Monitoramento eletrônico – Teses.
5. Entropia (Teoria da informação) – Teses. 6. Doce, Rio, Bacia (MG e
ES) – Teses. I. Silva, Talita Fernanda das Graças. II. Universidade
Federal de Minas Gerais. Escola de Engenharia. III. Título.

CDU: 628(043)



UNIVERSIDADE FEDERAL DE MINAS GERAIS
ESCOLA DE ENGENHARIA
PROGRAMA DE PÓS-GRADUAÇÃO EM SANEAMENTO, MEIO AMBIENTE E RECURSOS HÍDRICOS

FOLHA DE APROVAÇÃO

"PROGRAMAS DE MONITORAMENTO EMERGENCIAL DE QUALIDADE DA ÁGUA: Redesign de rede através de métodos baseados na Entropia de Shannon"

FÁBIO MONTEIRO CRUZ

Tese defendida e aprovada pela banca examinadora constituída pelos Senhores:

Professora Talita Fernanda das Graças Silva

Professor Eber José de Andrade Pinto

Professora Gilmare Antônia da Silva

Professor Sergio Koide

Professor Alberto de Freitas Castro Fonseca

Aprovada pelo Colegiado do PG SMARH

Versão Final aprovada por

Profa. Priscilla Macedo Moura

Profª. Talita Fernanda das Graças Silva

Coordenadora

Orientadora

Belo Horizonte, 25 de janeiro de 2024.



Documento assinado eletronicamente por **Talita Fernanda das Graças Silva, Professora do Magistério Superior**, em 25/01/2024, às 19:39, conforme horário oficial de Brasília, com fundamento no art. 5º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Eber Jose de Andrade Pinto, Professor do Magistério Superior**, em 25/01/2024, às 20:35, conforme horário oficial de Brasília, com fundamento no art. 5º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Sérgio Koide, Usuário Externo**, em 26/01/2024, às 11:17, conforme horário oficial de Brasília, com fundamento no art. 5º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Gilmare Antônio da Silva, Usuária Externa**, em 27/01/2024, às 08:30, conforme horário oficial de Brasília, com fundamento no art. 5º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Alberto de Freitas Castro Fonseca, Usuário Externo**, em 29/01/2024, às 09:30, conforme horário oficial de Brasília, com fundamento no art. 5º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Eduardo Coutinho de Paula, Coordenador(a) de curso de pós-graduação**, em 20/03/2025, às 07:42, conforme horário oficial de Brasília, com fundamento no art. 5º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



A autenticidade deste documento pode ser conferida no site https://sei.ufmg.br/sei/controlador_externo.php?acao=documento_conferir&id_orgao_acesso_externo=0, informando o código verificador **2974050** e o código CRC **B3E1DE2F**.

Aos meus amados João Pedro, Bruno Enrico, Isa Belly e Karoline Cruz (in memoriam), que, em momentos distintos, adicionaram novas cores à paleta da minha vida.

AGRADECIMENTOS

Agradeço a minha orientadora, professora Talita Fernanda das Graças Silva, pela confiança e empenho dispensados no desenvolvimento dessa pesquisa. Igualmente agradeço pelos conselhos, paciência, compreensão e competência com que conduziu essa pesquisa, que foram determinantes para a consolidação dessa tese.

A meus professores da Escola de Engenharia que me inspirarão por toda uma vida de atuação profissional.

Aos colegas de curso, pela solidariedade e troca de experiências.

A todos os mestres que tive durante todas as fases de minha educação que me despertaram o sonho de construir uma vida profissional calcada no conhecimento.

A minha família, que sempre acreditou no papel transformador da educação, capaz de prover conhecimento, oportunidades e uma vida digna. Em especial agradeço a meu avô Pedro Mendes, cuja sabedoria de pescador inspirou em mim o amor e o respeito pelas águas.

Aos meus irmãos Cléber e Edson; a minha tia Clara e meu pai João que foram e são meu maior exemplo e inspiração.

A todos os amigos que fiz nessa vida, especialmente aqueles com quem dividi um ombro nos momentos de pesar e um abraço apertado nas conquistas e vitórias.

Ao IFMG, por todo o apoio que me permitiu dedicar-me integralmente ao doutorado.

A UFMG pelo acolhimento e pela experiência enriquecedora. Sempre terei imenso respeito e orgulho por ter sido parte dessa comunidade acadêmica.

A todos que direta ou indiretamente contribuíram com essa jornada!!!

*“(...) Por que se chamavam homens
Também se chamavam sonhos
E sonhos não envelhecem.
(...) De tudo se faz canção
E o coração na curva de um rio rio rio rio...”*

Milton Nascimento/Lô Borges/Márcia Borges

RESUMO

A ocorrência de desastres ambientais de grande escala colocou a gestão ambiental, especialmente a gestão de recursos hídricos, no centro das atenções em todo o mundo. Dependendo da dimensão dos impactos e de sua extensão geográfica, o processo de recuperação ambiental torna-se bastante complexo e desafiante, exigindo muitas vezes a construção de soluções que não estão prontamente disponíveis. No caso da gestão de recursos hídricos, toda a cadeia de ações necessárias à recuperação dos ecossistemas impactados passa pelo estabelecimento de programas e redes de monitoramento da qualidade da água. Apesar de sua importância, as redes de monitoramento emergencial têm sido pouco exploradas nos estudos voltados ao *redesign* (redimensionamento) de redes de monitoramento. Menor ainda, tem sido o uso de abordagens baseadas na Entropia de Shannon, embora haja grande potencial para seu uso. Esta pesquisa avaliou a adequação de métodos para o redimensionamento de redes de monitoramento, baseados na Entropia de Shannon, visando subsidiar o aprimoramento de programas de monitoramento emergencial de qualidade da água em bacias impactadas por desastres ambientais. A bacia do Rio Doce, impactada pelo rompimento de uma barragem de rejeitos de mineração compreendeu o estudo de caso. Foram aplicados métodos calcados na análise do conteúdo de informação através da entropia de Shannon. A segmentação da bacia, por meio de análise de *cluster*, proporcionou a incorporação de características socioambientais subjacentes das subregiões nas análises, além de reduzir a complexidade dos problemas de otimização. Contudo, impôs a necessidade de se evidenciar de forma mais clara a efetividade do método e o efeito da segmentação da rede. “Variáveis não-críticas” foram selecionadas em proporção que variou de 32 a 50% em cada sub-região na bacia do Rio Doce, ilustrando oportunidades para redução da frequência de monitoramento dessas variáveis. Além disso, a entropia de Shannon foi capaz de capturar a variabilidade espaço-temporal do conteúdo de informação das variáveis, mesmo daquelas que podem estar em proporção muito pequena no ambiente. Os métodos de otimização da rede aplicados demonstraram potencial para redução de 20% das estações, sem prejuízo à geração da informação útil à gestão da qualidade da água. Contudo, os resultados sugerem que o método de otimização proposto demanda aprimoramento para melhorar seu desempenho em redes com número expressivo de estações e torná-lo mais objetivo. Tomadores de decisão podem encontrar nesses métodos estratégias de apoio ao gerenciamento de programas de monitoramento emergencial de qualidade da água, podendo melhorar o desempenho das redes e a relação custo-benefício da geração dos dados.

Palavras-chave: redes de monitoramento emergencial da qualidade da água; variáveis não-críticas de qualidade da água; bacia do Rio Doce; desastres ambientais; configuração espacial de redes de monitoramento.

ABSTRACT

The occurrence of large-scale environmental disasters has placed environmental management, particularly water resource management, at the forefront of global attention. Depending on the magnitude of the impacts and their geographical extent, the process of environmental recovery becomes quite complex and challenging, often requiring the development of solutions that are not readily available. In the case of water resource management, the entire chain of actions necessary for the recovery of impacted ecosystems involves establishing programs and networks for monitoring water quality. Despite its importance, emergency monitoring networks have been underexplored in studies focused on resizing monitoring networks. Even less common has been the use of approaches based on Shannon's Entropy, although there is significant potential for its application. This research evaluated the suitability of methods for resizing monitoring networks based on Shannon's Entropy, aiming to support the enhancement of emergency water quality monitoring programs in watersheds affected by environmental disasters. The study case focused on the Doce River watershed, which was impacted by the rupture of a mining waste dam. Methods based on the analysis of the information content of water quality variables monitored by an emergency network were applied to select the set of variables referred to as "non-critical." Furthermore, through the application of an original optimization methodology, an effort was made to select the set of stations corresponding to the arrangement of the best performing network. The prior segmentation of the watershed, through cluster analysis, allowed for the incorporation of underlying socio-environmental characteristics of the sub-regions into the analyses, in addition to reducing the complexity of the optimization problems. However, it imposed the necessity to clearly demonstrate the effectiveness of the method and the impact of segmenting the networks. "Non-critical variables" were selected in proportions ranging from 32% to 50% in each sub-region of the Doce River watershed, illustrating opportunities to reduce the frequency of monitoring these variables. The implemented methodology was able to capture the spatiotemporal variability of the informational content of the variables, even those that may be present in very small proportions in the environment. Regarding the spatial configuration of the analyzed network, there is potential for a 20% reduction of the stations, without compromising the generation of useful information for water quality management. However, the results suggest that the proposed optimization method requires improvement to enhance its performance in networks with a significant number of stations and to make it more objective. Decision-makers may find in these methods strategies to support the management of emergency water quality monitoring programs, potentially improving network performance and the cost-benefit ratio of data generation.

Keywords: emergency water quality monitoring networks; non-critical water quality variables; Doce River watershed; environmental disasters; spatial configuration of monitoring networks.

LISTA DE FIGURAS

Figura 1 - Relação entre diferentes formas de entropia*	44
Figura 2 - Localização geográfica da bacia do Rio Doce	49
Figura 3 – Distribuição dos tipos de solos da bacia do Rio Doce	51
Figura 4 – Usos da terra na bacia do Rio Doce	52
Figura 5 – Trecho impactado pela onda de rejeitos da barragem de Fundão até a UHE Risoleta Neves	56
Figura 6 – Metodologia proposta para o redimensionamento de redes emergenciais de qualidade da água	65
Figura 7 - Estações de monitoramento da qualidade da água da rede PMQQS selecionadas	68
Figura 8 - Comparação dos critérios de quantização das variáveis de qualidade da água	72
Figura 9 – Método de otimização baseado no algoritmo de busca exaustiva tomando funções monoobjetivo hierarquizadas	86
Figura 10 - Grupos de estações de monitoramento da rede PMQQS	88
Figura 11 – Distribuição de estações de subregiões considerando as unidades oficiais de gestão e planejamento hídrico da bacia do Rio Doce	89
Figura 12 - Entropia marginal (a) e entropia máxima (b) por subregiões de qualidade da água durante as estações seca e chuvosa	91
Figura 13 - Distribuição da redundância informacional. A linha vermelha representa o limiar para classificação de variáveis de qualidade da água como VBCI	93
Figura 14 - Número de variáveis de qualidade da água “não-críticas” por subregião	94
Figura 15 - Número de variáveis de qualidade da água “não-críticas” por subregião e período sazonal	97
Figura 16 - Distribuição espacial da redundância informacional de Arsênio dissolvido nas estações seca e chuvosa na bacia do Rio Doce	99
Figura 17 - Distribuição espacial da redundância informacional de Manganês total nas estações seca e chuvosa na bacia do Rio Doce	101
Figura 18 - Matriz de correlação de variáveis de qualidade da água na subregião A	102
Figura 19 - Contribuição de variáveis de qualidade da água para CP1 e CP2 (%), na subregião D	104
Figura 20 - Contribuição de variáveis de qualidade da água para CP1 e CP2 (%), na subregião D	106
Figura 21 - Box-Whisker de “variáveis não-críticas” de qualidade da água com maior contribuição para CP1 por subregião	107
Figura 22 - Box-Whisker de “variáveis não-críticas” de qualidade da água com maior contribuição para CP2 por subregião	107
Figura 23 - Box-Whisker de EWQI na bacia do Rio Doce	109
Figura 24 - Box-Whisker de EWQI na bacia do Rio Doce por período sazonal	111

Figura 25 - Distribuição das classes de qualidade da água definidas para EWQI por período sazonal na bacia do Rio Doce	113
Figura 26 - Contribuição de variáveis de qualidade da água para EWQI na subregião D*	116
Figura 27 - Contribuição de variáveis de qualidade da água para EWQI na subregião L*	118
Figura 28 - Contribuição de variáveis de qualidade da água para EWQI na subregião G*	120
Figura 29 - Contribuição de variáveis de qualidade da água para EWQI na subregião B*	123
Figura 30 - Contribuição de variáveis de qualidade da água para EWQI na subregião M*	124
Figura 31 - Cenários para otimização da rede PMQQS na subregião B*, método FMOHBE/ função monoobjetivo 1	128
Figura 32 - Cenários para otimização da rede PMQQS na subregião C*, método FMOHBE/ função monoobjetivo 1	129
Figura 33 - Cenários para otimização da rede PMQQS na subregião D*, método FMOHBE/ função monoobjetivo 1	130
Figura 34 - Cenários para otimização da rede PMQQS na subregião E*, método MGAMO	131
Figura 35 - Cenários para otimização da rede PMQQS na subregião E, método FMOHBE/ função monoobjetivo 1 (a), 2 e 3 (b)	132
Figura 36 - Séries temporais das estações da subregião E	134
Figura 37 – Distribuição espacial das estações situadas na subregião E	136
Figura 38 - Cenários para otimização da rede PMQQS na subregião G, método MGAMO	138
Figura 39 - Cenários para otimização da rede PMQQS na subregião G, método FMOHBE/ função monoobjetivo 1 (a), 2 (b) e 3 (b)	139
Figura 40 - Cenários para otimização da rede PMQQS na subregião J, método MGAMO	140
Figura 41 - Cenários para otimização da rede PMQQS na subregião J, método FMOHBE/ função monoobjetivo 1 (a), 2 e 3 (b)	141
Figura 42 - Distribuição espacial das estações situadas na subregião J	143
Figura 43 - Séries temporais das estações da subregião J	144
Figura 44 - Cenários para otimização da rede PMQQS na subregião M, método MGAMO	145
Figura 45 - Cenários para otimização da rede PMQQS na subregião M, método FMOHBE/	147
Figura 46 - Distribuição espacial das estações situadas na subregião M	148
Figura 47 - Séries temporais das estações da subregião M	149
Figura 48 – Rede de monitoramento do PMQQS proposta após otimização	150

LISTA DE TABELAS

Tabela 1 – Exemplo de variáveis discretas originais	41
Tabela 2 – Exemplo da etapa 1 da geração de variáveis conjuntas	41
Tabela 3 – Exemplo da etapa 2 de geração de variáveis conjuntas	42
Tabela 4 - Exemplo da etapa 3 de geração de variáveis conjuntas	42
Tabela 5 - Classes de qualidade da água	80
Tabela 6 - Variáveis de qualidade da água “não-críticas” removidas por apresentar medidas em não-conformidade com normas ambientais.....	95
Tabela 7 - Distribuição espacial das variáveis selecionadas como “não-críticas” à qualidade da água	95
Tabela 8 - “Variáveis não-críticas” de qualidade da água investigadas pela ACP	103
Tabela 9 - Variáveis de qualidade da água “não-críticas” listadas dentre as de maior contribuição para CP1	104
Tabela 10 - Variáveis de qualidade da água “não-críticas” listadas dentre as de maior contribuição para CP2	105
Tabela 11 - Estatísticas descritivas de EWQI ao longo da bacia do Rio Doce	108
Tabela 12 - p-valor dos testes de normalidade de Shapiro-Wilk para séries temporais do EWQI em todas as subregiões	110
Tabela 13 - p-valor de teste de Mann Whitney aplicado sobre séries temporais de EWQI por período sazonal	115
Tabela 14 - Largura de <i>bins</i> utilizada na discretização dos valores do EWQI em todas as subregiões	127

LISTA DE ABREVIATURAS E SIGLAS

AC - Análise de *Cluster*

ACP – Análise de Componentes Principais

AD – Análise Discriminante

AF – Análise Fatorial

AHP - Analytic Hierarchy Process

CCME - Water Quality Monitoring Canadian Council of Ministers of Environmental

CT-SHQA - Câmara Técnica de Segurança Hídrica e Qualidade da Água

DBO – Demanda Bioquímica de Oxigênio

ETA – Estação de Tratamento de Água

ETE – Estação de Tratamento de Esgoto

EWQI – Índice de Qualidade da Água Ponderado por Entropia

FDP – Função Densidade de Probabilidade

FMOHBE – Método de Otimização Hierárquico de Funções Monoobjetivo Baseado em Busca Exaustiva

IEMA/ES - Instituto Estadual de Meio Ambiente e Recursos Hídricos do Espírito Santo

IGAM – Instituto de Gestão das Águas de Minas

IQA – Índice de Qualidade da Água

IQAFNS - Índice de Qualidade da Água da Fundação Nacional de Saneamento

LDA – Limite de Detecção Analítica

MGAMO – Método de Otimização Baseado em Procedimento Guloso Aditivo para Função Monoobjetivo

MIMR – Máxima informação/Mínima redundância

MPF – Ministério Público Federal

MT - Mersenne Twister

PMQA – Programas de Monitoramento da Qualidade da Água

PMQQS - Programa de Monitoramento Quali-Quantitativo Sistemático de Água e Sedimentos

POME – Princípio da Máxima Entropia

POMO – Problema de Otimização Multiobjetivo

RMQA – Redes de Monitoramento da Qualidade da Água

SGB – Serviço Geológico do Brasil

SIG – Sistemas de Informação Geográfica

SMC – Simulação de Monte Carlo

SRDD - Scottish Research Development Department Index

TTAC – Termo de Transação e Ajustamento de Conduta

UHE – Usina Hidroelétrica

USGS - United States Geological Survey

VBCI – Variável de Baixo Conteúdo de Informação

WFD - Europe Union Water Framework Directive

ZCAS – Zona de Convergência do Atlântico Sul

SUMÁRIO

1	INTRODUÇÃO	17
1.1	Estrutura e organização do texto	19
2	JUSTIFICATIVA	20
3	OBJETIVOS	22
3.1	Objetivo Geral	22
3.2	Objetivos Específicos	22
4	REFERENCIAL TEÓRICO	23
4.1	Redes e programas de monitoramento da qualidade da água	23
4.2	Planejamento das redes de monitoramento da qualidade da água	24
4.3	Projeto técnico de redes de monitoramento da qualidade da água	26
4.4	Classificação de PMQA	30
4.5	Principais métodos utilizados em projetos técnicos de PMQA	33
4.5.1	Otimização mono e multiobjetivo	37
4.5.2	Entropia de Shannon	39
5	Área de estudo	48
5.1	Caracterização geral	48
5.2	Desastre de Fundão	55
5.3	Considerações sobre o monitoramento e a qualidade da água pré-desastre na bacia do Rio Doce	59
6	MATERIAL E MÉTODOS	63
6.1	Seleção de “variáveis não-críticas” de qualidade da água	63
6.1.1	Seleção e tratamento de dados de qualidade da água	63
6.1.1.1	Critérios de análise espaço-temporal dos dados	68
6.1.1.2	Determinação da redundância informacional	69
6.1.1.3	Quantização das variáveis	70
6.1.1.4	Determinação da entropia máxima usando simulação de Monte Carlo	73
6.1.1.5	Classificação de “variáveis não-críticas” de qualidade da água	74
6.1.1.6	Validação do método para seleção de variáveis de qualidade da água “não-críticas”	75
6.1.2	Seleção da melhor configuração espacial de estações de monitoramento da qualidade da água	76
6.1.2.1	EWQI das subregiões da bacia do Rio Doce	76

6.1.2.2	Discretização de EWQI.....	80
6.1.2.3	Problema de otimização por múltiplas funções monoobjetivo hierarquizadas	81
6.1.2.4	Métodos de otimização aplicados.....	83
7	RESULTADOS E DISCUSSÃO	87
7.1	Seleção de “variáveis não-críticas” na bacia do Rio Doce	87
7.1.1	Limites espaciais e variáveis de qualidade da água analisadas.....	87
7.1.2	Entropia marginal e máxima.....	90
7.1.3	“Variáveis não-críticas” na bacia do Rio Doce.....	92
7.1.4	Variabilidade espaço-temporal das “variáveis não-críticas”	97
7.1.5	Validação da seleção de “variáveis não-críticas”	102
7.2	Seleção da melhor configuração espacial de estações de monitoramento da qualidade da água	108
7.2.1	Subregiões de qualidade da água investigadas e análise exploratória de EWQI	108
7.2.2	Variabilidade espaço-temporal da qualidade da água.....	110
7.2.3	Variáveis de maior contribuição à EWQI	115
7.2.4	Configuração espacial de melhor desempenho para estações da rede do PMQQS.....	126
8	CONCLUSÕES	155
9	RECOMENDAÇÕES.....	157
	REFERÊNCIAS.....	160
	APÊNDICES.....	195

1 INTRODUÇÃO

A água é um dos principais recursos naturais afetados por desastres ambientais, como ruptura de barragens de rejeitos de mineração (Baudson *et al.*, 2021; da Cunha Richard *et al.*, 2020), vazamento ou derramamento de substâncias tóxicas e nocivas (Raven; Georg, 1989; Hou, 2012; Pedrosa, 2007), perfuração de gasodutos (Graham; Wilcox, 2021), derramamento de petróleo e derivados (He *et al.*, 2023), vinhoto (Da Silva *et al.*, 2022a), substâncias radioativas (Koo; Yang; Song, 2014) e outros resíduos perigosos de atividades antropogênicas (por exemplo, Cacciuttolo e Cano (2022) e Guo e Duan (2021)).

Esses desastres são frequentemente súbitos, difíceis de prever e causam sérios impactos ambientais, sociais e econômicos (Hou *et al.*, 2014). Em relação ao comportamento espaço-temporal dos poluentes em corpos hídricos, os impactos podem variar de curto a longo prazo e podem abranger áreas pequenas ou extensas (Gabriel *et al.*, 2021; Semenova, 2020).

Quanto ao efeito, a poluição decorrente de desastres pode promover mudanças na qualidade da água, alterações na morfologia do leito dos rios e biodisponibilidade de poluentes (Miller; Ghisolfi; Barroso, 2023). Os ecossistemas aquáticos também podem sofrer perturbações em seu equilíbrio pela mortalidade de peixes e outros organismos aquáticos, destruição de habitats, mutagenicidade, bioacumulação de poluentes, carcinogenicidade e genotoxicidade (Lusweti *et al.*, 2022; Graham; Wilcox, 2021; Deonarine *et al.*, 2013; Da Silva *et al.*, 2022a).

Além disso, dependendo das características dos poluentes liberados por desastres, pode haver persistência no ambiente e risco para a saúde humana através dos usos da água, sobretudo, o abastecimento humano (Lusweti *et al.*, 2022; Graham; Wilcox, 2021; Deonarine *et al.*, 2013).

Como resposta aos desastres, é crucial que haja rápida tomada de decisão dos gestores, pois a implementação precoce de protocolos de ações pode em certa medida reduzir a intensidade dos impactos e, quando possível, evitar perdas adicionais (Tang *et al.*, 2019; He *et al.*, 2023). Nesse contexto, os dados de monitoramento ambiental, especialmente os dados de qualidade da água,

desempenham um papel importante. Em curto prazo, o monitoramento pode auxiliar na compreensão da dispersão dos poluentes (Ding; Fang, 2019; Pereira *et al.*, 2021). A médio e longo prazo, o monitoramento pode auxiliar a análise do processo de recuperação da qualidade da água, sedimento, ecossistemas aquáticos e a segurança hídrica de bacias impactadas por desastres (Wild *et al.*, 2023; Santos *et al.*, 2023).

Como forma de melhorar o desempenho das redes de monitoramento muitos estudos têm sido conduzidos nas últimas décadas em torno dos principais aspectos que caracterizam a operação das redes, buscando aumentar a eficiência na produção dos dados e reduzir custos decorrentes de redundâncias de projetos (Karamouz *et al.*, 2009a). Eles concentram-se especialmente nas redes regulatórias, possivelmente por representarem a categoria de rede mais usual nos programas de monitoramento.

Essas abordagens têm foco na seleção do melhor conjunto de variáveis monitoradas (Khalil *et al.*, 2010; Barcellos; Souza, 2022), frequência de monitoramento (da Luz; Tobiasson; Kumpel, 2022; Do *et al.*, 2013), número de amostras e localização das estações (Reina-García *et al.*, 2020; de Almeida *et al.*, 2022; Nguyen *et al.*, 2019).

As redes de monitoramento emergencial da qualidade da água geralmente diferem das redes regulatórias. Com frequência, elas abrangem um espectro mais amplo de variáveis (alguns incomuns em redes regulatórias, mas diretamente relacionados a desastres), numa tentativa de abranger todos os possíveis impactos sobre a qualidade da água e os ecossistemas aquático (Wild *et al.*, 2023; Oehrig *et al.*, 2023).

Também é comum, que as redes emergenciais possuam maior densidade de estações de monitoramento com o objetivo de produzir um grande volume de informações espacialmente distribuídas (Shi *et al.*, 2018; UNEP, 2005; Manley *et al.*, 2020; Jing *et al.*, 2019; Mendes *et al.*, 2022; Pacheco *et al.*, 2023).

Face a esse contexto, entende-se que há espaço para contribuição na melhoria do desempenho de rede emergenciais, especialmente considerando que devido as suas características peculiares, pode haver em muitos casos, superdimensionamento dos principais aspectos que caracterizam os projetos das redes. De fato, a carência de estudos nessa tipologia de rede produz uma lacuna científica com potencial efetivo

para aprimoramento da operação das redes emergenciais de monitoramento da qualidade da água.

1.1 Estrutura e organização do texto

Essa tese está organizada da seguinte maneira. O capítulo 1 compreende a introdução do trabalho. Nessa seção são apresentados os conceitos básicos e principais características das redes de monitoramento. O capítulo 2 descreve a questão científica ilustrada na lacuna atualmente existente na investigação científica de métodos voltados ao redimensionamento de redes emergenciais de monitoramento da qualidade da água e ao uso da entropia de Shannon. O capítulo 3 define o objetivo geral e os objetivos específicos da tese. O capítulo 4 faz uma imersão e transita sobre os programas e redes de monitoramento da qualidade da água, descrevendo aspectos técnicos e de planejamento dos projetos. Esta seção ainda apresenta alguns critérios de classificação das redes de monitoramento, demonstra o estado da arte dos principais métodos e técnicas atualmente explorados no redimensionamento das redes e teoriza sobre o conceito das redes emergenciais de monitoramento da qualidade da água. O capítulo 5 apresenta uma descrição das características da bacia do Rio Doce mais pertinentes ao escopo da pesquisa. Adicionalmente há um subitem da seção onde é feita a caracterização do desastre de Fundão e descrita toda a cadeia de efeitos socioambientais que impactaram os recursos hídricos e motivaram o estabelecimento da rede emergencial em operação na bacia. Ao final dessa seção também é feita uma breve caracterização das condições da qualidade da água da bacia do Rio Doce, no período pré-desastre. O capítulo 6 descreve os procedimentos metodológicos propostos para o redimensionamento de redes emergências de qualidade da água. O capítulo 7 apresenta os resultados e analisa o desempenho dos métodos na aplicação realizada na bacia do Rio Doce, assim como suas limitações. Além disso, é feita a caracterização das principais variáveis e problemas relacionados à qualidade da água, no contexto pós-desastre de Fundão, como subsídio à discussão. No capítulo 8 aborda as conclusões deste estudo. O capítulo 9 faz recomendações para pesquisas futuras visando o aprimoramento dos métodos desenvolvidos e aplicados, e aos gestores de programas de monitoramento emergencial de qualidade da água. O capítulo seguinte apresenta as referências que apoiaram a pesquisa e a tese. Por fim, o apêndice contém gráficos e tabelas não

inseridos no texto principal, mas que podem promover melhor compreensão dos resultados e aspectos específicos da discussão.

2 JUSTIFICATIVA

Redes mal planejadas podem produzir informação sem valor analítico, avaliações equivocadas e dados financeiramente onerosos, sem contribuição efetiva ao gerenciamento da água (Strobl *et al.*, 2006a; 2006b). Ainda assim, não há uma metodologia universalmente aceita para o projeto das redes de monitoramento da qualidade da água (RMQA) (Strobl; Robillard, 2008). Não é incomum encontrar redes elaboradas de maneira arbitrária, sem seguir critérios metodológicos lógicos, estruturados e coerentes (Strobl *et al.*, 2006a; 2006b).

Idealmente, uma RMQA não deveria ter caráter definitivo. Isto porque, como destaca Strobl *et al.* (2006a, 2006b) é desejável que o projeto das redes tenha flexibilidade suficiente para permitir a realização de ajustes *à posteriori*.

O processo iterativo de geração de dados é capaz de retroalimentar os programas de monitoramento da qualidade da água (PMQA) com informações suficientes para a identificação de distorções ou redundância. Quando em face dessa análise há implementação de mudanças no conjunto de variáveis monitoradas, ajustes no número de estações ou em sua distribuição espacial ou mudança na frequência de monitoramento, por exemplo, esse processo pode ser chamado de *redesign* (redimensionamento) (Strobl *et al.*, 2006a; 2006b; Strobl; Robillard, 2008).

O redimensionamento de redes pode ter impacto positivo nos PMQA, pois excessos podem ser identificados e corrigidos, gerando racionalização na aplicação dos recursos e eventual realocação para áreas mais críticas (Strobl *et al.*, 2006b).

Muitos autores têm dedicado esforços e colaborado efetivamente com a investigação de métodos para esse fim (de Souza Fraga *et al.*, 2019; Camara *et al.*, 2020; Antanasijević *et al.*, 2018; Mavukkandy; Karmakar; Harikumar, 2014; de Almeida *et al.*, 2022; Singh *et al.*, 2019a; Puri; Karthikeyan, 2017; Salman *et al.*, 2021).

Apesar de avanços nesse campo, a despeito do ceticismo de parte dos gestores dos PMQA, tanto pela crença de que alterações nos programas podem comprometer a

capacidade de se determinar tendências na qualidade da água, quanto pela resistência à ideia de descontinuar o monitoramento de estações fixas (Strobl; Robillard, 2008), as redes emergenciais ainda permanecem à margem despertando pouco interesse da comunidade científica.

Em um estudo de revisão, Nguyen *et al.* (2019) identificaram 14 estudos que investigaram estratégias para a seleção de variáveis de qualidade da água em redes de monitoramento, de um total de 311. Desse universo não houve nenhuma menção explícita que contemplasse as redes emergenciais.

De forma semelhante, Jiang *et al.* (2020a) realizaram revisão sobre diversos aspectos relacionados ao projeto/redimensionamento de redes de monitoramento da qualidade da água e destacaram que do universo de 90 trabalhos inventariados houve ampla predileção pelas redes regulatórias com 78 contribuições identificadas, enquanto as redes emergenciais foram investigadas em apenas 9 casos.

Apesar de não haver uma definição formal de redes ou programas emergenciais de qualidade da água, eles podem ser caracterizados pela geração de informação sobre fontes de poluição decorrentes da ocorrência de emergências ambientais (CCME, 2015), independentemente do tempo de monitoramento necessário a essa finalidade. Nessa perspectiva, o caráter “emergencial” se dá pela motivação que levou ao estabelecimento das redes (no caso os desastres ambientais) e pelo gerenciamento a longo prazo de seus impactos, e não apenas pelo caráter de urgência do monitoramento.

3 OBJETIVOS

3.1 Objetivo Geral

Avaliar a adequação de métodos para redimensionamento de redes de monitoramento, como estratégia para o aprimoramento de programas de monitoramento emergencial de qualidade da água em bacias impactadas por desastres ambientais.

3.2 Objetivos Específicos

- Selecionar conjunto de variáveis de menor relevância no contexto do acompanhamento dos efeitos dos desastres ambientais.
- Analisar a aplicabilidade do índice de qualidade da água ponderado por entropia (EWQI) na caracterização sistemática da qualidade da água sob efeito da poluição decorrente de desastres ambientais.
- Analisar o desempenho de uma nova metodologia para otimização da configuração espacial de redes de monitoramento emergencial integrantes de PMQA.

4 REFERENCIAL TEÓRICO

4.1 Redes e programas de monitoramento da qualidade da água

As atividades econômicas e o crescimento populacional, se por um lado geram riqueza e desenvolvimento, por outro também geram passivos sobre os ecossistemas e serviços ambientais, desequilibram a relação demanda-disponibilidade dos recursos naturais, aumentam a produção de resíduos e a poluição ambiental (Baste; Watson, 2022; Kong; Khan, 2019; Chen *et al.*, 2020; Radelyuk *et al.*, 2021; Sun; Wang, 2023).

A escalada da degradação ambiental que chegou a patamares críticos no mundo contemporâneo motivou a realização de fóruns e encontros mundiais para debater as formas de enfrentamento às questões ambientais. Com efeito, foram estabelecidos acordos internacionais com propostas e metas para redução da poluição ambiental baseados no conceito de “*sustentabilidade*” (Aguila, 2020; Ojima; Melillo, 2022; Seyfang, 2003; Strong, 1992; Lau; Lee; Mohamed, 2012), onde destaca-se como uma das estratégias para seu alcance o gerenciamento integrado dos recursos naturais, incluindo-se os recursos hídricos (UNEP, 2012).

Ponto chave do chamado “*Gerenciamento integrado de recursos hídricos*” é a promoção de PMQA. Através deles são geradas informações confiáveis sobre o *status quo* da qualidade da água que subsidiam os tomadores de decisão no planejamento hídrico e ambiental das bacias hidrográficas (Behmel *et al.*, 2016; Vélez-upegui *et al.*, 2020; Bartram; Ballance, 1996; Strobl *et al.*, 2006b).

Os objetivos dos programas podem ser divididos para efeito conceitual em objetivo de gerenciamento e objetivos de qualidade. Os objetivos de gerenciamento referem-se aos aspectos técnicos do monitoramento que estão fortemente calcados no funcionamento do sistema ambiental investigado, enquanto os objetivos de qualidade relacionam-se ao conhecimento do sistema, propriamente (Soares, 2001).

O monitoramento da qualidade da água, para fins dessa pesquisa, compreende a caracterização e quantificação da poluição e seus efeitos, através de medidas de variáveis integrantes de três grandes grupos: variáveis físicas, químicas e biológicas. A principal importância dessa atividade deve-se à capacidade em apoiar efetivamente o gerenciamento da água e das bacias hidrográficas, inclusive permitindo a

comparação com padrões de referência para avaliação do *status quo* da qualidade da água (Marinho *et al.*, 2022; Razman *et al.*, 2022; Von Sperling, 2007; Tan, 2020).

A informação sobre a qualidade da água das bacias abrangidas por PMQA vem da operação de RMQA. Essas redes se materializam fisicamente no conjunto de estações de monitoramento espacialmente distribuídas na bacia com fins de produzir informações sistematizadas sobre a qualidade da água (Soares, 2001).

O projeto das redes deve atender a certos requisitos básicos e recursos disponíveis. Esses aspectos são fundamentais para tornar as informações confiáveis, representativas das diferentes regiões monitoradas na bacia e comparáveis entre si (Behmel *et al.*, 2016; Bartram; Ballance, 1996; Vélez-upegui *et al.*, 2020).

Soares (2001) observa que o projeto de redes de monitoramento pode ser dividido em etapa de planejamento e etapa técnica. A estratégia de monitoramento e os objetivos das redes compõe a etapa de planejamento, enquanto a etapa técnica se ocupa dos critérios técnicos das redes.

4.2 Planejamento das redes de monitoramento da qualidade da água

A definição dos objetivos do monitoramento influenciam o dimensionamento dos recursos materiais e humanos necessários à implantação/operação das redes (Behmel *et al.*, 2016). Não há de fato, uma única motivação para esse fim, mas frequentemente as redes são planejadas para atender múltiplos objetivos específicos (Strobl; Robillard, 2008; Bartram; Ballance, 1996). A análise dos dados gerados pelo monitoramento pode revelar novos *insights* não observados nos objetivos inicialmente propostos.

Alguns desses objetivos são: estabelecimento de condições de referência dos sistemas hídricos, identificação de não-conformidades com as normas e padrões ambientais, regulação dos usos da água e lançamentos de efluentes, subsídio a programas de controle ambiental de atividades poluidoras na bacia hidrográfica, avaliação de mudanças de longo prazo nas condições naturais da qualidade da água, avaliação da efetividade de ações específicas de recuperação ambiental nas bacias e avaliação do impacto de poluição acidental nos mananciais e ecossistemas aquáticos

(desastres ambientais), dentre outros (Strobl; Robillard, 2008; European Union, 2003; Bartram; Ballance, 1996).

Face aos objetivos definidos é necessário planejar as melhores estratégias para o monitoramento e o sistema de armazenamento e manipulação dos dados de qualidade da água. As estratégias de monitoramento incluem (Soares, 2001):

- a) A forma como os dados gerados pela rede serão tratados, sistematizados e armazenados;
- b) A elaboração de relatórios para apresentação aos gestores dos PMQA;
- c) A frequência com que serão elaborados os relatórios.

Os procedimentos e infraestrutura de manipulação dos dados compõe o sistema de informações de dados de qualidade da água dos PMQA (Soares, 2001), que além de incorporar a etapa técnica do projeto das redes, prevê:

- a) Identificação da demanda por informações para o planejamento, tomada de decisão e atendimento à legislação hídrica e ambiental vigente;
- a) Inventário de todas as informações disponíveis sobre a bacia onde será implantada a rede de monitoramento;
- b) Definição dos métodos de tratamento dos dados derivados do monitoramento;
- c) Definição dos procedimentos de coleta, preservação e transporte de amostras;
- d) Definição dos critérios analíticos das variáveis selecionadas para o projeto de rede;
- e) Definição da infraestrutura de tecnologia da informação necessária ao armazenamento, proteção e recuperação dos dados
- f) Realização da análise custo-benefício do PMQA;
- g) Realização de eventuais ajustes no PMQA.

A informação armazenada pode incluir as informações que balizaram a concepção do projeto, pois isso pode melhorar a tomada de decisão quanto a eventual necessidade de ajustes e correções nos projetos iniciais das redes, reduzindo a possibilidade de erros futuros (Soares, 2001).

4.3 Projeto técnico de redes de monitoramento da qualidade da água

A etapa técnica do projeto de RMQA deve conter no mínimo: a configuração espacial e/ou o número de estações, a seleção das variáveis a monitorar e a frequência de amostragem (Strobl; Robillard, 2008; CCME, 2015; Soares, 2001).

A seleção dos locais a monitorar e, por extensão a dimensão das redes, é considerado o aspecto mais crítico dos projetos técnicos, visto que envolve frequentemente a alocação do maior volume dos recursos de implantação dos programas. O próprio custo de coletas e análises laboratoriais depende diretamente desse aspecto (Palmer; Asce; Mackenzie, 1985).

Esse fato também é refletido nos estudos científicos, pois analisando *vis-à-vis* cada aspecto do projeto básico das redes é evidente a predileção dos pesquisadores pela temática de seleção dos locais de monitoramento (Nguyen *et al.*, 2019; Jiang *et al.*, 2020a).

As primeiras contribuições, nesse contexto, visando dar uma caráter mais impessoal e menos subjetivo ao projeto das redes referem-se aos trabalhos de Shreve (1967) *apud* Strobl e Robillard (2008), Sharp (1971), Sanders e Adrian (1978) e Sanders *et al.* (1983) *apud* Vilmin *et al.* (2018). Tanto os trabalhos de Shreve (1967) quanto Sharp (1971) partiram do conceito de hierarquização de trechos de rios proposto por Horton (1945).

Horton (1945) propôs um critério de hierarquização das redes fluviais, em que trechos de rio sem tributários são considerados de primeira ordem. De maneira geral, trechos de ordem u podem ter tributários de ordem $u - 1$ até 1. Esse procedimento garante que o rio principal terá sempre a maior ordem, desde a foz até a nascente principal (Soares, 2001).

Sharp (1971) utilizou a abordagem de Horton para seleção dos locais a monitorar, considerando a localização das fontes de poluição (Soares, 2001). Esse método permite isolar as fontes de poluição partindo da sub-divisão sistemática da rede hidrográfica em porções relativamente semelhantes quanto ao número de tributários contribuintes (Strobl; Robillard, 2008).

O critério de Sanders *et al.* (1983), por sua vez é uma extensão da proposta de Sharp (1971). O autor adicionou ao método Sharp o conceito de fontes de poluição como tributários. Todos os métodos descritos estão baseados na topologia das redes hidrográficas, não demandando informações prévias das condições de qualidade da água constituindo, portanto, métodos aplicáveis a bacias não monitoradas (Soares, 2001; Jiang *et al.*, 2020a).

O estado atual de evolução dos métodos para determinação do melhor arranjo espacial ao monitoramento da qualidade da água, entretanto, tem sido dominado por outras abordagens com destaque para otimização mono e multiobjetivo, análise multivariada, entropia de Shannon, geoestatística e métodos multicriteriais. (Ghorbani *et al.*, 2020; Liyanage; Marasinghe; Yamada, 2016; Reina-García *et al.*, 2020; Sabzipour; Asghari; Sarang, 2019; Lee *et al.*, 2014; Alves *et al.*, 2018; de Almeida *et al.*, 2022; Singh *et al.*, 2019a; Shi *et al.*, 2018).

A seleção das variáveis e sua frequência, historicamente, não recebeu a mesma atenção e cuidado nos projetos das redes que a seleção dos locais para monitoramento. Dessa forma o projeto técnico desses aspectos sempre esteve envolto em decisões ortodoxas, convenientes ou permeadas de subjetividade (Dixon; Chiswell, 1996; Strobl *et al.*, 2006a; 2006b).

Apenas nas últimas décadas, muito em decorrência do novo paradigma “*rico em dados, pobre em informação*” é que essa realidade tem sido superada através da busca do melhor custo-benefício das redes pela eliminação das redundâncias, com avanços efetivos nos projetos (Ward; Loftis; McBride, 1986; Timmerman *et al.*, 2010).

Ao analisar 59 estudos sobre projetos de redes de monitoramento, Nguyen *et al.* (2019), evidenciaram que apenas 14 tratavam da seleção de variáveis de qualidade da água, com abosoluta predominância de abordagens quantitativas (13 estudos)

sobre as qualitativas (1 estudo). Essa baixa atratividade é corroborada por Jiang *et al.* (2020a), ao demonstrar que de um universo de 107 estudos compilados, em sua análise sobre a distribuição dos estudos voltados ao projetos de redes de monitoramento da qualidade da água, apenas 8 se ocupavam da seleção das variáveis de qualidade da água; e destes 7 só foram publicados após 2013.

Menor número de variáveis facilita a análise dos dados demandando menor tempo e esforço por parte da equipe técnica dos programas. Entretanto, a descontinuação do monitoramento de variáveis de qualidade da água deve ocorrer de maneira cuidadosa e responsável para que as informações úteis ao gerenciamento da qualidade da água sejam preservadas. (Strobl; Robillard, 2008).

Historicamente, os programas baseiam seu monitoramento em variáveis físicas, químicas e biológicas, como pH, turbidez, DBO, coliformes fecais e outras consagradas mundialmente (Nguyen *et al.*, 2019). Contudo, há uma tendência recente de aprimoramento dos programas via adição de novas variáveis que refletem um conceito de qualidade ampliado e sistêmico. Essas variáveis podem incluir micropoluentes, contaminantes emergentes, novos indicadores microbiológicos e ainda análises ecotoxicológicas (Strobl; Robillard, 2008).

A principal limitação da expansão do número de variáveis está no aumento potencial dos custos operacionais do monitoramento (Collins *et al.*, 2012; Graveline *et al.*, 2010). Todavia, caso esse processo seja acompanhado de análise de desempenho das redes, pode haver maior liquidez ao programa uma vez que, os custos da inclusão de novas variáveis podem ser compensados pela eliminação ou redução da frequência de monitoramento de variáveis redundantes e pouco informativas.

Análises estatísticas podem auxiliar a identificação de grupos de variáveis semelhantes. Dessa forma, ao reduzir esse conjunto a apenas algumas variáveis, garante-se que a informação útil seja obtida sem a necessidade de manutenção do conjunto completo. Esse tipo de análise inciou-se pela aplicação de métodos como os propostos por Sanders, em 1983 (Strobl; Robillard, 2008).

O estado da arte revela estudos para a seleção de variáveis em projetos técnicos de PMQA baseados em análise multivariada e correlação-regressão (Alves *et al.*, 2018;

Nguyen *et al.*, 2019; Calazans *et al.*, 2018; da Luz; Tobiasson; Kumpel, 2022). Atualmente vê-se também aumento de estudos baseados nos princípios da entropia de Shannon e inteligência artificial (redes neurais artificiais) (Keum *et al.*, 2017; Jiang *et al.*, 2020a; Mirauda; Ostoich, 2020; Villas-Boas; Olivera; de Azevedo, 2017).

A frequência de monitoramento da qualidade da água, a despeito de sua importância secundária atribuída pelas primeiras contribuições científicas aos projetos técnicos das redes, tem grande relevância no alcance dos objetivos e impacto sobre os custos operacionais das redes (Soares, 2001; Strobl; Robillard, 2008).

Apesar de Jiang *et al.* (2020a), em seu estudo de revisão sobre os métodos de projeto e redimensionamento de redes de monitoramento da qualidade da água, não revelar resultados que atestem representatividade desprezível para a frequência de monitoramento nos projetos técnicos, é incontestável o fato de que esse tema ainda é muito pouco explorado. Dos 107 trabalhos compilados pelos autores apenas 22 se concentraram na determinação da frequência ideal de monitoramento.

Para Strobl e Robillard (2008) nos primórdios dos PMQA, não era incomum que a frequência fosse definida em função da capacidade técnica e física dos laboratórios responsáveis pelas análises, e não pelos objetivos do monitoramento, gerando assim distorções, as quais os autores denominaram de amostragem infrequente.

Frequências grosseiras podem não capturar as variações de qualidade da água ou os efeitos de fontes específicas de poluição que, por outro lado, podem representar os principais objetivos de programas de monitoramento investigativo. Esse é o caso, por exemplo, de amostragens realizadas a nível mensal, em bacias em que se busca avaliar o efeito da lixiviação de poluentes de solos mantidos sob uso de agricultura intensiva. Para medir adequadamente esse efeito do uso da terra sobre a qualidade da água seria oportuno o aumento da frequência de monitoramento na transição da estação seca para a chuvosa, uma vez que as piores condições são esperadas durante as primeiras chuvas.

De fato, quando a frequência de monitoramento está em nível menor que o ideal, as séries temporais podem não capturar o comportamento estatístico da qualidade da água. Apesar disso, eventuais ajustes necessitam considerar a relação custo-

benefício, uma vez que por vezes o aumento no número de amostragens pode produzir um impacto muito pequeno no erro padrão da média das variáveis analisadas, dado que ele é inversamente proporcional à raiz quadrada do número de amostras (Strobl; Robillard, 2008).

Uma das primeiras contribuições acerca desse aspecto dos projetos técnicos das redes foi dado por Sanders e Adrian, em 1978. Os autores apresentaram um método estatístico para definição da melhor frequência de amostragem, baseado na construção de intervalos de confiança (Jiang *et al.*, 2020a). Essa abordagem, embora tenha sido proposta para redes hidrométricas possui um nível de generalização que permite sua aplicabilidade para programas e redes de qualidade da água.

Em seu estado da arte, a frequência de amostragem, quando comparada aos demais critérios de projeto técnico das redes, possui menor diversidade de técnicas, sendo amplamente dominado por métodos estatísticos, embora a Entropia de Shannon apresente potencial para boas aplicações, visto as experiências recentes nesse tema (Peña-Guzmán; Soto; Diaz, 2019; Khalil *et al.*, 2014; Liu *et al.*, 2014; Karamouz *et al.*, 2009b).

4.4 Classificação de PMQA

Os programas de monitoramento da qualidade da água podem ser diferenciados entre si pelas características de operação das redes e seus objetivos. A pluralidade de critérios que norteiam o planejamento das redes também é refletido no número de critérios de classificação, de forma que não há uma única convenção adotada internacionalmente.

Apesar disso, os critérios propostos pelo Serviço Geológico dos Estados Unidos (United States Geological Survey - USGS), pelo Conselho Canadense de Ministros do Meio Ambiente (Canadian Council of Ministers of the Environment - CCME, Framework for Water Quality Monitoring) e pela União Europeia (European Union, Water Framework Directive -WFD) têm grande relevância no cenário mundial pois não estão atrelados a aspectos específicos de um país ou uma região.

O critério do USGS divide o monitoramento em três grandes grupos: de *status*, tendência e conformidade (USGS, 1995). Os programas de *status* devem fornecer

informação atual sobre as condições de qualidade da água e sanidade dos ecossistemas aquáticos, enquanto os programas de tendência devem revelar o comportamento da variabilidade espaço-temporal da qualidade da água.

Já os programas de conformidade têm forte caráter regulatório visando principalmente avaliar se as condições de qualidade da água estão em acordo com as normas e padrões estabelecidos, o controle da qualidade da água a longo prazo, avaliar a capacidade de atendimento dos mananciais ao uso para diluição de efluentes e os efeitos decorrentes de projetos na bacia durante e após sua implementação (USGS, 1995).

A WFD propõe em seu critério, tal como a USGS, três grupos de programas de monitoramento incluindo além dos objetivos do monitoramento, o horizonte de implementação dos procedimentos de monitoramento. São eles: programas de vigilância, programas operacionais e programas investigativos (European Union, 2003).

Os programas de vigilância, de acordo com a WFD (European Union, 2003), são projetados para operação em longo prazo e para cobrir os objetivos de: complementar e validar os procedimentos de avaliação de impactos ambientais, balizar de forma mais eficiente e efetiva outros futuros programas de monitoramento e detectar mudanças de longo prazo nas condições naturais dos mananciais e decorrentes de atividades antrópicas gerais.

Programas de monitoramento operacional, por sua vez, são orientados para: demonstrar o *status* de corpos hídricos na iminência de descumprir seus objetivos ambientais. Essa categoria de programas tem forte apelo voltado à confirmação da condição de risco de mananciais previamente identificada. Esse tipo de monitoramento de curto prazo é focado especialmente nas variáveis de qualidade da água indicativas dos efeitos da poluição que esteja exercendo pressão nos mananciais (European Union, 2003).

A WFD (European Union, 2003) propõe ainda uma categoria de programas chamados investigativos, que podem ser implementados durante um curto período de tempo justificado em função de certas condições excepcionais como: quando não há

atendimentos aos padrões ambientais dos mananciais por razões desconhecidas e o monitoramento operacional ainda não foi implementado, embora possa haver monitoramento de vigilância, e para determinar a magnitude e os impactos de emergências ambientais decorrentes de desastres.

Os critérios do CCME resumizam os critérios da União Europeia e do USGS condensando as tipologias dos programas de monitoramento em apenas duas categorias: i) monitoramento de *status* e tendência de longo prazo; ii) pesquisa de curto prazo (investigativo) ou monitoramento de conformidade (CCME, 2015). A primeira tipologia não apresenta nenhuma particularidade, de forma que são definidas da mesma maneira que os programas de *status* e tendência propostos pelo USGS, representando apenas uma junção de ambos (CCME, 2006).

De forma semelhante, vale para a segunda tipologia os mesmos conceitos de programas de monitoramento investigativo e de conformidade propostos pela WFD e pelo USGS de forma condensada (CCME, 2006). Os programas dessa tipologia são focados na investigação de fontes de poluição associadas a comportamentos de não-conformidade dos padrões esperados de qualidade da água em um manancial e na avaliação dos efeitos de poluição hídrica de fonte difusa ou pontual, sendo caracterizados como programas de monitoramento de curto prazo.

Em relação ao monitoramento emergencial, nenhum dos critérios mencionados apresenta redes com esse caráter como uma classe particular. Apesar disso, as principais características dos programas de monitoramento investigativo; propostos pelo CCME, USGS e União Europeia os qualificam como a alternativa mais apropriada para classificação das redes voltadas ao contexto crítico de poluição hídrica decorrente de desastres ambientais.

O ponto de contradição, nesse contexto, fica por conta do horizonte temporal de monitoramento, pois embora os impactos socioambientais e a recuperação da qualidade da água e dos ecossistemas aquáticos de bacias impactadas por desastres ambientais de grande magnitude, não se esgotem em curto prazo, de fato, as redes investigativas são redes temporárias caracterizadas pela operação em curtos períodos de tempo.

Dessa forma, em função da ausência patente de uma definição que contemple integralmente as características das redes implantadas para monitoramento de longo prazo em cenários de desastres ambientais com impactos de grande extensão espaço-temporal, se propõe nesse trabalho uma nova definição. Redes emergenciais, portanto, são compreendidas como redes de monitoramento da qualidade da água e sedimentos, cuja motivação para seu planejamento e estruturação decorre dos efeitos diretos e indiretos de desastres ambientais de grande proporção sobre recursos hídricos.

Essas redes caracterizam-se por monitorar, além das variáveis típicas de programas de monitoramento para fins de regulação e controle, variáveis diretamente relacionadas aos poluentes vinculados aos desastres, sendo oportuno incluir complementarmente análises ecotoxicológicas. Elas devem minimamente possuir os seguintes objetivos:

- a) acompanhamento de tendências de longo prazo no comportamento da qualidade da água, considerando os poluentes específicos derivados do desastre;
- b) análise da dinâmica de retorno dos recursos hídricos às condições pré-desastre;
- c) análise da variabilidade sazonal da qualidade da água, especialmente, com relação aos poluentes específicos relacionados ao desastre;
- d) suporte para estabelecimento de sistemas de alerta de risco à qualidade da água, especialmente se os mananciais impactados forem utilizados para abastecimento de núcleos humanos e/ou os poluentes tiverem características de persistência. subsidiar análise do desempenho das ações de recuperação da qualidade da água e dos ecossistemas aquáticos impactados pelo desastre.

4.5 Principais métodos utilizados em projetos técnicos de PMQA

Os métodos mais utilizados nos projetos técnicos e redimensionamento das redes são: otimização mono e multiobjetivo, análise multivariada, entropia de Shannon, geoestatística, estatística e métodos multicriteriais.

A análise multivariada tem sido uma alternativa comum na seleção de variáveis de qualidade da água e dos locais de monitoramento, sobretudo ao utilizar análise de *cluster* (AC), análise de componentes principais (ACP), análise fatorial (AF) e análise discriminante (AD) (Alves *et al.*, 2018; Peña-Guzmán; Soto; Diaz, 2019; Calazans *et al.*, 2018). A adequabilidade desses métodos estatísticos na identificação de padrões e redução de dimensionalidade é bem conhecida em casos em que há necessidade de tratamento de base de dados densas e com grande proporção de variáveis (Maity, 2018; Hair Jr *et al.*, 2006).

A Entropia de Shannon tem sido também utilizada em estudos sobre aspectos de redes de monitoramento nas suas diversas tipologias, especialmente na seleção de estações e de variáveis para o monitoramento (Jiang *et al.*, 2020a; Nguyen *et al.*, 2019; Keum *et al.*, 2017).

A teoria da informação introduziu a entropia de Shannon como uma medida da incerteza e do conteúdo de informação de processos ou variáveis aleatórias (Shannon, 1948). Dessa forma, métodos que utilizam a entropia de Shannon são promissores na investigação de redes hidrológicas e, de fato, a literatura especializada demonstra a sua aplicabilidade (Keum *et al.*, 2017; Singh *et al.*, 2019a; Karamouz *et al.*, 2009b; Stosic; Stosic; Singh, 2017; Barbaros, 2022).

Apesar dos bons resultados e aplicabilidade, a entropia de Shannon também apresenta limitações. A primeira delas refere-se ao fato de que frequentemente há necessidade de discretização dos dados originais quando as variáveis são contínuas e, a depender do método usado a esse fim, há risco tanto de superestimar quanto subestimar o conteúdo de informação das variáveis (Alfonso *et al.*, 2010; Sreeparvathy; Srinivas, 2020). O segundo aspecto é que em essência, a entropia de Shannon pode ser considerada um método estatístico livre de distribuição sendo bastante influenciado por valores extremos, necessitando sempre atenção especial do analista quanto à presença de *outliers* nos dados (Nooghabi; Nooghabi, 2016).

Quanto à geoestatística, tal como a entropia de Shannon, ela possui boa representatividade e tem sido uma opção frequente para a seleção dos locais de monitoramento (Yenilmez; Düzgün; Aksoy, 2015; Sabzipour; Asghari; Sarang, 2019; Ali; Othman, 2018). O interesse na utilização de métodos geoestatísticos, tal como

aqueles apoiados em otimização, tem forte relação com o avanço da tecnologia da informação.

Esses métodos em geral são baseados em operações espaciais como interpolação (Krigagem, Spline, etc...) e análise de correlação espacial (Semivariograma), dentre outros, cuja implementação se dá através do uso de Sistemas de Informação Geográfica (SIG), exigindo hardware e software compatíveis com o volume de dados processados para sua implementação (Oliver; Webster, 2014).

O atual cenário da tecnologia da informação favorece o desenvolvimento dos métodos geoestatísticos. Apesar da implementação desses métodos não demandar grande volume de dados, a consistência de seus resultados depende diretamente da distribuição espacial dos pontos amostrais (Jiang *et al.*, 2020a), nesse caso, representados pela configuração espacial das redes.

Métodos de otimização tem atraído o interesse de parcela significativa dos pesquisadores durante as últimas décadas, especialmente no redimensionamento de redes de monitoramento da qualidade da água (Jiang *et al.*, 2020a; Ning; Chang, 2005; Kao *et al.*, 2008; Zhu *et al.*, 2018; Salman *et al.*, 2021).

Esses métodos possuem considerável versatilidade, pois são aplicáveis a problemas de otimização que envolvem tanto variáveis contínuas quanto discretas (Jalili; Hosseinzadeh, 2018; Kaveh; Ilchi Ghazaan, 2014). No caso monoobjetivo é grande a probabilidade de se encontrar uma solução ótima global para os problemas de otimização, contudo, no caso multiobjetivo os métodos logram êxito quanto retornam um conjunto de soluções não-dominadas chamadas Pareto-ótimo (Brown; Dempster; Adby, 1974).

Nesses casos há necessidade de utilização de critérios adicionais de tomada de decisão, como os métodos *Analytic Hierarchy Process* (AHP) ou Promethee, para selecionar no conjunto Pareto-ótimo a solução mais adequada ao problema investigado (Ghorbani *et al.*, 2020).

Algoritmos baseados em inteligência artificial exercem forte dominância nos estudos voltados à otimização de redes da qualidade da água, tanto nas abordagens

multiobjetivo quanto monoobjetivo (Liyanage; Marasinghe; Yamada, 2016; Ghorbani *et al.*, 2020; Puri; Karthikeyan, 2017; Ouyang *et al.*, 2008; Mahjouri; Kerachian, 2011; Karamouz *et al.*, 2009a; Lee *et al.*, 2014; Telci *et al.*, 2009; Zhu *et al.*, 2018; Zhu *et al.*, 2019; Park *et al.*, 2006; Icaga, 2005; Pérez *et al.*, 2017; Salman *et al.*, 2021; Vega-Rodríguez *et al.*, 2021).

A evolução da arquitetura de hardware e software dos computadores convencionais, da computação paralela e surgimento do processamento em nuvem ocorrido nos últimos anos possibilitou a superação das limitações do poder de processamento demandado por sofisticados mecanismos de busca de soluções aos problemas de otimização (Dixon; Chiswell, 1996; Mladen, 1997; Fagerholm *et al.*, 2002).

O exame dos critérios que compõe frequentemente os problemas de otimização, ilustrados pelas funções objetivo, também corrobora as projeções promissoras do crescimento de aplicações de abordagens multicriteriais e baseadas na entropia de Shannon, especialmente em relação à configuração espacial das redes. Essas abordagens normalmente visam maximizar a informação e minimizar a redundância das redes (Liyanage; Marasinghe; Yamada, 2016; Ghorbani *et al.*, 2020; Puri; Karthikeyan, 2017; Mahjouri; Kerachian, 2011; Lee *et al.*, 2014; Park *et al.*, 2006; Icaga, 2005; Salman *et al.*, 2021).

O gerenciamento ambiental e, nesse caso o gerenciamento de recursos hídricos, está passando por um processo de mudança visando a adoção de abordagens integradoras e sistêmicas que permitem uma análise mais ampla dos processos ambientais, com efeitos também sobre os projetos e redimensionamento das redes (CCME, 2015; Collins *et al.*, 2012).

Com efeito, novos métodos têm proposto critérios a partir da integração de múltiplas abordagens como geoestatística, entropia de Shannon, variáveis socioeconômicas, indicadores de poluição hídrica e relações topológicas nas bacias, de forma a alcançar uma perspectiva mais ampla, não somente restrita à qualidade da água (Chang; Lin, 2014; Do *et al.*, 2011; Asadi *et al.*, 2021).

Dois fatos marcam, portanto, o estágio atual dos projetos e redimensionamento das redes de monitoramento da qualidade da água. Primeiro, a tendência cada vez mais

expressiva no uso de otimização; e segundo a evidente sofisticação na concepção dos problemas de otimização, personificados por múltiplas funções objetivos com elevado grau de complexidade.

Dada a importância recente atribuída aos métodos para projeto técnico das redes de monitoramento baseados em otimização e na Entropia de Shannon, especialmente quanto à seleção dos locais e das variáveis a monitorar, é conveniente descrever mais detalhadamente a teoria e os conceitos que os caracteriza.

4.5.1 Otimização mono e multiobjetivo

Embora todos os aspectos de projeto das redes de monitoramento possam ser tratados como problemas de otimização, tem sido mais frequente o uso dessa estratégia na seleção dos locais a monitorar, incluindo o redimensionamento das redes (Ning; Chang, 2005; Kao; Li; Hu, 2012; Puri *et al.*, 2017; Salman *et al.*, 2021).

O conceito de otimização, à luz das engenharias, compreende a determinação da melhor alternativa de projeto considerando para isso critérios pré-estabelecidos, restrições, e os recursos disponíveis. Nesse processo algumas etapas são comuns a qualquer aplicação (Brasil; da Silva, 2019).

Primeiramente deve-se identificar as variáveis envolvidas. Em seguida definir uma ou várias funções objetivo (otimização de objetivo único ou multiobjetivo), as quais propõe-se maximizar ou minimizar (dependendo das características das variáveis), sob determinadas restrições. Estas inclusive podem ser de natureza física. Por fim, cabe ao analista definir o método de otimização mais adequado ao problema de otimização (Brasil; da Silva, 2019).

Diferentemente dos tradicionais problemas de otimização monoobjetivo, um problema de otimização multiobjetivo (POMO) se propõe a otimizar um grupo de funções usualmente conflitantes entre si. Como consequência, não existe uma única solução para o problema, mas um conjunto de soluções que representam diferentes alternativas. Obviamente, por tratar-se de um problema multiobjetivo cada solução quando maximiza um objetivo, inevitavelmente tende a não ter o mesmo desempenho nos demais (Yu; Zhang, 2017; Wang; Rangaiah, 2017; Afsar; Miettinen; Ruiz, 2021; Allmendinger *et al.*, 2017; Li; Yao, 2019).

O gerenciamento de recursos hídricos requer em muitas situações a tomada de decisões baseada em múltiplas variáveis, que por vezes são de difícil harmonização. Devido aos altos custos de monitoramento e restrições orçamentárias, o número de estações em um sistema é geralmente limitado. Uma rede de monitoramento adequada deve fornecer o máximo de informações com número mínimo de estações de monitoramento (Yazdi, 2018).

O objetivo central da otimização de redes de monitoramento é a determinação de um conjunto de estações eficientes, de tal forma a evitar manter na rede estações com considerável redundância através da maximização e/ou minimização de métricas adequadas (Keum; Coulibaly, 2017a).

Problemas de otimização baseados em medidas do conteúdo de informação têm sido bastante explorados em redes de monitoramento da qualidade da água. A depender da abordagem escolhida pode-se tanto tratar o problema por uma abordagem monoobjetivo quanto multiobjetivo (Foroozand; Weijs, 2021).

Particularmente, a entropia de Shannon tem sido bastante explorada na concepção de funções monoobjetivo ou multiobjetivo (Foroozand; Weijs, 2021). Os algoritmos de otimização buscam maximizar e/ou minimizar a informação ou redundância da informação contida em conjuntos de estações candidatas à seleção na configuração ótima proposta para as redes (Máxima informação/Mínima redundância-MIMR) (Keum *et al.*, 2017).

Dentre as formas de entropia utilizadas destaca-se a entropia conjunta, transinformação (informação mútua), correlação total, valor da informação e entropia condicional (Salman *et al.*, 2021; Ursulak; Coulibaly, 2021; Stosic; Stosic; Singh, 2017; Pourshahabi, 2018; de Pádua *et al.*, 2019; Alfonso *et al.*, 2013; Keum; Coulibaly, 2017a).

A combinação de duas ou três formas de entropia compondo as funções mono ou multiobjetivo tem prevalecido nessas aplicações (Singh, 2013; Jiang *et al.*, 2020a; Alameddine *et al.*, 2013; Karamouz *et al.*, 2009a), caso do critério de máxima informação/mínima redundância (MIMR):

$$\max: H(X_{S1}, X_{S2}, \dots X_{Sk}) + \sum_{i=1}^m T(X_{S1;Sk}; X_{Fi}) \quad (1)$$

$$\min: C(X_{S1}, X_{S2}, \dots X_{Sk}) \quad (2)$$

Onde,

$X_{S1;Sk}$: denota as séries temporais conjuntas de $X_{S1}, X_{S2}, \dots X_{Sk}$

$H(X_{S1}, X_{S2}, \dots X_{Sk})$: entropia conjunta

$T(X_{S1;Sk}; X_{Fi})$: transinformação

$C(X_{S1}, X_{S2}, \dots X_{Sk})$: correlação total

4.5.2 Entropia de Shannon

A publicação do trabalho intitulado “*A Mathematical Theory of Communication*”, rendeu a Claude Elwood Shannon o título de criador da área que se convencionou chamar “Teoria da informação”. Diferentemente da conotação intuitiva da palavra “informação”, a definição de Shannon para este conceito desconsidera os aspectos semânticos. De forma mais clara pode-se pensar “informação” como o quanto é aprendido, ao invés de o que é aprendido (Dharmapalani *et al.*, 2018; Shannon, 1948).

Uma das principais inovações propostas por Shannon remete ao conceito de entropia informacional, capaz de medir de forma concreta e livre de quaisquer subjetividades o conteúdo real de informação associado a uma mensagem. A teoria da informação apesar de ter sido pensada por Shannon para limitar-se a aplicações voltadas à engenharia e comunicação, expandiu-se para outras áreas e tem amplo potencial de aplicação (Aftab *et al.*, 2011).

A entropia de Shannon está vinculada a probabilidades associadas a ocorrência de possíveis estados ou eventos. Se iterativamente qualquer novo dado a ser fornecido não apresenta incerteza alguma, então esse dado não agrega informação adicional em relação aos dados pré-existentes. Por outro lado, se há grande incerteza, então há grande possibilidade do novo dado trazer ganho quanto ao conteúdo de informação (Keum; Coulibaly, 2017b).

A entropia de Shannon é descrita como:

$$H(X) = -\sum_{j=1}^N p(x_j) \log[p(x_j)] \quad (3)$$

Onde,

j : denota um intervalo de dados discretos

x_j : denota o resultado correspondente ao intervalo j

$p(x_j)$: a probabilidade de ocorrência do resultado x_j

A entropia marginal $H(X)$ é a medida do conteúdo médio de informação de uma variável aleatória X com distribuição de probabilidade $p(x)$. Sua utilização possui viabilidade em qualquer tipo de conjunto de dados, como anual, mensal, sazonal, dentre outros (Singh *et al.*, 2019a).

Dado que a entropia é definida em termos de escala logarítmica, a opção do número utilizado como base define por consequência a unidade em que o conteúdo de informação é calculado. Se a base 2 é usada as unidades resultantes podem ser chamadas de dígitos binários. Um dispositivo com duas posições estáveis pode armazenar um bit de informação (Shannon, 1948). Outras opções são o número de Euler (base natural) e base 10, casos em que as unidades são chamadas *nats* e *hartley*, respectivamente (Foroozand; Weijs, 2021).

Entropia tem sido aplicada para avaliar redes de coleta de dados que atendem a uma variedade de propósitos, incluindo medição de chuva, monitoramento da qualidade da água e monitoramento de vazão. Nesse âmbito, algumas formas de entropia informacional, mais frequentemente utilizadas são: i) entropia marginal; (ii) entropia conjunta, (iii) transinformação e (iv) correlação total (Foroozand; Weijs, 2021).

A entropia conjunta representa a informação total não redundante, ou compartilhada, de múltiplas variáveis aleatórias (Li; Singh; Mishra, 2012). A título de exemplo, tomando o caso bivariado, ela é expressa por:

$$H(X_1; X_2) = -\sum_{x_1} \sum_{x_2} p(x_1; x_2) \log p(x_1, x_2) \quad (4)$$

Onde:

$p(x_1, x_2)$: distribuição conjunta das variáveis X_1 e X_2

$H(X_1; X_2)$: entropia conjunta das variáveis X_1 e X_2

Quantitativamente, a entropia conjunta é menor ou igual a soma de suas entropias marginais unidimensionais. Caso as variáveis aleatórias sejam completamente independentes a entropia conjunta deverá corresponder a soma das entropias marginais de ambas (Li; Singh; Mishra, 2012).

Em termos práticos a entropia conjunta impõe certa dificuldade à sua determinação, pois caso as variáveis aleatórias não sejam completamente independentes a soma de suas entropias marginais não é suficiente para determiná-la. Uma alternativa para esse fim, apresentada por (Singh, 2013) é gerar variáveis conjuntas através da agregação das variáveis originais. Para tanto deve-se proceder as seguinte etapas (Tabelas 1 a 4):

- 1) Agregar cada registro de ambas as variáveis originais em um novo registro correspondente à nova variável, como demonstrado a seguir.

Tabela 1 – Exemplo de variáveis discretas originais

Variável A	Variável B
1	5
3	9
2	7
1	5

Tabela 2 – Exemplo da etapa 1 da geração de variáveis conjuntas

Variável agregada
15
39
27
15

- 2) Definir uma nova classificação baseada na ordenação da variável agregada,

abaixo.

Tabela 3 – Exemplo da etapa 2 de geração de variáveis conjuntas

Variável agregada	Nova classificação
15	1
27	2
39	3

- 3) Reclassificar as variáveis agregadas correspondentes à nova variável conjunta, abaixo.

Tabela 4 - Exemplo da etapa 3 de geração de variáveis conjuntas

Variável conjunta
1
3
2
1

A entropia marginal de variáveis conjuntas geradas dessa maneira corresponde à entropia conjunta das variáveis aleatórias discretas originais. No caso de múltiplas variáveis deve-se fazer a agregação aos pares, sendo que a partir da primeira repetição dos procedimentos haverá sempre a agregação de uma variável conjunta e uma variável original. Ao final da última agregação, portanto, obter-se-á a variável conjunta correspondente à agregação de toda as variáveis originais.

A transinformação corresponde à informação compartilhada por um par de variáveis, também chamada de informação mútua. Ela é definida matematicamente pela expressão (Stosic; Stosic; Singh, 2017):

$$T(X_1; X_2) = \sum_{x_1} \sum_{x_2} p(x_1; x_2) \log \frac{p(x_1; x_2)}{p(x_1)p(x_2)} \quad (5)$$

Onde,

$T(X_1; X_2)$: Informação compartilhada pelas variáveis X_1 e X_2

$p(x_1)$: Probabilidade de ocorrência de x_1

$p(x_2)$: Probabilidade de ocorrência de x_2

$p(x_1; x_2)$: Probabilidade conjunta de x_1 e x_2

A transinformação também compreende uma medida do grau de dependência entre variáveis aleatórias, capaz de medir tanto a dependência linear quanto não-linear (Li; Singh; Mishra, 2012). Em aplicações práticas pode-se buscar a transinformação existente entre conjuntos específicos de variáveis aleatórias e uma variável simples, assim como pode-se determinar a transinformação entre diferentes conjuntos de variáveis aleatórias.

Em ambos os casos, pode-se contornar esse problema gerando variáveis conjuntas para grupos de variáveis aleatórias originais a partir dos mesmos procedimentos de agregação descritos para aplicações na determinação da entropia conjunta e, em seguida, calcula-se diretamente a transinformação entre os pares de variáveis conjuntas (Singh, 2013).

A correlação total corresponde a uma medida da informação redundante considerando um conjunto de distintas variáveis aleatórias. Em termos práticos ao se determinar a diferença entre a soma das entropias marginais e a entropia conjunta de diferentes variáveis aleatória chega-se à correlação total (Keum *et al.*, 2017):

$$C(X_1; X_2, \dots; X_N) = \sum_{i=1}^N H(X_i) - H(X_1; X_2, \dots; X_N) \quad (6)$$

Onde,

$C(X_1; X_2, \dots; X_N)$: Correlação total entre N variáveis

$H(X_i)$: Entropia marginal da variável X_i

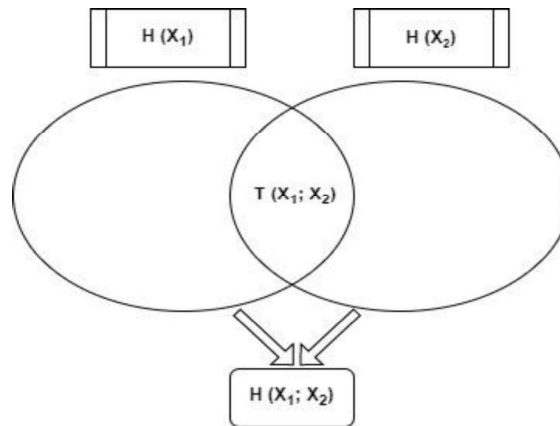
$H(X_1; X_2, \dots; X_N)$: Entropia conjunta das N variáveis

A determinação da correlação total entre um par de variáveis é equivalente à transinformação. A soma das entropias marginais de todas as variáveis nunca é inferior à sua entropia conjunta, então a correlação total não pode assumir valores

negativos. Se não houver dependência entre elas, a correlação total será zero (Keum *et al.*, 2017).

A relação entre as formas de entropia mencionadas para o caso bivariado é ilustrada pela Figura 1.

Figura 1 - Relação entre diferentes formas de entropia*



*** $H(X_1)$ e $H(X_2)$ são as entropias marginais das variáveis X_1 e X_2 . $H(X_1; X_2)$ é a entropia conjunta de X_1 e X_2 e $T(X_1; X_2)$ representa a transinformação das variáveis X_1 e X_2**

Os tipos de entropia descritos são as medidas básicas que têm sido tipicamente usadas no projeto de redes de monitoramento de água (Alfonso *et al.*, 2013; Ursulak; Coulibaly, 2021; de Pádua *et al.*, 2019). Enquanto muitos estudos desenvolveram abordagens específicas para estudos de caso usando os termos básicos de entropia, alguns estenderam os termos derivando ou combinando as medidas básicas (Keum *et al.*, 2017). Este é o caso do conceito de redundância à luz da Teoria da Informação.

Shannon (1948) propôs o conceito de redundância como a relação:

$$R = 1 - \frac{H(X)}{H_m(X)} = 1 - H_r(X) \quad (7)$$

Onde,

R : Redundância

$H(X)$: Entropia marginal da variável X

$H_m(X)$: Entropia máxima da variável X

$H_r(X)$: Entropia relativa da variável X

De forma a tornar mais clara a relação da redundância com os processos hidrológicos, Singh (2013) propôs um exemplo tomando E hipoteticamente como a resposta de uma bacia hidrográfica a fatores como temperatura, radiação, velocidade do vento e assim por diante. Como alguns desses fatores não são totalmente independentes, $H(X)$ tende a atingir um limite à medida que o número de fatores tende ao infinito, ou seja, à medida que novos fatores não transmite nenhuma informação nova, então a razão de $H(X)/H_m(X)$ expressa a restrição do ambiente e a subsequente perda de incerteza ou liberdade de escolha.

Um dos desdobramentos do conceito de entropia informacional introduzido por Shannon refere-se à capacidade de se estimar a condição relativa à máxima informação capaz de ser transmitida por uma variável aleatória. O princípio da máxima entropia (POME) formulado por Jaynes (1957a; 1957b), refere-se ao uso de uma distribuição de probabilidade apropriada que maximiza o valor da Entropia, consideradas as informações disponíveis sobre a variável aleatória em análise na forma de restrições (Singh, 2013; Yamaka; Sriboonchitta, 2020).

Se nenhuma informação além das restrições dadas estiver disponível, a distribuição de probabilidade deve ser selecionada de modo que seja menos tendenciosa em relação ao que não é conhecido. Assim, um dos pontos-chave é definir restrições, expressas em termos de momentos da distribuição (Singh, 2010; Wei; Song, 2019).

POME mostrou-se adequada a inúmeras aplicações voltadas à análise de dados e modelagem de variáveis aleatórias, incluindo-se as de natureza hidrológica. Isto decorre do fato deste princípio tomar inicialmente as incertezas como aspectos limitantes na modelagem e permitir ajustes à *posteriori* na medida em que novas informações são incorporadas decorrentes do aumento dos dados do monitoramento (redução da incerteza) (Perdigão *et al.*, 2020).

Apesar deste aspecto, o POME não descarta por completo os princípios físicos na estimativa dos limiares máximos ou superiores da entropia de Shannon. Ao contrário,

ele deve incorporar o conhecimento físico do comportamento das variáveis, na forma de restrições, para tornar sua utilização condizente com a realidade. Com efeito, reduz-se a possibilidade deste conceito ser utilizado indiscriminadamente gerando soluções que não passam de meros exercícios teóricos sem possuir valor prático (Perdigão *et al.*, 2020).

Gomes (2013) descreveu sucintamente as condições que as distribuições candidatas pelo POME devem satisfazer e os dois casos de restrições mais comuns:

Condições:

$$\blacksquare f_x(x) \geq 0, \text{ com igualdade válida apenas fora do suporte } S \text{ da variável} \quad (8)$$

$$\blacksquare \int f_x(x) dx = 1; \quad (9)$$

$$\blacksquare \int f_x(x) dx \int f_i(x) dx = k_i, \text{ para } 1 \leq i \leq k; \quad (10)$$

Onde,

k_i : momento centrado de i-ésima ordem

$f_x(x)$: função densidade de probabilidade

$f_i(x)$: função que faz $f_x(x)$ respeitar cada restrição

As condições podem ser resolvidas com uso dos multiplicadores de Lagrange. Dessa forma, o Lagrangiano é escrito como:

$$\begin{aligned} J(f_x(x)) &= - \int f_x(x) \ln[f_x(x)] dx + \beta_0 \left(\int f_x(x) dx \right) \\ &+ \sum_{i=1}^k \beta_i \left(\int f_x(x) dx \int f_i(x) dx - k_i \right) \end{aligned} \quad (11)$$

Onde,

$\beta_0, \dots, \beta_i$: multiplicadores de Lagrange

Derivando a equação 11 em ordem à distribuição $f_x(x)$, obtém-se:

$$\frac{\partial J(f_x(x))}{\partial f_x(x)} = -\ln[f_x(x)] - 1 + \beta_0 + \sum_{i=1}^k \beta_i f_x(x) \quad (12)$$

$\beta_0, \dots, \beta_i$ são escolhidos de forma que $f_x(x)$ satisfaça as restrições.

Caso 1: Suporte fixo ($S = [a, b]$)

Neste caso não há restrições relativas aos momentos, de forma que os multiplicadores de ordem 1 a k tomam todos valor 0. Assim, igualando a equação 12 a 0, obtém-se:

$$\int f_x(x) dx = 1 \Rightarrow \int_a^b e^{\beta_0-1} dx = 1 \Rightarrow e^{\beta_0-1}(b-a) = 1 \Rightarrow f_x(x) = \frac{1}{(b-a)} \quad (13)$$

Logo, sob a restrição de suporte fixo, a distribuição com máxima entropia é a uniforme.

Caso 2: Valor médio e variância fixos

Neste caso, há restrição quanto aos momentos, $\beta_0, \beta_1, \beta_2 \neq 0$. Da mesma forma que no caso anterior se a equação 12 for igualada a 0:

$$f_x(x) = \exp[\beta_0 + \beta_1 x + \beta_2 x^2 - 1] \quad (14)$$

Os valores dos escalares podem ser determinados através da resolução do sistema:

$$\int_{-\infty}^{+\infty} \exp[\beta_0 + \beta_1 x + \beta_2 x^2 - 1] dx = 1 \quad (15)$$

$$\int_{-\infty}^{+\infty} x \exp[\beta_0 + \beta_1 x + \beta_2 x^2 - 1] dx = k_1 \quad (16)$$

$$\int_{-\infty}^{+\infty} x^2 \exp[\beta_0 + \beta_1 x + \beta_2 x^2 - 1] dx = \sigma^2 \quad (17)$$

A partir da resolução da segunda equação do sistema, obtém-se os valores para os multiplicadores de Lagrange.

$$\beta_0 = -\ln[\sqrt[2]{2\pi\sigma}]; \beta_1 = k_i; \beta_2 = -\frac{1}{2\sigma^2} \quad (18)$$

Substituindo os valores na equação 14:

$$f_x(x) = \frac{1}{\sqrt[2]{2\pi\sigma}} \exp\left[-\frac{(x-k_1)^2}{2\sigma^2}\right] \quad (19)$$

Portanto, sob a restrição de valor médio e variância fixos, a distribuição de entropia máxima é a normal. Para os casos de variáveis discretas e ausência de restrições, a distribuição uniforme é a alternativa mais recomendável para a determinação da máxima entropia (Muñoz-Cobo *et al.*, 2017).

5 Área de estudo

5.1 Caracterização geral

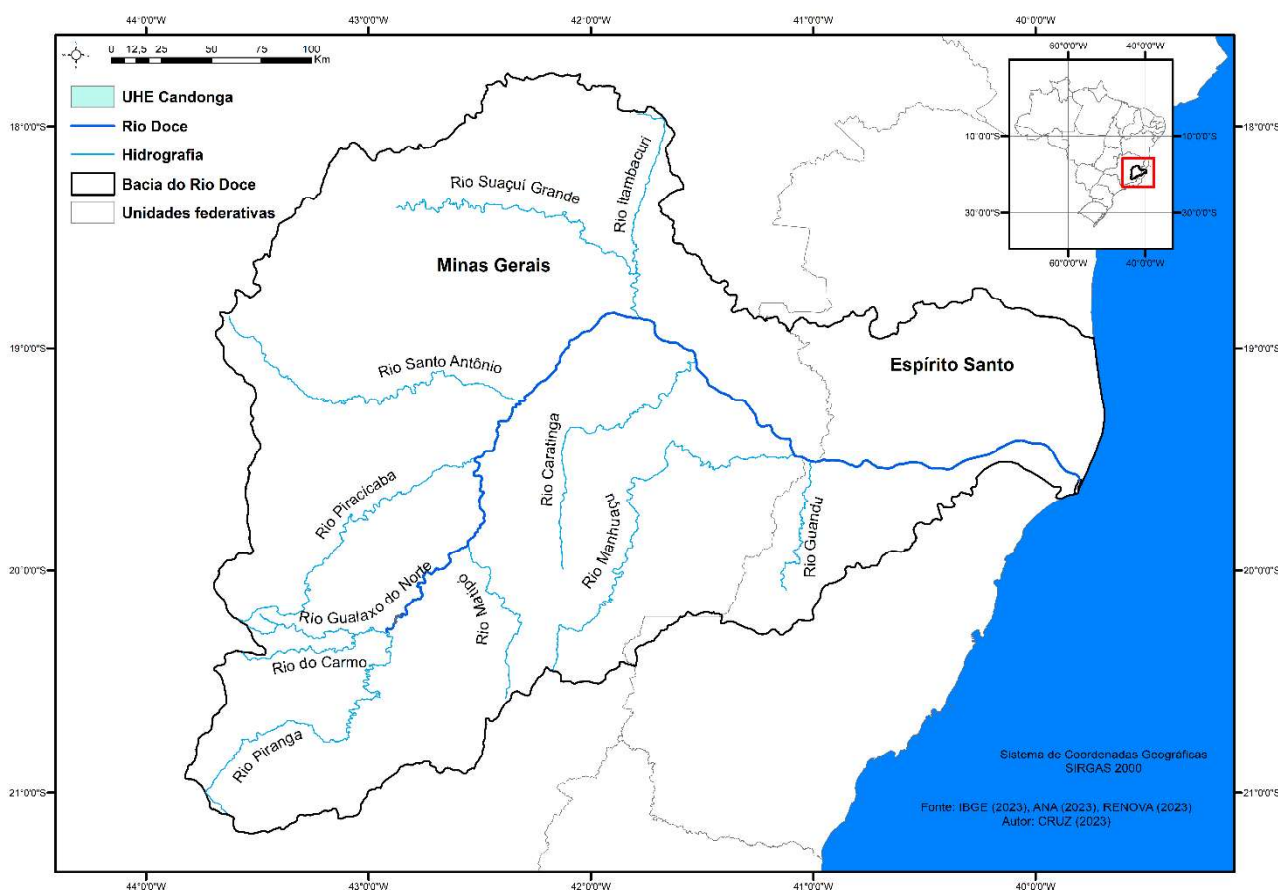
A área de estudo é compreendida pela bacia hidrográfica do Rio Doce (Figura 2), localizada na região sudeste do Brasil. Sua área de drenagem abrange aproximadamente 83.400 km², dos quais 86% fazem parte do Estado de Minas Gerais e os 14% restantes fazem parte do Estado do Espírito Santo (Aguiar *et al.*, 2020). A extensão do Rio Doce é de aproximadamente 853 km, e é formado pela confluência dos Rios Carmo e Piranga (Santolin *et al.*, 2015; dos Reis *et al.*, 2017).

A temperatura média anual na bacia varia entre 20°C e 22°C. As precipitações concentram-se no verão, com acumulado anual variando de 1.250 a 1.500 mm (Mello *et al.*, 1999). O padrão de chuvas na região é bem-marcado, com a ocorrência de dois períodos distintos. A estação chuvosa estende-se dos meses de outubro a março, 85% da precipitação anual distribui-se nesse período, e a estação seca vai de abril a setembro (Cupolillo; Abreu; Vianello, 2008; Kütter *et al.*, 2023).

A população total residente na bacia, até o ano de 2021, foi estimada em 3,6 milhões de habitantes, distribuídos em 225 municípios, dos quais 200 são mineiros e 25 capixabas. Desse total, apenas 44 tratam no mínimo 30% do esgoto coletado. Há predominância das soluções coletivas para a coleta do esgoto, que atendem 82% da população. Contudo, nem todo esgoto coletado é tratado, sendo que apenas 23,5%

da população atendida pelas soluções coletivas de coleta dispõe também de tratamento de esgoto. Cerca de 68% de todo o esgoto produzido na bacia é lançado diretamente nos mananciais, sem receber nenhum tipo de tratamento (ANA, 2016; ENGECORPS, 2021; ANA, 2023a).

Figura 2 - Localização geográfica da bacia do Rio Doce



O relevo da bacia do Rio Doce é bastante acidentado, caracterizado pelos chamados “mares de morro”, que condicionou a dinâmica de ocupação do espaço favorecendo o desenvolvimento das cidades junto à rede fluvial, e em áreas suscetíveis à inundação. As planícies fluviais são amplas e os terraços em geral possuem 3 metros de desnível. Quanto aos padrões geomorfológicos do Rio Doce, há segmentos de meandros, retilíneos e anastomosados, além de ilhas no médio Doce (ECOPLAN-LUME, 2010).

De montante à jusante há quatro usinas hidrelétricas instaladas no Rio Doce, pois seu declive de aproximadamente 1200 metros associado à elevada sinuosidade gera potencial favorável à produção de hidroeletricidade. São elas: Risoleta Neves

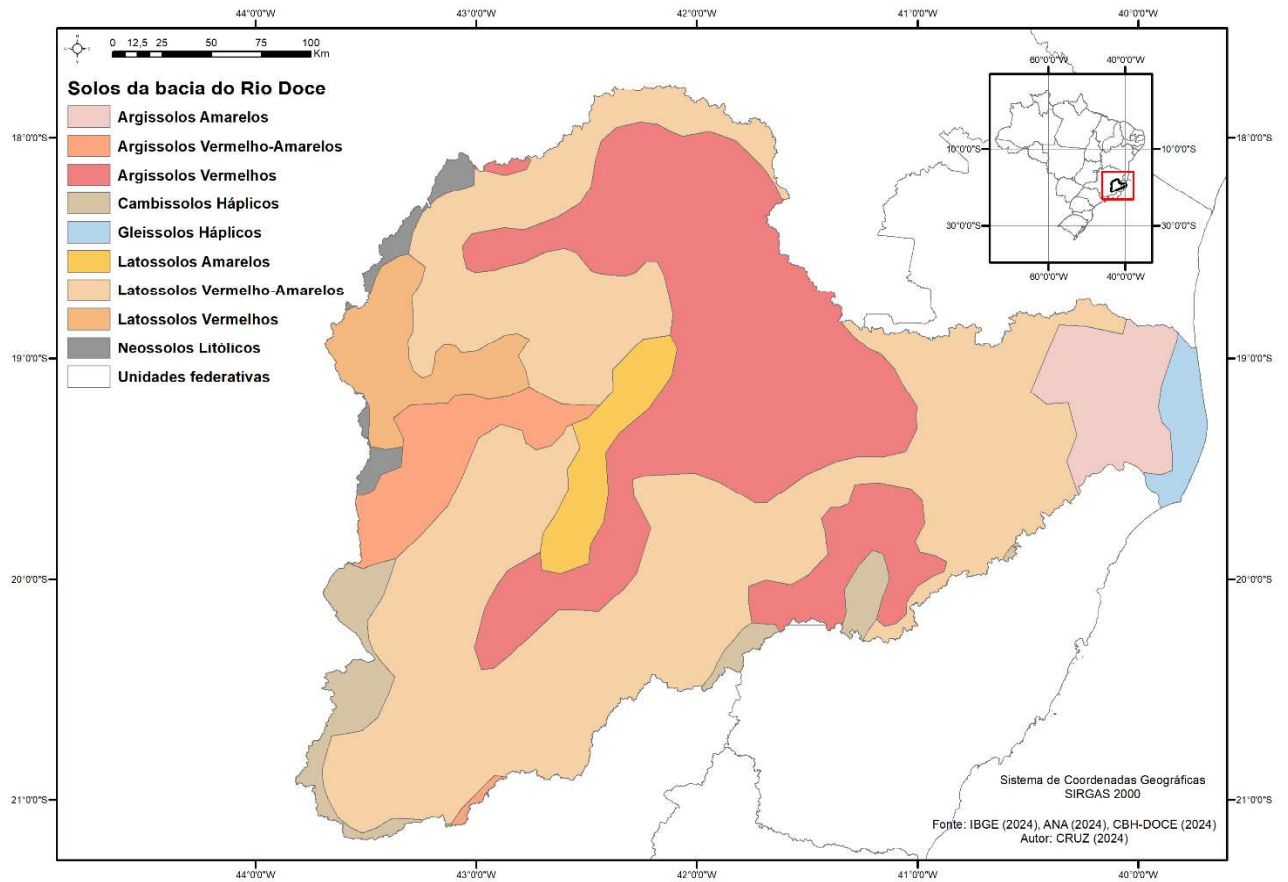
(Candongá) (Alto Doce), Baguari (Médio Doce), Aimorés (Médio Doce) e Mascarenhas (Baixo Doce). A operação das usinas depende diretamente do regime de vazões do Rio Doce, cuja média é de 900 m³/s, alcançando vazão máxima nos picos de cheia de 5000 m³/s (Kütter *et al.*, 2023).

A sinergia das alterações no escoamento da água e das taxas de sedimentação decorrentes das barragens, ao longo da calha principal do Rio Doce, e da baixa proteção dos solos da bacia, decorrente das profundas alterações da paisagem, promovem uma expressiva susceptibilidade à erosão, totalizando 58% da área total da bacia (Yamamoto *et al.*, 2023).

Essa fragilidade física dos solos da bacia decorre em parte de suas características pedológicas (Figura 3). Em toda a extensão da bacia do Rio Doce há 8 tipos de solos mais representativos: Argissolo amarelo, Argissolo vermelho-amarelo, Argissolo vermelho, Cambissolo háplico, Gleissolo háplico, Latossolo amarelo, Latossolo vermelho-amarelo, Latossolo vermelho e Neossolo litólico (CBH-DOCE, 2024). Destes, os Latossolos vermelho-amarelo e Argissolos vermelho constituem as classes com maior extensão, cobrindo respectivamente 27,5% e 24,4%. (ENGEORPS, 2021).

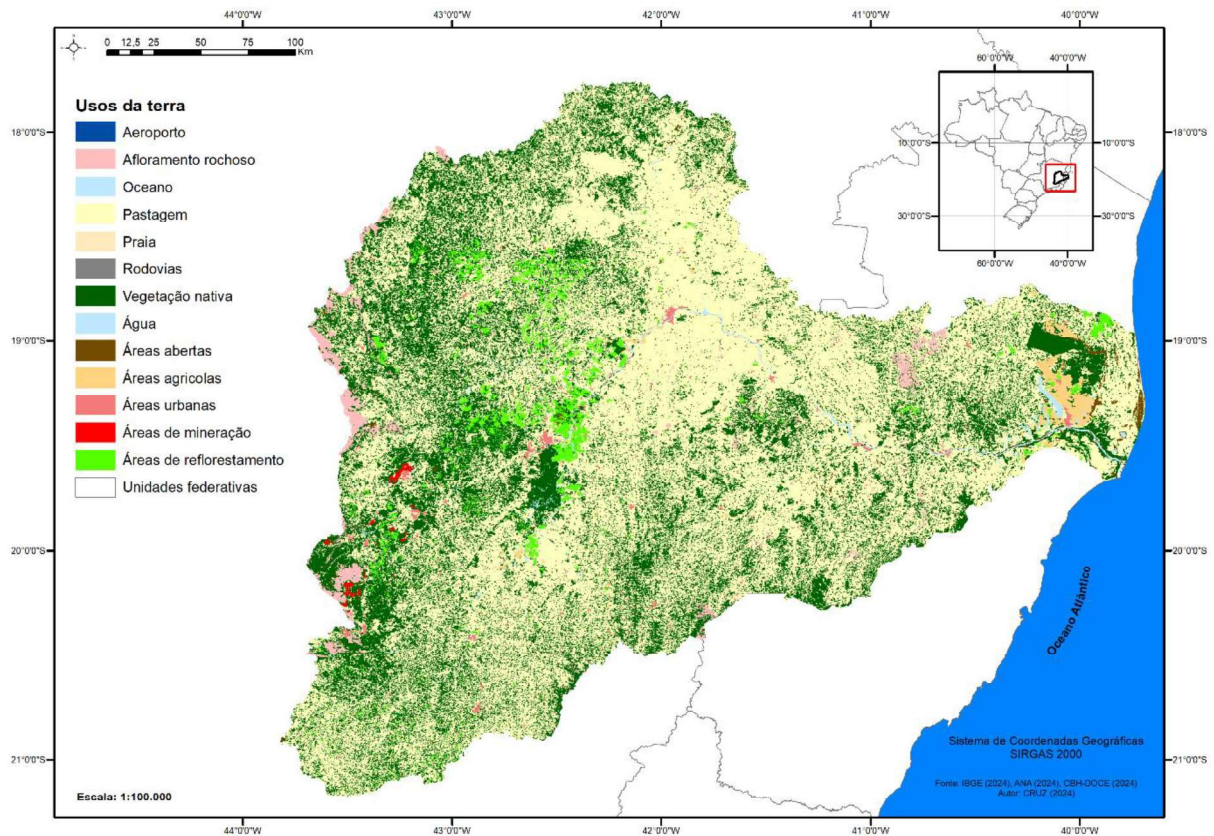
Os latossolos de maneira geral caracterizam-se por serem solos antigos, bastante evoluídos e intemperizados, por isso frequentemente são desprovidos de minerais primários e secundários (Jacomine, 2009). A classe particular dos Latossolos vermelho-amarelos tem predominância em terraços dissecados constituindo solos, solos profundos e acentuadamente drenados (IBGE, 2015). Os Argissolos, por outro lado, representam solos com elevada teor de argila, especialmente no horizonte B (Jacomine, 2009). Sua tipologia vermelha é caracterizada por solos de profundidade variável constituídos por material mineral oriundo de gnaisses, além de xistos e magmáticos (ENGEORPS, 2021; IBGE, 2015).

Figura 3 – Distribuição dos tipos de solos da bacia do Rio Doce



Além dos solos, a distribuição dos diferentes usos da terra (Figura 4) também tem impacto sobre a vulnerabilidade à erosão e o grau de conservação ambiental da bacia. Devido a seu processo histórico de sucessivos ciclos econômicos nocivos, que impuseram profundas mudanças na bacia com passivos ambientais que se perpetuam até os dias atuais, a bacia do Rio Doce é caracterizada por grande diversidade de usos alternativos da terra com pequena proporção do bioma original preservado (Fischer, 2018; Marinho Júnior; Espindola; Nodari, 2024; Marinho; Espindola, 2021; Espindola, 2015).

Figura 4 – Usos da terra na bacia do Rio Doce



A bacia ainda contém cerca de 33,4% de vegetação preservada, aí incluídos os biomas Mata Atlântica e Cerrado. Contudo, essas áreas não são contínuas, ao contrário, ilustram um cenário de intensa fragmentação da paisagem. Os 66,6% restantes são ocupados por usos alternativos da terra (ENGEORPS, 2021).

Os remanescentes florestais da bacia concentram-se nas áreas de maior declividade situadas na região das cabeceiras do Rio Doce. Pastagens e campos são os usos econômicos da terra predominantes em toda a bacia, representando cerca de 48,1% da área total. Os setores central e leste são as regiões com maior predominância desse uso, que está intimamente ligado à agropecuária extensiva praticada na bacia tanto para fornecimento de carne quanto leite. Os municípios de Governador Valadares e Linhares são os principais centros produtores desses produtos em âmbito regional (ENGEORPS., 2021).

Agricultura e silvicultura também têm representatividade expressiva no cenário atual da bacia, ocupando respectivamente 10,9% e 3,4 % da área total. A primeira concentra-se nas porções centro-sul e leste da bacia, com predominância de culturas

permanentes como o café, na zona da mata mineira e nos municípios de São Domingos do Norte, Rio Bananal, Governador Lindenberg, Vila Valério, Jaguaré, Linhares e Brejetuba, no Espírito Santo. Outras culturas importantes na bacia são pimenta-do-reino, cacau, coco-da-baía, de cana-de-açúcar (ENGECORPS., 2021).

Já a Silvicultura tem destaque especialmente no corredor ao redor do Rio Doce e na confluência dos Rios Piranga e Piracicaba. Há grande predominância de florestas plantadas de Eucalipto na região, que abastecem principalmente a indústria siderúrgica com madeira. Os demais usos da terra, sejam eles naturais ou antropogênicos, são menos representativos em área efetivamente ocupada individualmente, embora, economicamente alguns deles tenham grande relevância como é o caso da mineração (ENGECORPS., 2021).

A mineração tem destaque especialmente na região do Quadrilátero ferrífero. Essa atividade iniciou-se com a exploração de ouro pelos portugueses durante o século XVII, de maneira rudimentar, nos sedimentos de aluvião. Estima-se que aproximadamente 600 toneladas de ouro foram extraídas, desde então. Apesar dessa importância, a falta de manejo ambientalmente adequado levou à contaminação por Arsênio em diversos compartimentos ambientais, especialmente os sedimentos, águas superficiais, solos e águas subterrâneas, nas bacias do rio do Carmo e Gualaxo do Norte (de Carvalho *et al.*, 2018; Figueiredo; Borba; Angélica, 2007; Teixeira *et al.*, 2020; Silva *et al.*, 2018).

A mineração artesanal que havia na bacia foi gradualmente sendo substituída por forte mecanização. Além disso, outros minérios passaram a exercer interesse econômico na região, principalmente o Ferro, a partir da metade do século XX. Atualmente mais de 50% das exportações brasileiras de minérios deve-se ao minério de Ferro (Teixeira *et al.*, 2020). Com efeito, a partir da criação da empresa Vale do Rio Doce e do papel estratégico do Estado Brasileiro ao solucionar os problemas de financiamento e organização do setor minerário no Brasil, houve rápido crescimento dessa atividade acompanhado da implantação de infraestruturas tanto para a verticalização da produção quanto para seu escoamento, caso da Estrada de Ferro Vitória Minas (Espindola; Elisa; Mifarreg, 2017).

Esse cenário promoveu aumento do impacto da mineração na economia do Estado de Minas Gerais e do Brasil. De fato, dentre os 10 municípios com maior produção de minérios no Brasil, 7 estão localizados em Minas Gerais, incluindo o maior produtor – Itabira, situado na bacia do Rio Doce. Além disso, dentre os 41 municípios com maior produção de minério de Ferro no Brasil, 21 estão na bacia do Rio Doce. Contudo, apesar dos impactos positivos desse cenário, a verticalização da produção mineral no Quadrilátero Ferrífero impôs a necessidade de mecanismos de controle ambiental que permitissem a disposição do crescente volume de rejeito derivado da atividade de mineração do Ferro (Espindola; Elisa; Mifarreg, 2017; dos Santos *et al.*, 2023).

Dessa forma, ocorreu paulatinamente a construção de barragens de rejeitos ao longo da bacia. Barragens de rejeito são estruturas frequentemente construídas com material derivado do próprio local onde estão situadas, sendo comum o uso de solo (Kossoff *et al.*, 2014) ou material estéril gerado pela própria atividade de mineração para armazenamento dos rejeitos (Rico; Benito; Díez-Herrero, 2008; Cordeiro *et al.*, 2019). Os rejeitos de mineração têm grande diversidade de constituição física e química, podendo conter eventualmente metais e metaloides (Roche; Thygesen; Baker, 2017), representando um risco potencial especialmente aos recursos hídricos.

As barragens de contenção representam mecanismos de controle ambiental importantes para o armazenamento de rejeitos da atividade minerária, isolando-os de componentes como os solos e recursos hídricos. Apesar disso, tem-se evidenciado inúmeros episódios de rompimento de barragens com graves efeitos sobre os sistemas fluviais e ecossistemas aquáticos, muito em decorrência das escolhas técnicas equivocadas para prolongamento da vida útil dessas estruturas ou por negligência no controle dos riscos de acidentes (dos Reis Braga; Castro Fonseca, 2021; Halabi *et al.*, 2022; Armstrong; Petter; Petter, 2019).

O Brasil ao longo de sua história registrou 13 desses episódios (Halabi *et al.*, 2022). Dois deles são considerados os maiores desastres ambientais da atividade minerária no século XXI (Freitas; Andrade da Silva, 2019; Burritt; Christ, 2018).

Hoje há grande concentração dessas barragens na encosta leste da Serra do Espinhaço, particularmente nas cabeceiras dos principais afluentes à margem esquerda do Rio Doce, no caso os Rios Piracicaba, do Carmo, Santo Antônio,

Corrente, Suaçuí Grande e Suaçuí pequeno. Estima-se que haja cerca de 108 barragens de contenção de rejeitos de mineração na bacia do Rio Doce (Espindola; Elisa; Mifarreg, 2017; dos Santos *et al.*, 2023).

5.2 Desastre de Fundão

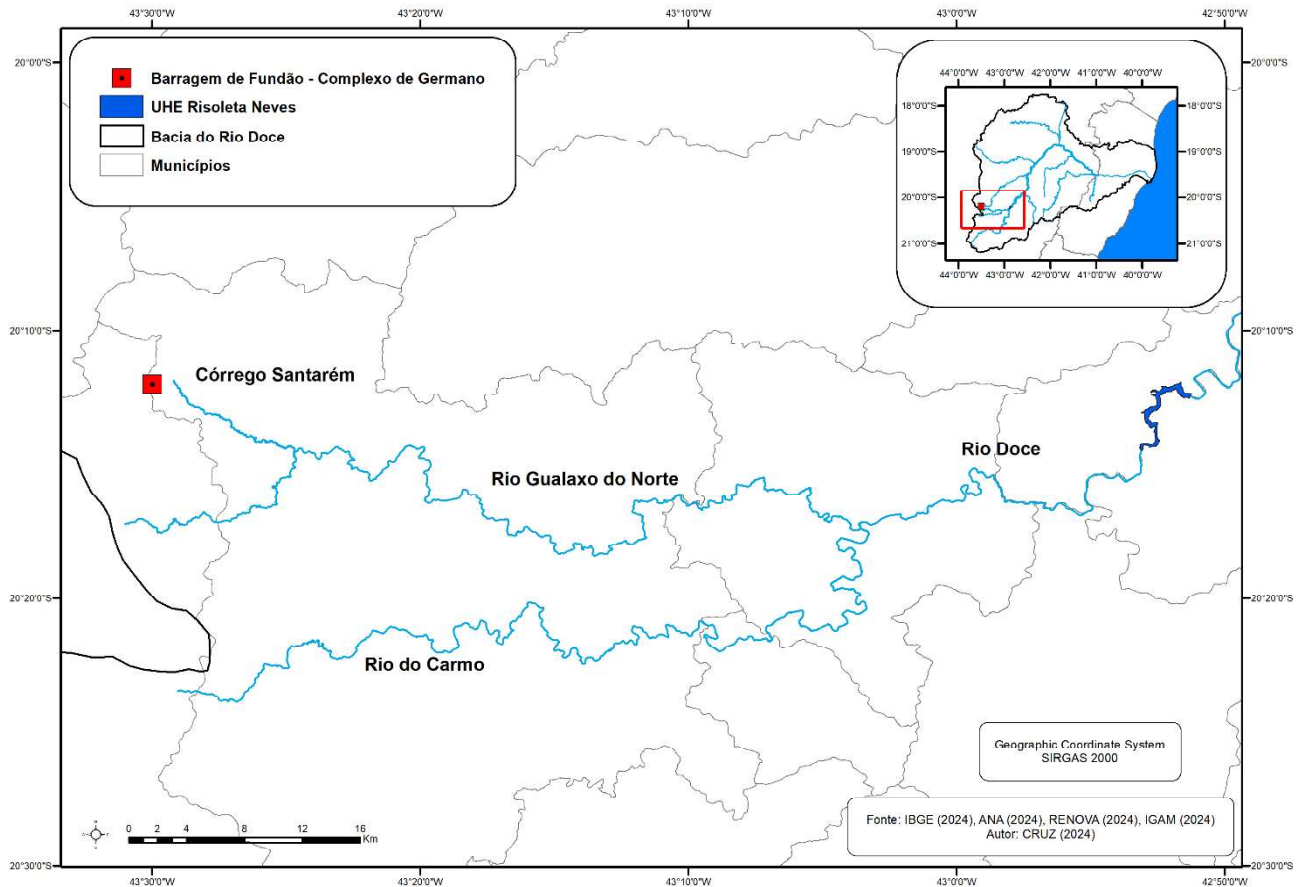
A barragem de Fundão, uma instalação de armazenamento de rejeitos de grande porte mantida pela empresa Samarco Mineração S.A., situada no distrito de Bento Rodrigues (Mariana/MG), contendo partículas minerais finas e grossas resultantes do processamento de minério de Ferro, nas cabeceiras da bacia do Rio Doce – se rompeu na tarde de 5 de novembro de 2015. Da barragem, classificada à época como Classe II (alto potencial de dano ambiental), 44 milhões de metros cúbicos de rejeitos liberados atingiram a barragem de Santarém, à jusante, causando seu galgamento e forçando a passagem da onda por 55 Km no Rio Gualaxo do Norte até desaguar no Rio do Carmo (Sánchez *et al.*, 2018; IGAM, 2020a; dos Santos *et al.*, 2023; Duarte *et al.*, 2021).

A partir daí, a pluma percorreu 22 Km até seu encontro com o Rio Doce e deste seguiu até a foz da bacia, onde por fim, os rejeitos foram lançados ao Oceano Atlântico em 21 de novembro de 2015, totalizando aproximadamente 663 Km de mananciais diretamente impactados (Sánchez *et al.*, 2018; IGAM, 2020a; Da Silva *et al.*, 2022b; Silva *et al.*, 2018). Estima-se que aproximadamente 10 milhões de metros cúbicos tenham sido depositados na calha dos rios e planície fluvial no trecho que vai da barragem de Fundão até a usina hidrelétrica (UHE) Risoleta Neves, distante aproximadamente 113 km do ponto de rompimento da barragem (Figura 5). A barragem de Candonga também agiu amortecendo a onda de rejeitos e favorecendo a deposição de cerca de 20 milhões de metros cúbicos do material mais fino constituinte dos rejeitos (Da Silva *et al.*, 2022b; Figueiredo *et al.*, 2020; Davila *et al.*, 2020).

Investigações realizadas posteriormente pelo Ministério Público Federal (MPF) dão conta que, sem mencionar as questões relacionadas à negligência do Estado no controle e monitoramento ambiental da mineração no Estado de Minas Gerais e a gestão obtusa dos riscos relacionados ao estado de conservação da barragem de conhecimento prévio da empresa, a barragem de Fundão colapsou como efeito de um

processo denominado liquefação associado a falhas na drenagem da bacia não corrigidas (Sapata Gonzalez *et al.*, 2022).

Figura 5 – Trecho impactado pela onda de rejeitos da barragem de Fundão até a UHE Risoleta Neves



A liquefação é o comportamento anômalo de perda das propriedades sólidas que faz com que um material se comporte fisicamente como um líquido decorrente da aplicação de tensão de alguma origem, como terremotos ou mesmo trânsito de maquinário pesado, tornando esse material instável e susceptível a falhas. Note-se ainda que a barragem de Fundão pertencia à categoria de alteamento à montante, cujos problemas de instabilidade e riscos de rompimento já são bem conhecidos pela engenharia (Kossoff *et al.*, 2014; Rico; Benito; Díez-Herrero, 2008; Sapata Gonzalez *et al.*, 2022).

Em decorrência do desastre registraram-se elevações significativas nas concentrações de metais e metaloides como Alumínio, Arsênio, Cádmio, Chumbo, Cobre, Cromo, Mercúrio e Níquel em diversas regiões da bacia (ANA, 2016). Quando

presentes na água esses poluentes oferecem elevado risco ao abastecimento público, uma vez que o tratamento convencional não os remove eficientemente (ANA, 2016).

Além dos impactos nos recursos hídricos foram evidenciados outros danos ambientais e sociais diretos, tais como a morte e desaparecimento de pessoas; isolamento de áreas habitadas; desalojamento de comunidades pela destruição de moradias e estruturas urbanas; fragmentação de habitats; destruição de áreas de preservação permanente e vegetação nativa; mortandade de animais de produção e impacto à produção rural e ao turismo, com interrupção de receita econômica; restrições à pesca; mortandade de animais domésticos; mortandade de fauna silvestre; dizimação de ictiofauna silvestre em período de defeso; dificuldade de geração de energia elétrica pelas hidrelétricas atingidas; além da sensação de perigo e desamparo da população em diversos níveis (IBAMA, 2015; Sánchez *et al.*, 2018; de Carvalho, 2019).

Como efeito indireto do desastre outros importantes impactos sociais, econômicos e culturais também acarretaram grande prejuízo à sociedade, com destaque para: perda de patrimônio cultural, perda de capacidade de geração hidrelétrica da UHE Risoleta Neves, interrupção da pesca no Rio Doce e no Oceano Atlântico na região estuarina, suspensão da atividade de mineração, redução do turismo, perda de empregos na atividade de mineração, redução de receita decorrente do pagamento de *royalties* aos municípios e interrupção do abastecimento de água em muitas cidades e comunidades ribeirinhas (Sánchez *et al.*, 2018; De Carvalho, 2019; Sapata Gonzalez *et al.*, 2022).

Nesse último caso, esse impactou devido-se à elevada concentração de sedimentos finos em suspensão nos mananciais impactados pelos rejeitos, que inviabilizou a operação das estações de tratamento da água (ETA) utilizadoras do Rio Doce e a captação feita no rio por comunidades rurais e outros usuários, como os agricultores do baixo Doce. No total, o abastecimento foi suspenso em 12 cidades ao longo do Rio Doce, afetando uma população estimada em 424.000 pessoas, com destaque para Governador Valares em Minas Gerais e Colatina no Espírito Santo (IBAMA, 2015; Sánchez *et al.*, 2018; de Carvalho, 2019; Sapata Gonzalez *et al.*, 2022).

Face à dimensão dos impactos do rompimento da barragem de Fundão, muitos autores têm se dedicado a caracterizar os rejeitos derivados da barragem. Apesar de haver concordância dos autores sobre a constituição física desse material na maioria dos casos, por vezes há pequenas divergências sobre a constituição química do material dispersado, com alguns estudos apontando alguns elementos significativos que não corroboram com a investigação de outros (Figueiredo *et al.*, 2020; Davila *et al.*, 2020; Duarte; Neves, 2023; Segura *et al.*, 2016).

Esse fato parece manter relação com a seleção dos locais de monitoramento e as condições das amostragens, uma vez que em todos os locais por onde os rejeitos foram depositados a lama resultante pode ter influência não apenas dos rejeitos, mas também dos sedimentos de fundo e solos das áreas marginais de toda a rede hidrográfica atingida (Santolin *et al.*, 2015; Bertoldo *et al.*, 2023).

A caracterização da lama realizada por Figueiredo *et al.* (2020), a partir de amostras coletadas à superfície e a 1 m de profundidade distribuídas ao longo de toda a extensão atingida pela onda de rejeitos de Fundão, sugere não haver toxicidade, corrosividade e reatividade na lama. Quimicamente, as amostras apresentam forte presença de Fe, Si e Al, como esperado para rejeitos de mineração de Ferro, com relativa proporção também de K, Mn e Ti. A granulometria das amostras foi bastante variável, ficando em termos gerais no espectro de 1 a 200 μm , com média de 55 μm . Partículas com tamanho de até 123 μm representaram 90% das amostras.

Em outro estudo conduzido por Davila *et al.* (2020) em que foram analisadas amostras de lama coletadas diretamente da barragem de Fundão além de haver predominância de Fe, em concordância com Figueiredo *et al.* (2020), foram identificados também níveis mais elevados de Mn e Cr. Da mesma forma, os resultados de Duarte e Neves (2023) destacam Fe, Si e Al como os principais constituintes da lama de Fundão. Embora em menores percentuais tenham sido detectados K, Ca, Ti e Mn em todas as amostras.

Em pesquisa semelhante, conduzida por Segura *et al.* (2016) com amostras de lama coletadas na superfície após o *topsoil* em Bento Rodrigues (distrito de Mariana - MG, devastado pela passagem da onda de rejeitos de Fundão), os autores encontraram resultados concordantes com Figueiredo *et al.* (2020) para a granulometria das

partículas e com Figueiredo *et al.* (2020), Davila *et al.* (2020) e Duarte e Neves (2023), em relação aos principais constituintes químicos dos rejeitos, apesar de que níveis mais elevados de Cu não foram observados pelos outros autores.

Além disso, o trabalho de Segura *et al.* (2016), ao realizar a extração de metais por diferentes métodos, alerta para potencial mobilidade de Ba, Sr, Fe, Mn e Al com possibilidade de contaminação de ecossistemas aquáticos e seres humanos, em condições ambientais específicas, com grave risco de cronicidade.

5.3 Considerações sobre o monitoramento e a qualidade da água pré-desastre na bacia do Rio Doce

A bacia do Rio Doce historicamente passou por um dramático processo de modificação de suas características naturais, fruto do padrão de ocupação predatório implantado pelos diferentes ciclos econômicos que se sucederam até os dias atuais (Espindola, 2015). Essas alterações, fatalmente também produziram efeitos sobre os mananciais que drenam a bacia, tanto em termos quantitativos quanto em níveis qualitativos (Felippe *et al.*, 2016).

Dentre os principais fatores que exercem pressão sobre a qualidade da água na bacia pode-se elencar a baixa cobertura vegetal da bacia impactada pela intensa fragmentação da paisagem, a falta de saneamento básico nos centros urbanos produto do crescimento desordenado das cidades, a exploração mineral (especialmente no Quadrilátero ferrífero), a implantação da indústria siderúrgica e a agricultura/pecuária extensionista praticada com baixo uso de tecnologia (ECOPLAN-LUME, 2010; ENGECORPS, 2021).

No período anterior ao desastre havia duas redes de monitoramento regulatórias em operação na bacia do Rio Doce, as redes do Instituto Mineiro de Gestão de Águas (IGAM) e a rede da Agência Nacional de Águas e Saneamento Básico (ANA) operada pelo Serviço Geológico do Brasil (CPRM).

No caso de qualidade das águas, os dados do IGAM representam a principal fonte de informação periódica, sistemática e confiável do período pré-desastre e tem sido bastante utilizada em pesquisas que buscam estabelecer um paralelo entre as

condições anteriores e atuais da qualidade da água na bacia (<http://www.repositorioigam.meioambiente.mg.gov.br/jspui/handle/123456789/411>).

Além dessa fonte também é possível identificar estudos independentes, que coletaram e analisaram amostras de água e sedimento do Rio Doce e/ou afluentes durante períodos menores de monitoramento.

Petrucio *et al.* (2005) ao realizarem campanhas de monitoramento em 7 estações durante os anos de 2000 e 2001, em sub-bacias do médio Rio Doce, observaram valores críticos de Oxigênio dissolvido (2 mg/L), Fósforo total (843 µg/L) e coliformes totais ($>160 \times 10^3$ NMP/100mL) no rio Ipanema. Clorofila a apresentou altos valores no rio Severo (53,5 µg/L), enquanto no Rio Piracicaba foram registrados, tal como no rio Ipanema, níveis elevados de coliformes totais ($>160 \times 10^3$ NMP/100mL). Ao considerar as características da região os autores atribuíram os resultados à carência de saneamento básico dos centros urbanos dos principais municípios e ao lançamento de efluentes industriais.

Esse mesmo cenário crítico foi corroborado pela investigação da qualidade da água através do Índice de Qualidade da Água da Fundação Nacional de Saneamento (IQAFNS), considerando os dados do IGAM medidos em 32 estações no período de 2000 a 2017. Os resultados demonstraram haver tendência à redução da qualidade da água em locais à jusante de grandes centros urbanos, atribuído à descarga de esgoto doméstico e industrial não-tratados diretamente nos mananciais da bacia do Rio Doce (Fraga *et al.*, 2021).

Essa hipótese é corroborada pela tendência de aumento de *E. coli* evidenciada em algumas estações, pela baixa cobertura relativa à presença de estações de tratamento de esgoto (ETE) e pelo crescimento da infraestrutura urbana acima da média nacional ocorrido na bacia no mesmo período (Fraga *et al.*, 2021). Os autores afirmam que o rompimento da barragem de Fundão e os rejeitos depositados na bacia não possuem ligação direta com as tendências evidenciadas.

Além da poluição orgânica, a presença de metais e metaloides na água e sedimentos já havia sido observada em diferentes regiões da bacia. Wild *et al.* (2023) compilaram dados de monitoramento do IGAM, do “Programa de Monitoramento Quali-

Quantitativo Sistemático Permanente de Água e Sedimentos da Bacia do Rio Doce” (PMQQS) e Instituto Estadual de Meio Ambiente do Espírito Santo (IEMA/ES) para fins de investigação da dinâmica espaço-temporal do retorno da qualidade da água do Rio Doce às condições pré-desastre. Os autores ao analisarem a excedência (taxa de não-conformidade com limites admissíveis de variáveis de qualidade da água) das variáveis frequentemente relacionadas à onda de rejeitos de Fundão, a partir dos dados pré-desastre; observaram que Manganês, Ferro dissolvido e Alumínio dissolvido apresentavam os percentuais mais expressivos, considerando a bacia de maneira geral.

Considerando a variabilidade espacial, as excedências de Alumínio aumentavam no sentido leste-oeste em direção à foz da bacia. O Arsênio, por sua vez, apresentava excedências limitadas apenas ao rio do Carmo e o Chumbo à montante e jusante de Governador Valadares. Manganês registrava elevadas taxas de excedência no alto Doce, sobretudo no Rio do Carmo (>70%), na transição alto-médio Doce e médio Doce (>50%). Ferro e turbidez apresentavam excedência mais uniforme ao longo de toda a bacia, à exceção do Ferro dissolvido no Rio do Carmo, que apresentou taxa um pouco menor (Wild *et al.*, 2023).

Outros autores identificaram padrões de enriquecimento de Ferro dissolvido e Turbidez, além de Alumínio dissolvido, em direção à foz da bacia no período pré-desastre, especialmente durante a estação chuvosa. Esse fato foi associado à carga de sólidos e a presença das fontes de metais na bacia (da Cunha Richard *et al.*, 2020). Na região do médio Rio Doce também há registros de anomalias em níveis perigosos à saúde humana de vários elementos, incluindo Ferro, Alumínio, Zinco e Manganês no Rio Doce, mas principalmente nos Rios Piracicaba e Ipanema (Jordão *et al.*, 1996).

Esse padrão foi diretamente relacionado aos impactos das indústrias instaladas há décadas nessa região. Essa região popularmente conhecida como Vale do Aço possui intensa atividade da indústria de transformação, especialmente a siderúrgica, marcada pela manipulação de metais. Em outras matrizes igualmente importantes como solos e sedimentos da região, também foram registradas concentrações elevadas de Cádmio, Chumbo, Cromo, Zinco e Cobre em níveis muito superiores

àqueles medidos em amostras coletadas em locais fora da área de influência da atividade industrial (Jordão *et al.*, 1996).

Na região das cabeceiras da bacia do Rio Doce, coletas de água realizadas no rio do Carmo em 12 pontos distintos, em Abril de 2015 haviam registrado concentrações de Arsênio em nível superior aos padrões de potabilidade adotados no Brasil. Amostras de sedimento de fundo coletadas na mesma oportunidade também apresentaram níveis igualmente elevados de Arsênio. Em ambos os casos, há evidências que associam esse comportamento ao efeito da mineração de Ouro ocorrida há séculos no Quadrilátero ferrífero, bem como as condições geológicas da região de Mariana e Ouro Preto, que favorecem o enriquecimento de Arsênio no Rio do Carmo ao drenar os municípios, sobretudo, na Mina da Passagem (Silva *et al.*, 2018).

Além das questões vinculadas às péssimas condições de saneamento e a influência da mineração, estudos anteriores já demonstravam o impacto da sazonalidade e das condições de uso/ocupação da terra também na qualidade da água na bacia. Nesse contexto, deve-se destacar o histórico frequente de episódios de floração de algas, que foi marcante no período pré-desastre. Esses episódios receberam bastante atenção vide os riscos à saúde da população relacionados ao abastecimento.

A investigação dos episódios ocorridos entre 2008 e 2012 demonstram haver grande influência da sazonalidade, com tendência a floração de algas durante a estação seca ou na transição da estação seca para a chuvosa, com presença de nitrogênio orgânico e amoniacal na água (Jardim *et al.*, 2014).

Mesmo em condições favoráveis a esse processo, quanto à presença de nutrientes, as condições de maior fluxo do período chuvoso atuaram como limitantes. A redução da turbidez na estação seca, ao contrário age como facilitadora desse processo. Além disso, as condições precárias de saneamento da bacia, já mencionadas, favorecem a descarga de efluentes diretamente nos mananciais, que somada ao aporte de nutrientes dos solos mantidos sob uso agrícola acabam também colaborando para condições propícias à floração de algas (Jardim *et al.*, 2014).

A sazonalidade marcante também está diretamente relacionada ao comportamento da Turbidez e carga de sedimentos em suspensão observados na bacia do Rio Doce.

Na estação chuvosa especialmente, o impacto de sistemas frontais e da Zona de Convergência do Atlântico Sul (ZCAS) na bacia promove chuvas intensas que aumentam dramaticamente o escoamento superficial. Com efeito, tanto há maior propensão à ocorrência de cheias, quanto a carga de sedimentos transportados assume proporções muito elevadas (Oliveira; Quaresma, 2017).

Esse último caso decorre das péssimas condições de conservação dos solos da bacia que por apresentarem baixa proteção e reduzida capacidade de infiltração, decorrente da baixa densidade de cobertura vegetal, favorecem o escoamento superficial e a exportação de sedimentos aos rios. Isso deve-se ao histórico padrão de ocupação da bacia em que a floresta foi substituída por usos alternativos como, agricultura, pastagens, mineração e uso urbano, tornando os solos mais propensos à erosão (Oliveira; Quaresma, 2017).

6 MATERIAL E MÉTODOS

A metodologia dessa pesquisa está apoiada em duas etapas distintas que convergem no atendimento ao objetivo geral (Figura 6). São elas: seleção de “variáveis não-críticas” de qualidade da água e seleção da melhor configuração espacial de estações de redes de monitoramento emergencial da qualidade da água.

6.1 Seleção de “variáveis não-críticas” de qualidade da água

6.1.1 Seleção e tratamento de dados de qualidade da água

Foram utilizados nessa pesquisa séries temporais de variáveis de qualidade da água integrantes do PMQQS, operado pela Fundação Renova na bacia hidrográfica do Rio Doce (RENOVA, 2017a).

A fundação RENOVA foi criada a partir do termo de transação e ajustamento de conduta (TTAC) firmado entre instituições federais, os Estados de Minas Gerais e Espírito Santo e as empresas SAMARCO MINERAÇÃO SA, Vale S.A. e BHP Billiton Brasil Ltd., definindo compromissos mútuos para restaurar a bacia hidrográfica à situação anterior ao rompimento da barragem (ENGECORPS, 2021).

O PMQQS foi desenvolvido pela Fundação Renova, em parceria com órgãos ambientais, agências de água, e membros da câmara técnica de segurança hídrica e

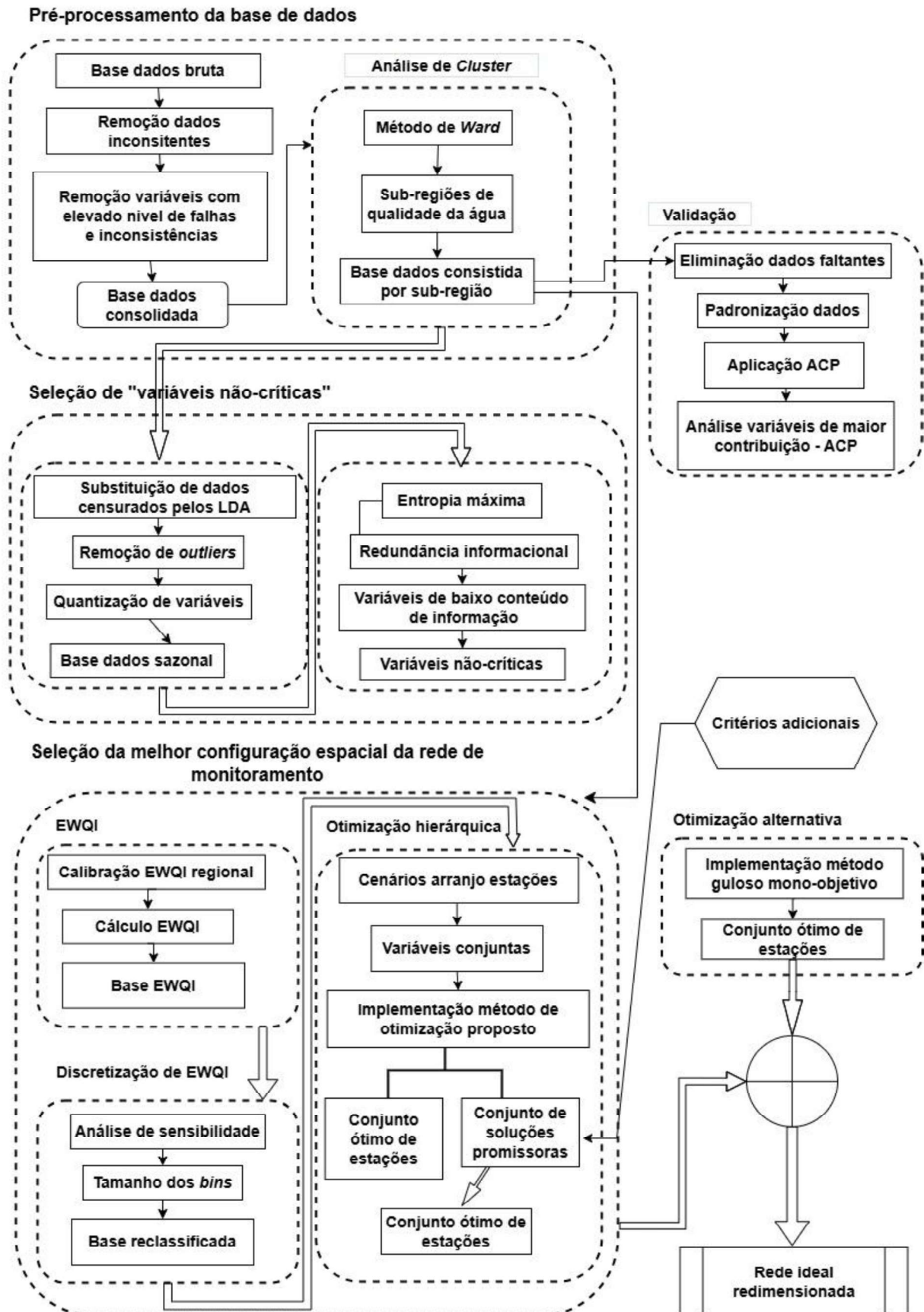
qualidade da água da estrutura de governança (CT-SHQA) criada a partir do TTAC, a fim de acompanhar a recuperação da bacia do Rio Doce e zona costeira e estuarina adjacentes (RENOVA, 2018).

Ele teve início em 31 de julho de 2017 e opera em caráter permanente. Atualmente, a rede opera em frequência mensal 92 estações distribuídas entre as lagoas do Rio Doce, zona costeira/estuarina, Rio Doce e seus afluentes. A seleção dos locais a monitorar pela fundação considerou os seguinte critérios (RENOVA, 2017b):

- Trechos da calha principal do Rio Doce e tributários atingidos diretamente pela pluma de rejeitos do rompimento da barragem de Fundão;
- Tributários com descarga líquida acima de 10% da vazão de referência Q_{95} estimada para o exutório da bacia;
- Locais próximos ao exutório de afluentes cuja disponibilidade hídrica seja inferior à demanda para diluição de efluentes, considerando como parâmetro de decisão a DBO_5 ;
- Trechos de tributários não atingidos pelos rejeitos, situados na região de montante da bacia;
- Locais já monitorados pela rede básica operada pelo IGAM e Instituto Estadual de Meio Ambiente e Recursos Hídricos do Espírito Santo (IEMA/ES), que apresentam séries históricas (*baseline*);
- Locais sugeridos por outras câmaras técnicas pertencentes ao sistema de governança criado pelo TTAC, pelos órgão ambientais dos Estados de Minas Gerais e Espírito Santo, e pela própria Fundação Renova.

Este estudo concentrou-se em corpos d'água lóticos, compreendendo um total de 39 estações de monitoramento utilizadas, da rede PMQQS (Figura 7) (Apêndice -Tabela A1).

Figura 6 – Metodologia proposta para o redimensionamento de redes emergenciais de qualidade da água



Para este estudo, todos os dados inválidos foram removidos. Assim como, foram descartadas as variáveis que apresentaram percentual de dados inválidos superior a 30% ou que apresentaram percentual de dados faltantes acima de 30%. Esse critério buscou prevenir possíveis problemas decorrentes da aplicação de métodos estatísticos aos dados. A ausência significativa de dados medidos e validados poderia refletir em grande incerteza sobre a confiabilidade dos resultados. Além disso, a imputação de grande proporção de dados em substituição a dados censurados ou faltantes deve ser tratado sempre com bastante cautela, sendo não recomendada em casos de elevado risco de adição de viés aos resultados (Sperling; Verbyla; Oliveira, 2020).

Para a seleção de “variáveis de qualidade da água não-críticas” adicionalmente os dados abaixo ou acima do limite de detecção analítica (LDA) (dados censurados) foram substituídos pelos respectivos limites de detecção, de forma a permitir a utilização do maior volume possível de dados e testar a capacidade do método em selecionar variáveis com baixo conteúdo informativo, muitas delas com elevado percentual de dados censurados.

Dados censurados de qualidade da água representam situações em que se procedeu a análise das variáveis, mas não se obteve uma medida quantificável. Eles podem ocorrer tanto acima quanto abaixo do LDA. Há diversas alternativas de tratamento aplicáveis a bases que contém dados censurados: eliminação de todos os dados censurados, substituição dos dados censurados por 0, substituição dos dados censurados por uma fração do LDA, imputação de valores aos dados censurados baseado em métodos mais sofisticados e substituição dos dados censurados pelo valor do LDA (Sperling; Verbyla; Oliveira, 2020).

Cada uma das alternativas deve ser cuidadosamente analisada de acordo com o contexto específico de investigação da qualidade da água e do tratamento estatístico que se pretende dar aos dados, pois há limitações que podem introduzir viés às análises. A simples substituição dos dados censurados pelos valores de LDA tende a gerar valores de estatística descritiva, como média, mais elevados que a realidade (Sperling; Verbyla ; Oliveira, 2020). Apesar disso, essa foi a escolha para o tratamento dos dados censurados neste estudo.

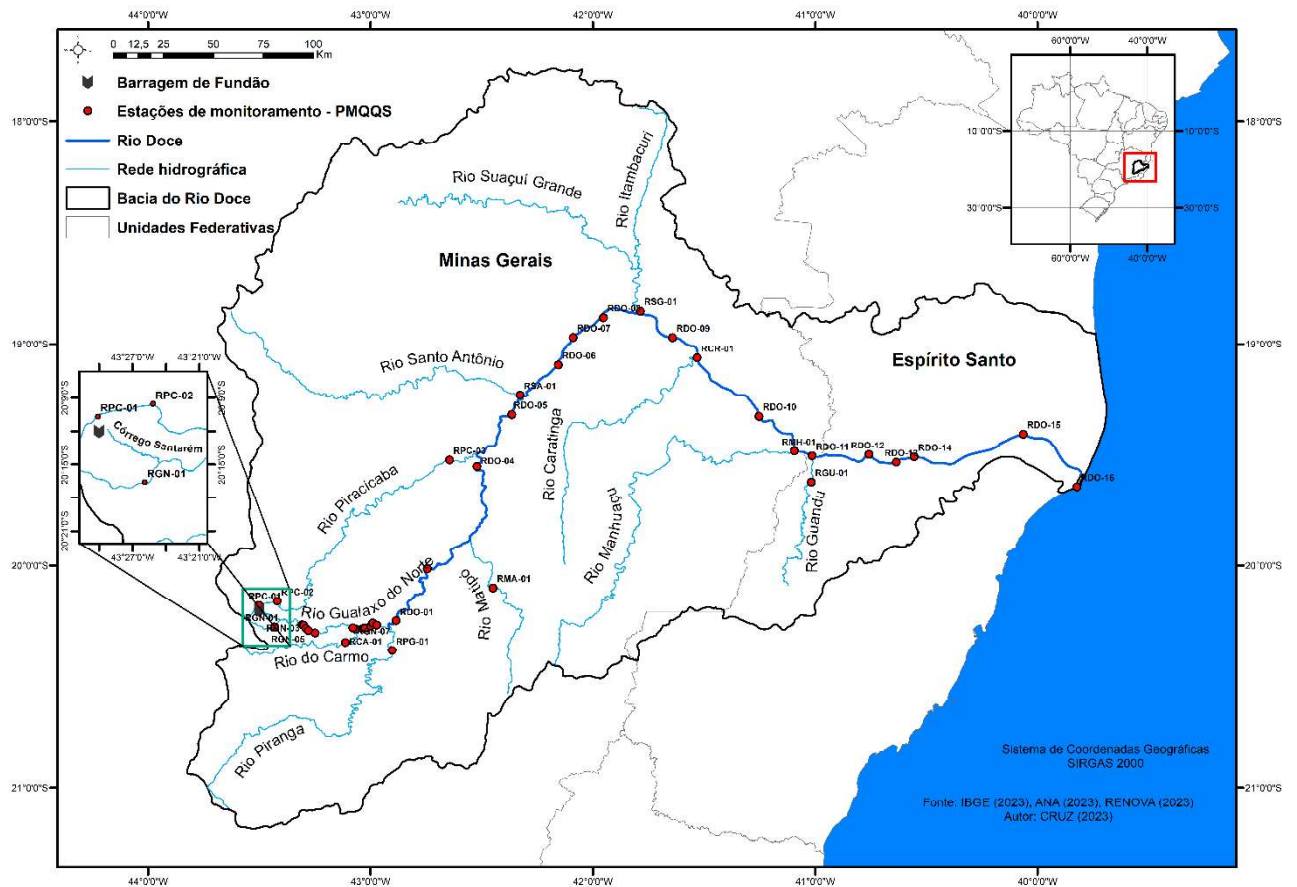
A justificativa para essa escolha está no fato de que a metodologia proposta para seleção das variáveis está centrada no uso de entropia de Shannon, e especialmente na determinação da entropia marginal e máxima de cada variável, como será descrito com maior riqueza de detalhes nas seções seguintes. Todas as variáveis serão quantizadas, passando de aleatórias contínuas a discretas. Dessa forma, mesmo os dados de variáveis situadas abaixo do LDA agregam uma informação útil, representando medidas situadas em uma faixa estreita de valores que possui um grau de incerteza associado e, portanto, entropia.

Ainda para essa etapa em particular, também foi necessário realizar um tratamento para identificar e remover *outliers* situados na cauda superior da distribuição, pois a entropia de Shannon, e particularmente a entropia máxima, são sensíveis a eles (Nooghabi; Nooghabi, 2016; Kafantaris *et al.*, 2020). Conforme será mais bem descrito nas seções posteriores, o critério proposto para seleção de variáveis não-críticas de qualidade da água passa pela determinação da entropia máxima de cada variável. No caso de variáveis categóricas, caso recorrente em virtude da quantização das variáveis de qualidade da água, a entropia máxima pode ser estimada pelo uso da distribuição uniforme (Singh, 2013).

Nesse cenário, dado que a distribuição uniforme assume como parâmetros os valores extremos da amostra, de fato, a presença de *outliers* na cauda superior da distribuição adiciona variância aos dados pelo aumento da amplitude amostral e, por extensão, do número de classes derivadas da quantização ou de qualquer outro critério de discretização. Com efeito, esse processo pode projetar valores de entropia máxima superior à condição real das variáveis de qualidade da água.

Esse fato foi confirmado nos primeiros exercícios de aplicação da metodologia para seleção das variáveis não-críticas, ocasião em que o número de variáveis selecionadas por esse critério foi bastante inferior ao esperado, tendo-se confirmado posteriormente, que essa distorção se deu em decorrência da presença dos *outliers*. Vários algoritmos para detecção de *outliers* foram testados e o método de Gini (NAIR, 1936) resultou nas melhores estimativas.

Figura 7 - Estações de monitoramento da qualidade da água da rede PMQQS selecionadas



6.1.1.1 Critérios de análise espaço-temporal dos dados

Em bacias hidrográficas extensas a existência de diferentes contextos regionais e a sazonalidade podem influenciar o conteúdo de informação característico de cada variável de qualidade da água investigada.

Dessa forma, a bacia hidrográfica do Rio Doce foi dividida em subregiões homogêneas de qualidade de água por meio da aplicação da análise de *cluster* para identificar grupos de estações de monitoramento com forte similaridade. A análise de *cluster* é um método estatístico multivariado usualmente utilizado em análises exploratórias para identificar padrões em conjuntos de dados e reunir observações semelhantes em grupos (Kettenring, 2006).

Neste estudo, adotou-se o método de Ward (Ward, 1963), que é uma técnica hierárquica aglomerativa, e tomou-se a distância euclidiana como medida de

dissimilaridade. O método de Ward é caracterizado pela formação de grupos bastante homogêneos com mínima variância interna (Szekely; Rizzo, 2005) e é extensivamente aplicado em estudos de qualidade de água (Szekely; Rizzo, 2005; Azhar *et al.*, 2015; da Silva, 2020; Lobo *et al.*, 2015; Hajigholizadeh; Melesse, 2017; Li; Tian; Liu, 2019; Kändler *et al.*, 2017; Pinto; Almeida; Oliveira, 2018; Jiang *et al.*, 2020b).

Os dados foram padronizados, conforme sugerido por Härdle e Simar (2015), para eliminar distorções decorrentes das diferentes escalas em que as variáveis de qualidade da água são medidas. A determinação do número ideal de *clusters* baseou-se na análise do comportamento de fusão no dendrograma (Hair Jr *et al.*, 2006), associada ao conhecimento sobre a distribuição espacial das estações.

Para incorporar os efeitos da variabilidade temporal, a base foi dividida em dois conjuntos representativos do período seco (abril a setembro) e chuvoso (outubro a março) da bacia.

6.1.1.2 Determinação da redundância informacional

A redundância informacional de cada variável de qualidade da água por estação de monitoramento foi determinada a partir de suas respectivas séries temporais. As etapas para isso foram: i) discretização dos dados; ii) cálculo das entropias marginais de acordo com a Equação 1; iii) cálculo da entropia máxima (seção 6.1.1.4); iv) obtenção da entropia relativa e redundância informacional (Equação 7).

O procedimento aplicado neste trabalho para discretização de variáveis de qualidade da água é descrito a seguir (seção 6.1.1.3). A discretização de variáveis aleatórias foi necessária porque é particularmente complexa a determinação da entropia marginal de variáveis contínuas. O uso de distribuições discretas, mesmo em redes de monitoramento onde predominam variáveis contínuas, tem se tornado cada vez mais frequente (Mirauda; Ostoich, 2020; Keum; Coulibaly, 2017b). A entropia máxima foi estimada utilizando simulação de Monte Carlo tomando a distribuição uniforme como referência, como explicado na seção 6.1.1.4

Foram elaborados gráficos de Box-Whisker para avaliar a entropia marginal e máxima considerando cada sub-região da qualidade da água e a sazonalidade dos dados.

6.1.1.3 Quantização das variáveis

A discretização de variáveis aleatórias pelo método do histograma é uma alternativa frequentemente adotada, onde os intervalos das classes do histograma são estabelecidos através de algum critério arbitrário (Sturges, 1926; Scott, 1979; Freedman; Diaconis, 1981; Terrell; Scott, 1985). Apesar de simples, como não há um critério universal para determinação do tamanho do *bin* ideal, diferentes dimensões podem ser obtidas, o que, por sua vez, leva a diferentes valores de entropia (Sreeparvathy; Srinivas, 2020; de Pádua *et al.*, 2019).

Alternativamente, a quantização de variáveis tem sido extensivamente explorada, com resultados considerados satisfatórios (Alfonso *et al.*, 2013; Keum; Coulibaly, 2017b; Foroozand; Weijs, 2021; de Pádua *et al.*, 2019). A quantização proposta por Alfonso *et al.* (2013) parte da aplicação de uma função matemática de base que transforma um conjunto de valores contínuos em um conjunto discreto. O valor resultante é então aproximado a seu menor múltiplo inteiro mais próximo (Equação 20).

$$x_q = a \left\lceil \frac{2x+a}{2a} \right\rceil \quad (20)$$

Onde,

x_q : valor discreto quantizado da variável x

$\lceil . \rceil$: função de base,

a : constante

A função de base não requer uma distribuição paramétrica e permite considerar aspectos físicos para estabelecer um valor para a (Mirauda; Ostoich, 2020). No entanto, determinar um valor apropriado para a é uma tarefa complexa e neste estudo a constante a assumiu o valor do *bin* estimado pelo critério de Scott (1979) (Equação 21), apoiado em sugestões e resultados satisfatórios obtidos em trabalhos anteriores (Mirauda; Ostoich, 2020; Singh, 2013).

$$a = 3.49\sigma(x)N^{-\frac{1}{3}} \quad (21)$$

Onde,

α : tamanho do bin,

σ : desvio padrão da variável aleatória x ,

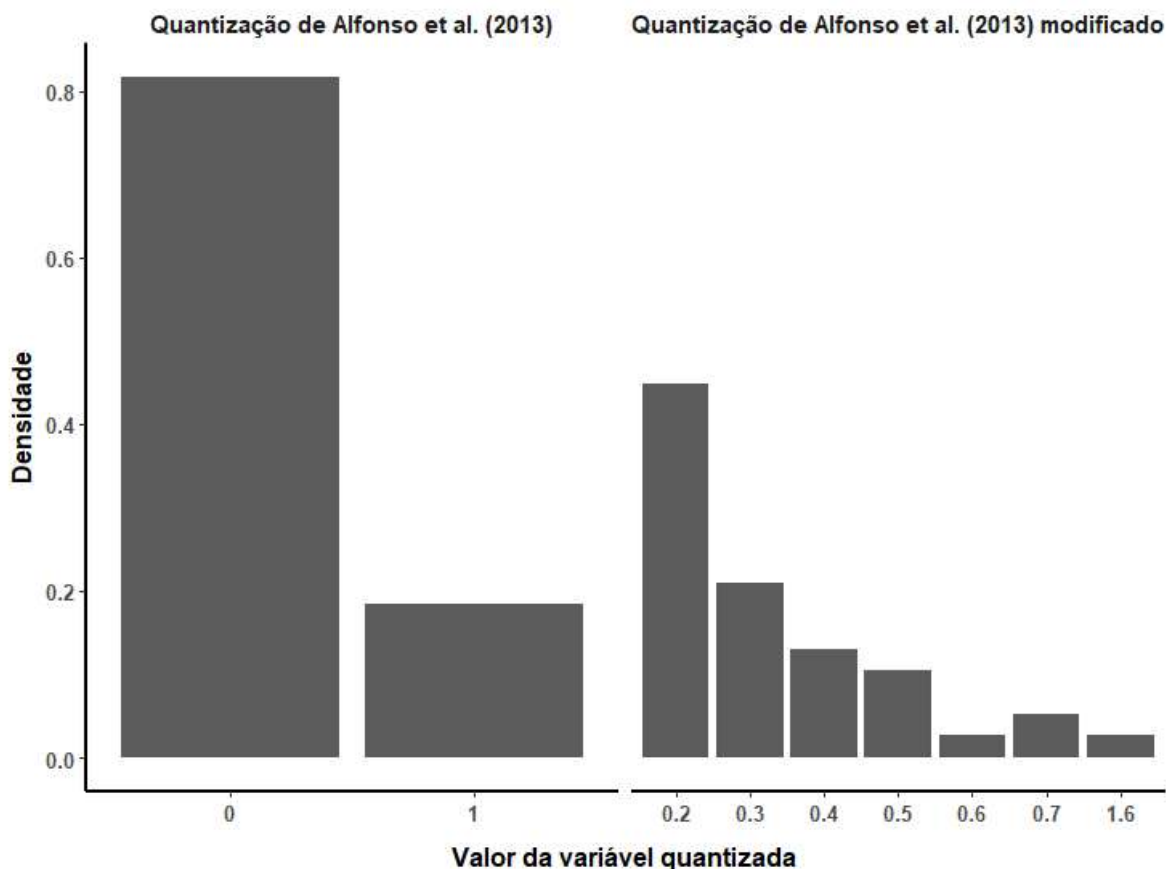
N : tamanho do conjunto de dados da variável aleatória x

No caso do programa PMQQS, muitas variáveis de qualidade da água dizem respeito a elementos-traço e apresentam valores censurados. O uso da Equação 29 proposta por Alfonso *et al.* (2013) distorceria a função densidade de probabilidade (FDP), levando a uma perda de variabilidade, impactando o número de estados discretos que por sua vez, afeta fortemente o cálculo da entropia marginal (Çengel, 2021).

Para superar essa limitação, foi proposto aproximar o valor quantizado de cada observação ao número de casas decimais uma unidade inferior ao nível de sensibilidade do método analítico da variável. Por exemplo, se a sensibilidade analítica na determinação de uma dada variável foi de quatro dígitos decimais, no processo de quantização, o terceiro dígito decimal foi assumido na determinação da variável quantizada. Isso garantiu que valores relativamente próximos representassem o mesmo estado discreto e que valores notavelmente diferentes se tornassem estados distintos da nova variável discreta após a quantização.

Para fins de demonstração, ambos os critérios foram aplicados na quantização da variável Ferro II (mg/L) medida na estação Governador Valadares - 07 (Figura 8). Na figura estão representados o histograma da distribuição dos valores de Fe quantizados pelo método de Alfonso *et al.* (2013) e segundo a modificação proposta nessa pesquisa, esquerda e direita respectivamente.

Figura 8 - Comparação dos critérios de quantização das variáveis de qualidade da água



Em ambos os critérios ocorre generalização da informação original, decorrente da discretização das variáveis, que tende a mascarar as diferenças menos expressivas entre os valores da variável. No entanto, no caso do critério de Alfonso esse processo é muito mais pronunciado e apenas dois estados discretos foram definidos. Isso ocorre porque o critério de aproximação da variável quantizada ao menor número inteiro mais próximo não permite capturar as pequenas variações no valor da variável. No caso do critério de Alfonso modificado, houve maior número de estados discretos, e por extensão, uma suavização menos abrupta levando a menor perda do comportamento distributivo da variável original, fato ilustrado pelo maior número de classes e suas respectivas frequências no histograma.

6.1.1.4 Determinação da entropia máxima usando simulação de Monte Carlo

Para obter a entropia máxima de cada variável, considerou-se a distribuição uniforme como aquela que maximiza a entropia marginal. A distribuição uniforme foi selecionada porque: (i) as variáveis utilizadas foram previamente discretizadas por meio de um método de quantização, (ii) adotou-se um domínio para as simulações $[a,b]$ com base no intervalo original dos dados e, (iii) nenhuma suposição foi feita sobre o comportamento da média e variância dos dados, já que as séries temporais consideradas são relativamente curtas.

A simulação de Monte Carlo (SMC) foi aplicada para obter a entropia máxima através do gerador de números pseudoaleatórios *Mersenne Twister* (MT) (Matsumoto; Nishimura, 1998), que foi desenvolvido no final da década de 1990 para uso em SMC. O MT gera uma sequência pseudoaleatória estatisticamente equidistribuída de alta qualidade e muito eficiente que o credencia como uma excelente alternativa para grandes simulações estatísticas (Mazzotti, 2009).

Para cada variável de qualidade da água, foram simulados 100 diferentes conjuntos de dados sintéticos (*Ensemble*) derivados da distribuição uniforme com domínio definido pelo mesmo intervalo dos dados $[a,b]$ (suporte fixo) e mesmo tamanho da variável original, a fim de preservar as mesmas características das amostras. Assim, considerando o número de variáveis e estações da rede do PMQQS investigada nesta pesquisa, foram realizadas 304.200 simulações.

Cada conjunto de dados sintético foi quantizado usando os mesmos critérios estabelecidos para a quantização das variáveis originais e, posteriormente, sua entropia marginal foi calculada. Cada entropia marginal calculada nesta etapa representa uma das muitas possibilidades de entropia máxima da variável original que foi estimada através da Equação 22.

$$H_m(X) = P90[H(X_q)] \quad (22)$$

Onde,

$H_m(X)$: entropia máxima da variável X ,

$H(X_q)$: entropias marginais das variáveis sintéticas quantizadas,

$P90[H(X_q)]$: percentil 90 das entropias marginais das variáveis sintéticas quantizadas.

O pacote nativo *Stats* do software R foi usado para as simulações de Monte Carlo (R Core Team, 2021).

6.1.1.5 Classificação de “variáveis não-críticas” de qualidade da água

As variáveis que apresentam redundância informacional maior ou igual a um determinado limiar foram classificadas como “*variáveis de baixo conteúdo de informação*” (VBCI). Em seguida, para cada sub-região de qualidade da água estabelecida de acordo com a análise de *cluster*, a representatividade relativa das VBCI (Rep_{VBCI}) foi calculada por meio da Equação 23.

$$Rep_{VBCI} = \left(\frac{N_{VBCI}}{N} \right) * 100 \quad (23)$$

Onde,

Rep_{VBCI} : representatividade relativa das VBCI em cada sub-região e período sazonal,

N_{VBCI} : número de estações onde a variável foi classificada como VBCI,

N : número total de estações na sub-região.

O limiar para classificar as variáveis como VBCI tomou a redundância informacional de todas as variáveis de qualidade da água em ambos os períodos sazonais ao longo da bacia investigada. O limiar foi definido pela análise gráfica do comportamento distributivo da redundância informacional, assumindo que o último patamar antes da redundância de valor 100, em que aumento do percentil não implica mais alterações significativas do valor da redundância, representa o valor ideal para esse critério.

Variáveis classificadas como VBCI em todas as estações de monitoramento de uma sub-região ($Rep_{VBCI} = 100$), em ambos os períodos sazonais, foram confrontadas com

os limites máximos admissíveis para águas doces classe 2 estabelecidos pela resolução nº 357/2005 do Conselho Nacional do Meio Ambiente (BRASIL, 2005). Para essa etapa de consistência, a série VBCI incluiu até mesmo medidas identificadas como *outliers*.

As VBCI em que todas as medidas não ultrapassaram os limites da legislação foram classificadas como “variáveis não-críticas”, passíveis de redução na sua frequência de monitoramento. As variáveis monitoradas na forma dissolvida sem padrões ambientais estabelecidos foram analisadas tomando como referência os limites estabelecidos para a sua concentração total.

6.1.1.6 Validação do método para seleção de variáveis de qualidade da água “não-críticas”

ACP foi utilizada para validar as “variáveis não-críticas” de qualidade da água. Espera-se que, devido ao seu baixo conteúdo de informação, essas variáveis apresentem pequena contribuição as duas primeiros componentes, mostrando que não são determinantes para as características da qualidade da água em cada sub-região.

A ACP é amplamente utilizada no tratamento de dados multivariados e isso se deve em grande parte à sua capacidade de extrair a maior parte das informações contidas em um conjunto de variáveis originais em um conjunto menor de novas variáveis ortogonais chamadas componentes principais. As componentes, embora sejam combinações lineares das variáveis originais, são independentes e não correlacionadas. Além disso, a soma de suas variâncias é igual à das variáveis originais (Abdi; Williams, 2010; Olsen; Chappell; Loftis, 2012).

Os dados utilizados na ACP foram os mesmos descritos na seção 6.1.1. O conjunto de dados foi padronizado para eliminar distorções decorrentes das diferentes escalas de medição das variáveis de qualidade da água. No entanto, visto que dados faltantes não podem ser padronizados, foi necessário eliminar os campos da base de dados que se enquadraram nessa situação.

Quanto à correlação entre as variáveis, foi eliminada uma variável de cada par cujo coeficiente de correlação de Pearson apresentou valor maior que 0,8. Esse

procedimento visou reduzir variáveis redundantes, otimizar o conjunto de variáveis analisadas (Hair Jr *et al.*, 2006) e evitar distorções nos resultados que pudessem atribuir maior importância a variáveis multicolineares. Os *outliers* não foram eliminados para preservar ao máximo as informações contidas no conjunto de dados e auxiliar na interpretação dos resultados.

No desenvolvimento de todas as etapas metodológicas, os modelos e funções foram implementados na linguagem de programação R utilizando pacotes *entropy*, *FactoMineR* e *MultivariateAnalysis* (Hausser; Strimmer, 2021; Le; Josse; Husson, 2008; Azevedo, 2021).

6.1.2 Seleção da melhor configuração espacial de estações de monitoramento da qualidade da água.

6.1.2.1 EWQI das subregiões da bacia do Rio Doce

Embora o monitoramento da qualidade da água seja fundamental ao gerenciamento hídrico, a extração de informações sumarizadas a partir dos dados de monitoramento não é tarefa fácil, pois as redes em geral monitoram conjuntos de variáveis de diferentes naturezas e escalas de medida.

Em resposta a essa problemática uma abordagem que tem sido incorporada ao redimensionamento das redes é o uso de índices de qualidade da água (de Almeida *et al.*, 2022; Salman *et al.*, 2021; Goudarzi *et al.*, 2022).

Alguns dos índices de qualidade (IQA) intensivamente aplicados em estudos de qualidade da água são: Malaysian index, Scottish research development department index (SRDD), Indian pollutant index, NSF index, EQ index, Smith Index, Almeida Index, Iranian Water Quality Index, Water Quality Index (WQI), dentre outros (Sutadian *et al.*, 2016; Uddin; Nash; Olbert, 2021; Manna; Biswas, 2023).

Os índices baseiam-se no processo de agregação da informação a partir do uso de funções. Eles devem representar uma síntese do estado da qualidade da água preservando a variabilidade espaço-temporal que caracteriza as variáveis monitoradas. Uma das principais vantagens do uso de índices é a sua capacidade de

converter dados complexos em uma linguagem inteligível mesmo para as pessoas que não possuem conhecimento técnico específico (Uddin; Nash; Olbert, 2021).

IQA são muito populares nas organizações voltadas ao gerenciamento de recursos hídricos, contudo, não há um único IQA universalmente aceito. Na verdade todos possuem aspectos positivos e limitações particulares (Sutadian *et al.*, 2016).

Os IQA partem de transformações das variáveis individualmente com suas próprias unidades, para uma mesma escala. Normalização ou padronização são procedimentos frequentes nesse sentido. Após transformados os valores são chamados subíndices, e posteriormente são agregados para então se definir o índice, propriamente dito. Por fim, são estabelecidas as classificações com níveis de qualidade (Sutadian *et al.*, 2016; Dash; Kalamdhad, 2021; Uddin; Nash; Olbert, 2021).

Apesar de haver vantagens no uso de índices como dados de entrada para o redimensionamento de redes, podem ocorrer entraves na sua aplicação em virtude de fatores específicos nas seguintes situações:

- a) A variável componente do índice não está incluída no programa de monitoramento e o índice é rígido não permitindo sua utilização na ausência de qualquer das variáveis preconizadas;
- b) O índice possui pesos iguais ou pré-determinados para todas as variáveis não permitindo ajustes para as características hidroambientais específicas da região em que se pretendia utilizá-lo;
- c) O conjunto de variáveis de qualidade da água monitoradas é muito maior que o número total de variáveis pré-estabelecidas pelo índice, de forma que embora sua utilização seja viável, grande parte da informação da qualidade da água disponível não poderá ser utilizada, podendo inclusive que variáveis relacionadas a poluentes que apresentam risco aos ecossistemas e aos usos da água não sejam incluídas no índice, levando assim a propostas para novas configurações de rede inconsistentes.

Como resposta a essas limitações, novos modelos de índices que contemplem mecanismos de contorno a algumas de suas características indesejáveis, têm sido

propostos. Nessa nova geração de índices alternativos aos convencionais encontra-se o índice de qualidade da água ponderado por entropia (EWQI).

Um dos fatores que tem motivado a sua utilização é que ele não baseia a atribuição de pesos às variáveis em percepções pessoais, ainda que relativa à consulta de técnicos e especialistas, fato que muitas vezes pode conduzir precocemente à avaliação da qualidade da água em um sentido e desconsiderar as características locais (Singh *et al.*, 2019a; Singh *et al.*, 2019b; Amiri; Rezaei; Sohrabi, 2014; Fagbote; Olanipekun; Uyi, 2014; Unigwe; Egbueri, 2023).

A determinação do EWQI em um conjunto de dados de qualidade da água deve ocorrer em várias etapas sequenciais (Li; Qian; Wu, 2010).

- A primeira etapa envolve a construção de uma matriz de dados relativos às variáveis de qualidade da água monitoradas.
- A segunda etapa envolve a construção de uma matriz contendo os valores normalizados de cada variável para eliminar erros causados por diferentes dimensões e unidades.
- A terceira etapa envolve o cálculo da entropia de Shannon de cada variável (Shannon, 1948):

$$H = - \left(\frac{1}{\log n} \right) \sum_{i=1}^m V_{ij} \log V_{ij} \quad (24)$$

Onde,

n : número de locais de amostragem

V_{ij} : probabilidade de ocorrência do valor normalizado da variável j avaliada na amostra ith , como segue:

$$V_{ij} = \frac{v_{ij}}{\sum v_{ij}} \quad (25)$$

- A quarta etapa envolve o cálculo dos pesos de entropia (W) de modo que os parâmetros com menor entropia ou desordem recebem mais peso da seguinte maneira:

$$W_j = \frac{(1-H)}{\sum_{j=1}^t (1-H)} \quad (26)$$

- Por último, a agregação de pesos de entropia e uma escala de classificação de qualidade resulta no EWQI e é expressa da seguinte forma:

$$EWQI = \sum_{j=1}^n W_j U_j \quad (27)$$

Onde, U_j para cada parâmetro é dado como a razão do valor monitorado do j th parâmetro (I_j) pelo seu valor normalizado (S_j):

$$U_j = \left(\frac{I_j}{S_j} \right) * 100 \quad (28)$$

A implementação do índice EWQI foi realizada seguindo as etapas propostas por (Li; Qian; Wu, 2010). Na determinação dos pesos de entropia os limites admissíveis tomados como referência compreenderam aqueles definidos pela resolução CONAMA 357/2005 (BRASIL, 2005) e pela Deliberação Normativa Conjunta Copam/CERH-MG nº 01/2008 (COPAM-CERH, 2022).

Para as variáveis monitoradas e analisadas na forma dissolvida, em que não há limite admissível previsto nas normativas, foram utilizados os valores de referência previstos

para a forma total. A segmentação da base de dados inicial, por sub-região, foi utilizada na criação de índices de qualidade da água regionalizados.

Os EWQI regionais foram utilizados no cálculo de EWQI de cada campanha de monitoramento realizada, considerando os dados de cada estação integrante de sub-região. Também para efeito de discussão remeteu-se o valor dos índices a sua respectiva classe de qualidade da água proposta, de acordo com os critérios de Li, Qian e Wu (2010) (Tabela 5).

Tabela 5 - Classes de qualidade da água	
EWQI	Classes de qualidade da água
<50	Excelente
50-100	Boa
100-150	Média
150-200	Ruim
>200	Extremamente ruim
Fonte: Adaptado de Li, Qian e Wu (2010)	

Adicionalmente as séries temporais de EWQI foram analisadas pelo uso de estatística descritiva e exploratória, além de gerar gráficos de séries temporais úteis na interpretação de alguns resultados. A normalidade do comportamento distributivo das séries temporais de EWQI foi avaliada pelo teste Shapiro-Wilk a 5% de significância estatística (Maity, 2018).

6.1.2.2 Discretização de EWQI

O uso da entropia de Shannon nos procedimentos de otimização demandou a discretização das séries temporais do EWQI. Na tentativa de manter os mesmos princípios adotados na seleção de “variáveis não-críticas” foram aplicados os critérios de quantização de Alfonso *et al.* (2013) e Alfonso modificado. Contudo, os resultados não atestaram a viabilidade desse processo, pois as variáveis quantizadas demonstraram tendência a maximizar exageradamente todas as formas de entropia, não alcançando a generalização desejada.

A busca de uma solução a esse entrave levou à decisão de se utilizar critérios de discretização baseados no método do histograma. Para contornar os efeitos de escolhas arbitrárias para o tamanho dos *bins* (intervalos de classes), que podem ter impacto bastante significativo na determinação das diferentes formas de entropia e,

por consequência, na configuração ideal de estações em uma rede de monitoramento, optou-se por realizar uma análise de sensibilidade modificando o critério proposto por de Pádua *et al.* (2019).

de Pádua *et al.* (2019) determinaram a entropia marginal de variáveis hidrométricas para diferentes valores da constante a , utilizada na quantização de variáveis aleatórias contínuas por meio de uma função de base. Para os autores, ao analisar o comportamento gráfico dessa relação (entropia vs valor para a) o valor ideal para a compreende aquele em que mantendo-se a variação de a a entropia correspondente já não sofre impacto significativo.

O mesmo princípio foi implementado nessa pesquisa, entretanto a foi substituído por larguras (classes do histograma) e a entropia utilizada na análise de sensibilidade foi a entropia conjunta de cada sub-região. O valor inicial dos *bins* foi selecionado tomando o menor valor na comparação entre os critérios de Sturges e Scott (Sturges, 1926; Scott, 1979), com aumento progressivo fixado arbitrariamente em saltos de 10 unidades, considerando que na maioria dos casos o máximo EWQI ultrapassou 1000 unidades.

Definido o tamanho ideal dos *bins*, foi procedida a discretização das variáveis.

6.1.2.3 Problema de otimização por múltiplas funções monoobjetivo hierarquizadas

O problema de otimização proposto teve por finalidade determinar o conjunto ideal de estações de monitoramento de cada sub-região de qualidade da água do Rio Doce, através da maximização da entropia conjunta e transinformação, (Shannon, 1948), e minimização do número de estações por meio de funções monoobjetivo hierarquizadas.

A escolha dessa abordagem se deu por dois fatores. O primeiro deles deve-se ao fato de que a segmentação da bacia, através da determinação de subregiões de qualidade da água, reduziu consideravelmente o número total de estações incluídos em cada sub-região e, por consequência, o número de cenários potencialmente candidatos a melhor configuração de rede. De fato, a sub-região com maior número de estações as possui em número de 6.

Esse procedimento viabilizou a implementação da estratégia de busca exaustiva para solução ao problema multiobjetivo proposto. A principal vantagem dessa estratégia é varrer todo o espectro possível de soluções aos problemas de otimização. Por outro lado, a sua maior limitação na otimização de redes de monitoramento é que cada nova estação que compõe a rede produz aumento exponencial no número de cenários potenciais, de forma que, redes com elevado número de estações tornam intratável o problema de otimização por essa abordagem, visto o enorme esforço computacional e tempo demandado em sua implementação (Fernandes *et al.*, 2023; Hlavác, 2000).

De fato, nessas circunstâncias tratar problemas de otimização triviais com algoritmos inspirados na natureza seria desnecessário, pois devido à baixa complexidade resultante dos recortes regionais adotados outros métodos clássicos de otimização são suficientes a essa finalidade. Portanto, considerando esses aspectos o uso do algoritmo de busca exaustiva demonstrou ser uma alternativa de otimização viável.

O segundo aspecto é que ao investigar as principais formas de entropia que compõe funções multiobjetivo na otimização de redes de monitoramento de qualidade da água Foroozand e Weijs (2021) concluíram que o uso da correlação total na otimização das redes é desnecessário, visto que o impacto total da redundância já é contabilizado na entropia conjunta. Além disso os autores afirmam que funções monoobjetivo baseadas apenas na entropia conjunta são suficientes para gerar resultados satisfatórios na otimização das redes, visto que a comparação realizada com outros métodos multiobjetivo não demonstraram diferenças estatisticamente significativas.

Dessa forma, considerando que abordagens mais simples e parcimoniosas devem ter prioridade na solução dos problemas de engenharia, apoiando-se na proposição dos autores, optou-se por utilizar uma estratégia de otimização tomando funções monoobjetivo hierarquizadas (método FMOHBE), assim como não incluir a correlação total no problema de otimização.

As funções monoobjetivo propostas tomaram como princípio a ideia de que, as diferentes medidas do conteúdo de informação e o quantitativo de estações não possuem idealmente o mesmo peso na configuração ótima das redes de monitoramento, mas que há hierarquia no grau de importância dos critérios. Nessa perspectiva foi assumido que o conteúdo de informação conjunto tem prioridade sobre

o conteúdo de informação transferida (transinformação) e sobre o número de estações e que, o potencial de transferência da informação se superpõe ao objetivo de minimizar o número de estações na rede (Equação 29).

$$\begin{array}{l}
 \text{Max: } H(X_{S1}, X_{S2}, \dots, X_{Sk}) \\
 \downarrow \\
 \text{Max: } T(X_{S1}:X_{Sk}, \dots, X_{F1}:X_{Fm}) \\
 \downarrow \\
 \text{Min : } N
 \end{array} \tag{29}$$

Onde,

$\text{Max: } H(X_{S1}, X_{S2}, \dots, X_{Sk})$: maximização da entropia conjunta para um dado conjunto de estações selecionado;

$\text{Max: } T(X_{S1}:X_{Sk}, \dots, X_{F1}:X_{Fm})$: maximização da transinformação entre um conjunto de estações selecionado e o conjunto remanescente de estações;

$\text{Min : } N$: minimização do número de estações

Essa abordagem poderia ainda incorporar o uso de pesos e condensar as três funções objetivo propostas em uma única função, contudo, não há um único critério a esse fim e todas as diversas abordagens que poderiam ser implementadas ou são arbitrárias, imputando ao problema mais um conteúdo de subjetividade não desejado, ou implicaria em outro problema de otimização adicional de investigação da combinação entre os pesos que maximiza o desempenho da solução Pareto-ótima decorrente da implementação do método (Singh, 2013; Castelletti; Pianosi; Restelli, 2013; Yuan *et al.*, 2016).

6.1.2.4 Métodos de otimização aplicados

O método proposto para otimização de funções monoobjetivo hierarquizadas por meio de busca exaustiva, FMOHBE, pode também ser chamado de otimização em cascata, visto sua lógica de operação. Além das funções objetivo demonstradas assumiu-se como restrição que a configuração ótima de rede não deve possuir entropia conjunta inferior a 90% da entropia do conjunto completo de estações.

Alternativamente, nos casos em que eventualmente só existem duas estações, o método tomou apenas a primeira função objetivo nos procedimentos de otimização, a entropia conjunta foi substituída pela entropia marginal e foi aplicado somente o método FMOHBE.

Inicialmente a otimização busca a solução que produz o maior conteúdo total de informação, caso o algoritmo encontre uma solução ótima global o método é interrompido, caso não haja essa solução, do subconjunto de soluções mais promissoras é analisada a transinformação, como critério secundário. Caso exista uma solução de maior potencial de transferência da informação então o método é interrompido e é apresentada a solução ideal para rede. Contudo, se essa etapa igualmente não retornar o ótimo global o algoritmo gera o subconjunto das melhores configurações e segue buscando a solução que minimiza o número de estações.

Pode haver necessidade de se utilizar critérios adicionais na investigação da configuração ideal das redes. Contudo, os critérios adicionais de análise deverão considerar apenas os cenários mais promissores à configuração ótima da rede, visto a filtragem realizada pelos métodos de otimização.

A Figura 7 descreve o método de otimização FMOHBE. A entropia conjunta foi determinada conforme proposto por Singh (2013) (seção 4.5.2). A transinformação compreendeu a informação transferida pelo conjunto de estações de cada cenário analisado, para conjuntos de estações correspondentes não incluídas nesses mesmos cenários, seguindo os mesmo procedimentos de Singh (2013) já mencionados.

De forma, a comparar o desempenho de otimização hierarquizado foi implementada alternativamente uma estratégia baseada no método guloso aditivo, tomando uma única função objetivo que visa maximizar a entropia conjunta (método MGAMO) (Hlavác, 2000; Foroozand; Weijs, 2021).

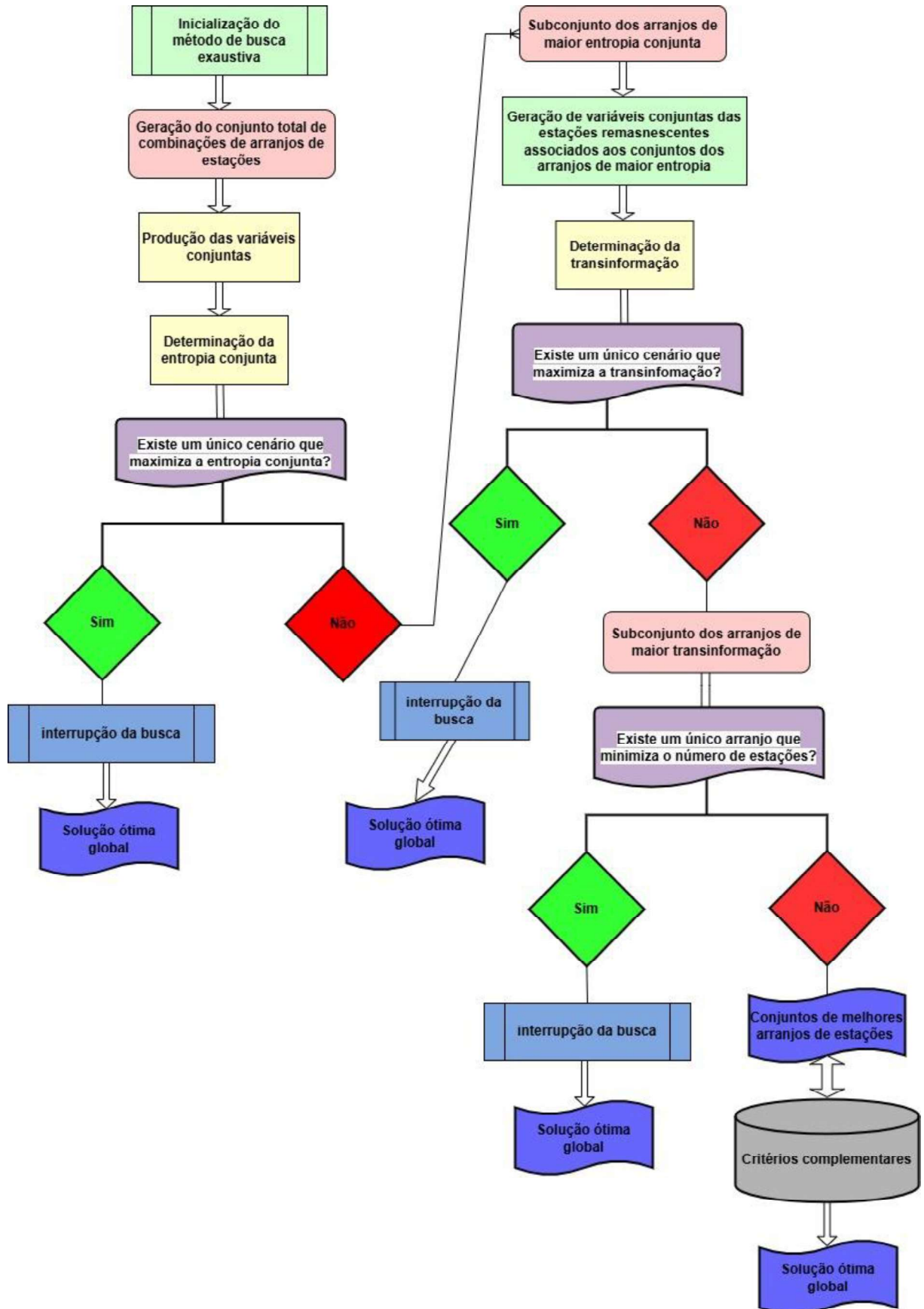
Contudo, ao invés de iniciar o algoritmo a partir de uma estação central de maior entropia marginal, optou-se por testar uma abordagem considerando a ordem das estações de montante à jusante, ou seja, a estação mais à montante foi escolhida

como principal e a cada cenário foi adicionada a estação mais próxima, no sentido montante à jusante.

Complementarmente também foram utilizados teste de hipótese não-paramétricos Mann-Whitney e Kruskal-Wallis na comparação entre séries temporais de EWQI não discretizadas, de estações indicadas como redundantes nos procedimentos de otimização, a fim de avaliar se essas estações são excessivamente semelhantes entre si, o que corroboraria seu elevado grau de redundância que motivou a indicação de remoção da rede (Sperling; Verbyla; Oliveira, 2020).

Os procedimentos foram todos implementados por meio da linguagem de programação *R* e dos pacotes *entropy*, *combinat* e *infotheo* (R Core Team, 2021; Hausser; Strimmer, 2021; Chasalow, 2012; Meyer, 2022).

Figura 9 – Método de otimização baseado no algoritmo de busca exaustiva tomando funções monoobjetivo hierarquizadas



7 RESULTADOS E DISCUSSÃO

7.1 Seleção de “variáveis não-críticas” na bacia do Rio Doce

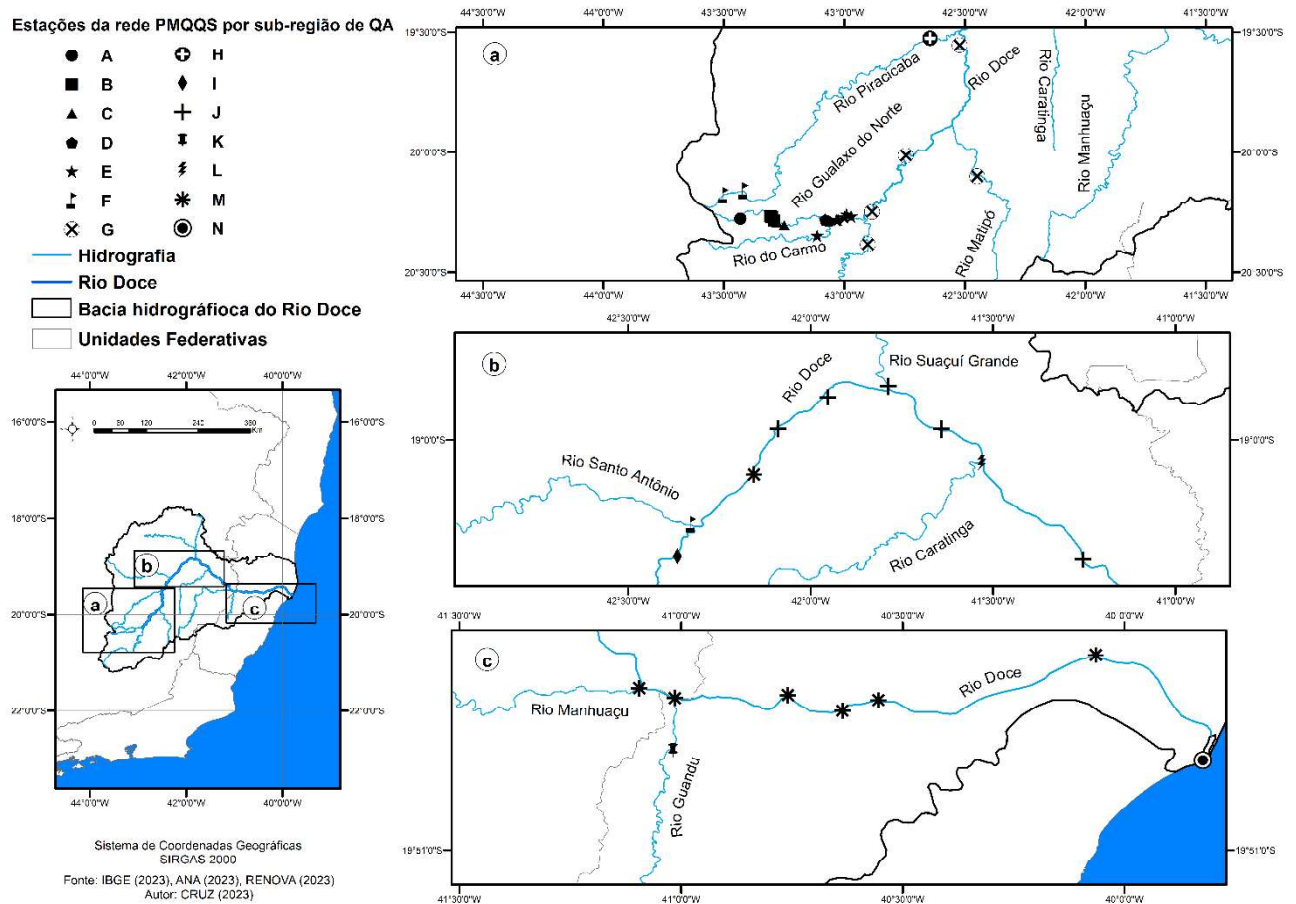
7.1.1 Limites espaciais e variáveis de qualidade da água analisadas

Dentre as 90 variáveis de qualidade da água potencialmente aptas para essa pesquisa (Apêndice - Tabela A2), 78 foram selecionadas e 11 variáveis foram eliminadas porque dados faltantes ou inválidos excederam 30% das medidas (Apêndice - Tabelas A3 e A4). A variável Temperatura ambiente, embora monitorada em todas as estações e campanhas, foi excluída por não constituir estritamente uma variável de qualidade da água. Assumindo todos os registros não faltantes como dados válidos, foram utilizados nesta pesquisa 107.585 valores, correspondendo ao monitoramento mensal de agosto de 2017 a dezembro de 2020.

A análise de *cluster* pelo algoritmo de Ward resultou na segmentação das 39 estações de monitoramento em 14 grupos principais, considerados subregiões de qualidade de água na bacia hidrográfica (Figuras 10). Os grupos foram bastante heterogêneos em termos de número de estações, que variou de uma a seis. As subregiões foram denominadas de A a N, de montante a jusante na bacia hidrográfica (Figura 11) (Apêndice - Tabela A5) (Apêndice - Figura A1).

A subregião A é formada por uma estação localizada no Rio Gualaxo do Norte, a montante da barragem de Fundão. As subregiões B, C e D são formadas por estações de monitoramento no Rio Gualaxo do Norte, a jusante da barragem de Fundão. A subregião E é formada por estações de monitoramento no Rio do Carmo. As subregiões F, H, K e L são formadas por estações de monitoramento em afluentes do Rio Doce (Rios Piranga, Santo Antônio, Piracicaba, Guandu e Caratinga), enquanto as subregiões G, J e M são formadas por estações de monitoramento localizadas tanto no Rio Doce quanto em seus afluentes (Rios Piranga, Matipó, Suaçuí Grande e Manhauçu). As subregiões I e N possuem uma estação de monitoramento cada, no Rio Doce.

Figura 10 - Grupos de estações de monitoramento da rede PMQQS

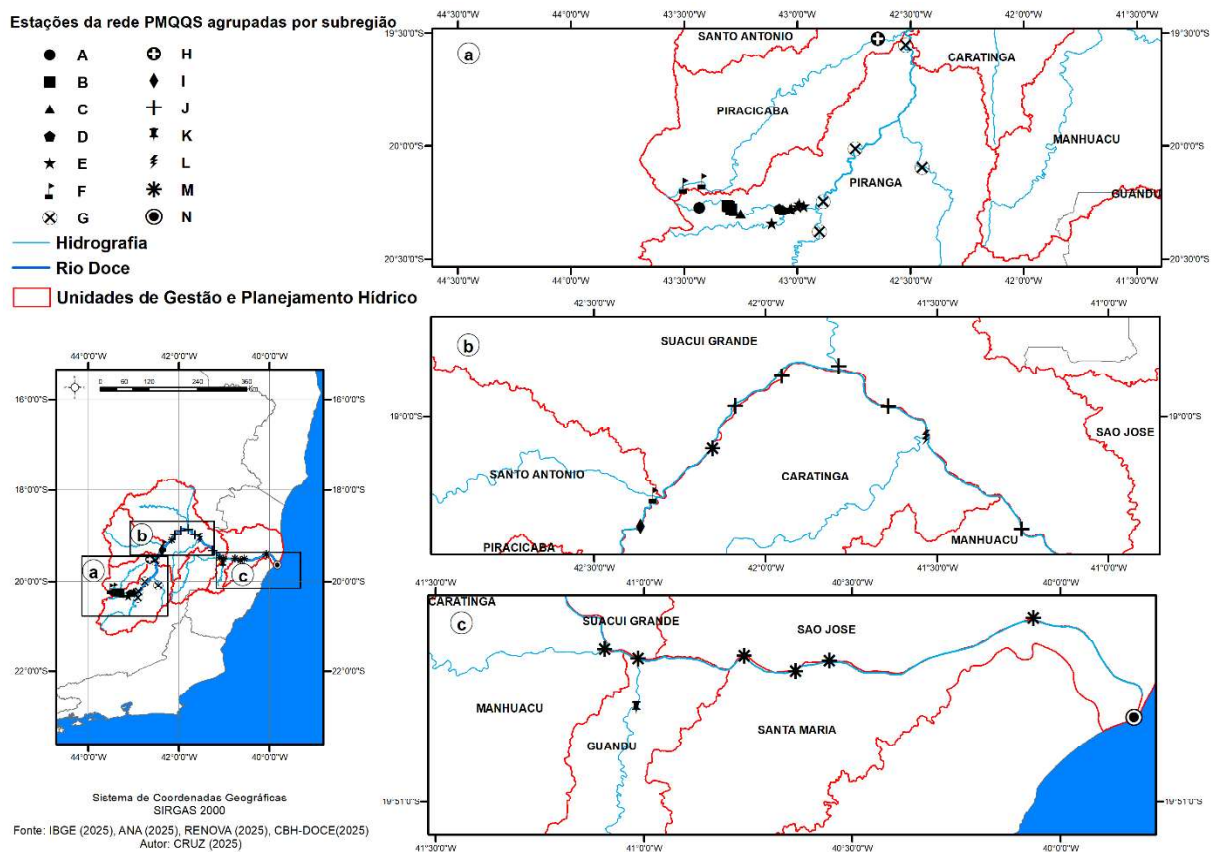


A bacia do Rio Doce adota uma divisão oficial para fins de planejamento e gestão hídrica. Os critérios estabelecidos nessa regionalização não são exclusivamente de ordem física, ou seja, os limites de cada uma dessas unidades não são necessariamente coincidentes com os divisores topográficos das bacias. Embora, esse seja um dos critérios adotados, ele é acrescido de outros critérios socioeconômicos, culturais e políticos (ANA, 2023b).

Ao todo a bacia possui nove unidades de gestão, sendo seis localizadas na porção mineira e três na região capixaba. Em Minas Gerais foram definidas as unidades: Caratinga, Manhuaçu, Piracicaba, Piranga, Santo Antônio e Suaçuí Grande. O estado do Espírito Santo adota as unidades: Guandu, São José e Santa Maria. Essas unidades possuem comitês de bacia e compreendem os perímetros de implementação de programas e projetos definidos pelo Plano Integrado de Recursos Hídricos da Bacia do Rio Doce (ANA, 2023b).

Embora, a regionalização aplicada nessa pesquisa tenha indicado 14 subregiões de qualidade da água, em detrimento das 9 adotadas pelos critérios oficiais de gestão hídrica, é possível destacar que há coerência na distribuição espacial das estações de cada subregião quando analisadas conjuntamente com as unidades oficiais de gestão e planejamento hídrico da bacia do Rio Doce (Figura 11).

Figura 11 – Distribuição de estações de subregiões considerando as unidades oficiais de gestão e planejamento hídrico da bacia do Rio Doce



As estações das subregiões A, B, C, D, E e G estão distribuídas entre os Rios Gualaxo do Norte, Carmo, Piranga e Matipó. Elas situam-se nas regiões das cabeceiras do Rio Doce e dentro do perímetro da unidade Piranga. As estações da subregião J estão todas situadas na região do médio Rio Doce, à exceção de uma estação que se situa no Rio Suaçuí Grande. O Rio Caratinga é coberto por apenas uma estação (subregião L), assim como o Rio Santo Antônio (subregião F). Apesar de não estarem na calha do Rio Doce essas estações estão em regiões próximas à confluência dos Rio Suaçuí Grande, Caratinga e Santo Antônio com o próprio Doce, na região mais ao centro da

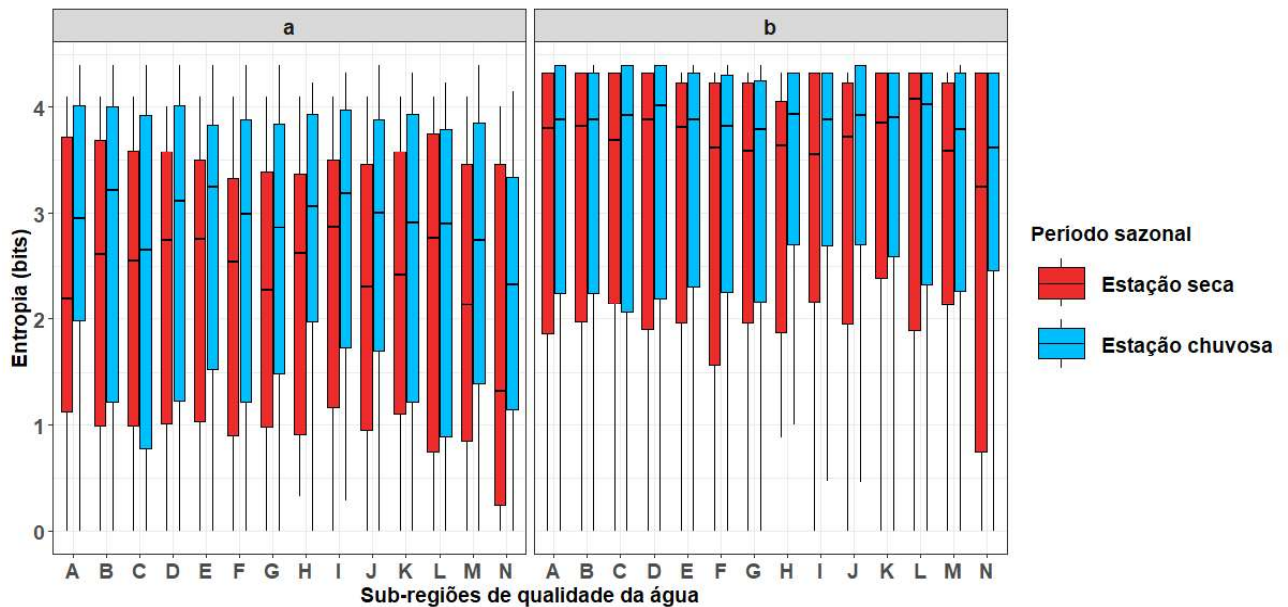
bacia. Conjuntamente essas estações fazem parte das unidades Suaçuí Grande, Caratinga e Santo Antônio.

Já as unidades Guandu, Manhuaçu, Santa Maria e São José, todas situadas na região mais próxima à porção final da bacia, são cobertas pelas estações das subregiões K, M e N, que possuem relativa proximidade entre si. Apesar das incongruências evidenciadas; como o fato de a unidade Piracicaba ser coberta com três estações, mas não integrantes de uma única subregião (subregiões F e H) e da subregião M estar predominantemente no baixo Rio Doce mas possuir uma estação no médio Doce; de maneira geral, pode-se compreender essas subregiões como uma divisão mais detalhada das unidades de planejamento e gestão, como efeito da variabilidade das características físicas, socioambientais e econômicas a nível interno de cada uma das unidades oficiais, que são de fato bem extensas.

7.1.2 Entropia marginal e máxima

A entropia marginal, a entropia máxima e, por extensão, a redundância de informação das variáveis de qualidade da água foram influenciadas pela sazonalidade na bacia hidrográfica do Rio Doce. O valor máximo observado de entropia marginal variou de 4,09 bits durante o período seco a 4,39 bits durante o período chuvoso, correspondendo a um incremento de 7%. Observou-se ganho no conteúdo de informação durante a estação chuvosa em todas as subregiões de qualidade da água, o que foi destacado pelo deslocamento dos patamares relativos aos percentis 50 e 75 de entropia marginal para níveis associados a entropias mais expressivas, quando comparados à estação seca (Figura 12a). O efeito da sazonalidade sobre a entropia de Shannon das variáveis de qualidade da água já havia sido demonstrado em estudos anteriores em recursos hídricos, igualmente no sentido de tornar as variáveis de qualidade da água mais informativas durante a estação chuvosa (Tanos *et al.*, 2015).

Figura 12 - Entropia marginal (a) e entropia máxima (b) por subregiões de qualidade da água durante as estações seca e chuvosa



A entropia parece distribuir-se assimetricamente, com maior dispersão até o percentil 50 e maior densidade na cauda superior da distribuição, onde obviamente se situa o maior conteúdo informativo (Figura 12). A menor entropia marginal obtida (Figura 12a), tanto na estação seca quanto na chuvosa, foi de valor zero ou muito próximo de zero, o que indica incerteza desprezível ou de baixa expressividade ou, em outros termos, conhecimento completo e elevada previsibilidade quanto ao conteúdo de informação a ser transmitido em eventos futuros pela variável aleatória (Çengel, 2021). A variância dos dados foi representada por estados discretos após a quantificação das variáveis. Quanto menor a variância de uma dada variável, menor o número de estados discretos para descrevê-la, podendo atingir um único estado e zero incerteza (entropia nula).

A entropia máxima foi limitada ao maior valor de entropia marginal (Figura 12b). Durante a estação seca, a entropia máxima atingiu 4,09 bits e na estação chuvosa 4,39 bits. Assim como a entropia marginal, a sazonalidade favorece níveis de entropia máxima mais expressivos durante a estação chuvosa. A similaridade entre os valores mais expressivos das entropias marginal e máxima já era esperada, pois ao utilizar a distribuição uniforme com intervalo finito obtido a partir dos dados observados, as características estatísticas das variáveis originais são preservadas nas simulações de Monte Carlo.

Apesar não haver nenhuma restrição teórica quanto ao limite superior do valor das variáveis monitoradas, o que levaria a possibilidade de as simulações realizadas assumirem valores de qualquer ordem de grandeza, no ambiente natural frequentemente há limites físicos para o comportamento das variáveis. Esse processo pode ser exemplificado pelo caso da concentração de saturação do oxigênio dissolvido nas águas de um rio, que é limitado minimamente pelas condições de temperatura e pressão atmosférica (Chapra; Camacho; McBride, 2021).

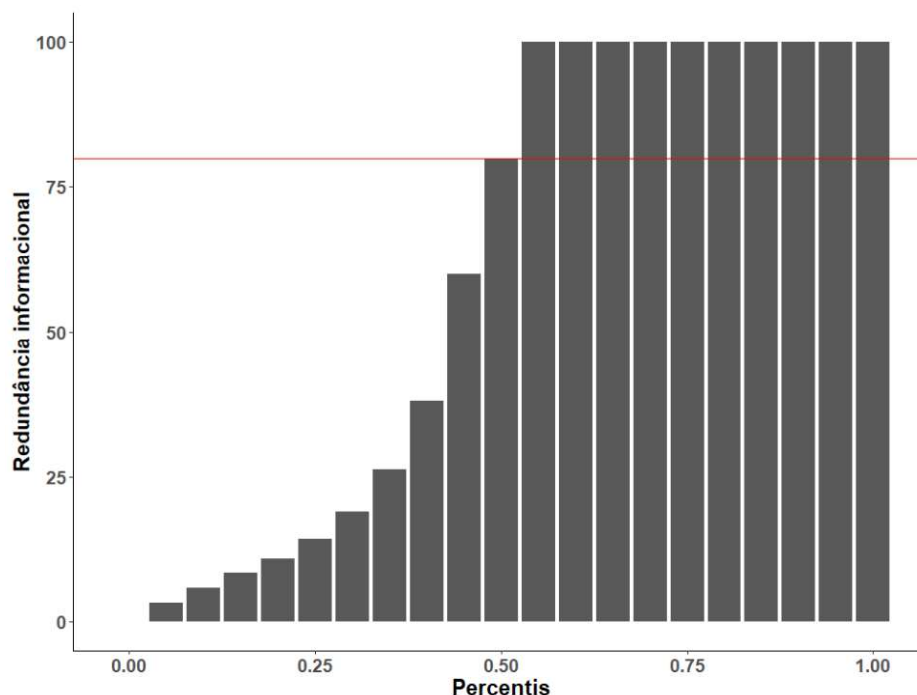
Portanto, o critério adotado de domínio definido para as SMC, por se ater aos limites superior e inferior observados das variáveis de qualidade da água cria condições para que as séries sintéticas possuam coerência com as condições físicas dos sistemas hídricos naturais representados nas séries temporais, minimizando os riscos de os dados das simulações corresponderem a meros exercícios de modelagem computacional.

7.1.3 “Variáveis não-críticas” na bacia do Rio Doce

A distribuição da redundância informacional considerando todas as variáveis de qualidade da água e estações de monitoramento, demonstra que o valor em que alterações nos percentis não impactaram mais o valor da redundância informacional ocorreu próximo ao percentil 50, o que corresponde a redundância informacional da ordem de 80% (Figura 13). Esse valor foi utilizado como limiar para classificar as VBCI.

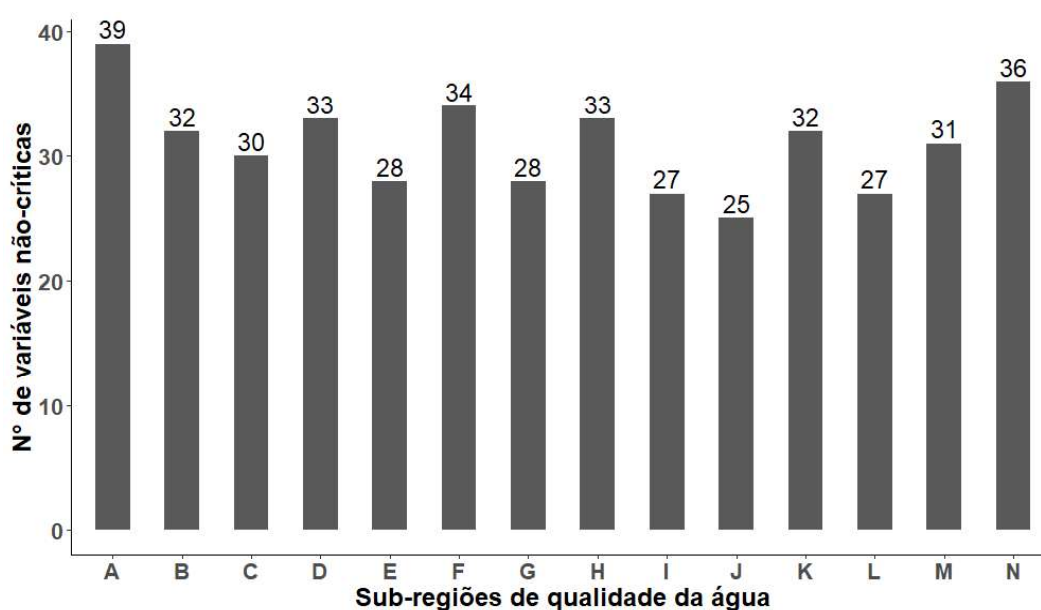
Foram classificadas como VBCI 51 variáveis na estação chuvosa e 62 variáveis na estação seca, ao longo das 14 subregiões de qualidade da água. O subconjunto daquelas que apresentaram redundância informacional em nível superior a 80%, em ambos os períodos sazonais, incluíram o total de 41 variáveis. Essas últimas, como preconizado na metodologia, tiveram cada um de seus dados confrontados com os limites admissíveis definidos pelas normas ambientais brasileiras para avaliação de eventuais não conformidades. Nesses casos, apesar do baixo conteúdo de informação, elas não foram rotuladas como “variáveis não-críticas”.

Figura 13 - Distribuição da redundância informacional. A linha vermelha representa o limiar para classificação de variáveis de qualidade da água como VBCI



A subregião L (no Rio Caratinga) foi responsável pelo maior número de variáveis não conformes, quatro no total; enquanto as subregiões D (no Rio Gualaxo no Norte) e J (no Rio Doce) registraram três variáveis não conformes cada; subregiões B (no Rio Gualaxo do Norte e Rio Suaçuí Grande), C (no Rio Gualaxo no Norte), E (no Rio do Carmo), G (no Rio Doce), M (no Rio Manhuaçu e no Rio Manhuaçu) e N (no Rio Doce) registraram duas variáveis não conformes cada; F (no Rio Santo Antônio), H (no Rio Piracicaba), I (no Rio Doce) e K (no Rio Guandu) apresentaram o menor número (uma variável não conforme cada). Na subregião A (estação RGN-01 no Rio Gualaxo do Norte, a montante da barragem de Fundão), todas as variáveis mantiveram-se dentro dos limites ambientais. O número de “variáveis não-críticas” de qualidade da água variou de 25 a 39, de acordo com a subregião (Figura 14).

Figura 14 - Número de variáveis de qualidade da água “não-críticas” por subregião



Em termos percentuais, há de 32 a 50% de variáveis “não-críticas” sendo monitoradas em cada sub-região. A gravidade dos impactos, especialmente aos recursos hídricos, decorrentes do desastre de Fundão levou aos tomadores de decisão relativos ao planejamento do PMQQS a optarem por monitorar o maior número possível de variáveis, de forma, a cobrir ao máximo os possíveis elementos inseridos ou remobilizados aos sistemas hídricos da bacia.

Os gestores desse programa de monitoramento emergencial de qualidade da água assumiram, portanto, o risco de superdimensionar a rede. Apesar das razões para isso serem plenamente justificáveis, pelo cenário crítico gerado pelo desastre de Fundão, os resultados sugerem haver oportunidades para ajustes no projeto da rede PMQQS.

Ainda em relação a variáveis em não-conformidade com as normas, aquela que mais se destacou foi a Demanda Bioquímica de Oxigênio (DBO), com medidas não conformes em 8 subregiões, Fósforo total em 6 subregiões, Chumbo total em 4 subregiões, Mercúrio total e Fósforo dissolvido em duas subregiões cada e Boro dissolvido, Boro total, Cádmio total e Níquel total, com uma subregião cada (Tabela 6).

Variáveis de qualidade da água	Subregiões de qualidade da água													
	A	B	C	D	E	F	G	H	I	J	K	L	M	N
Ferro II														
Nitrito														
Nitrogênio orgânico														
Polifosfato														
Sólidos sedimentáveis														
Sulfetos como H ₂ S indissociável														
Antimônio total														
Berílio total														
Boro total														
Cádmio total														
Cromo total														
Nitrogênio Kjeldahl total														
Chumbo total														
Mercúrio total														
Molibdênio total														
Níquel total														
Fósforo total														
Selênio total														
Prata total														
Sulfetos totais														
Vanádio total														

As “variáveis não-críticas” são caracterizadas por serem pouco informativas (baixo conteúdo de informação), o que significa pequena incerteza em seus valores esperados a cada nova campanha de monitoramento. A incerteza está ligada ao grau de redundância, neste caso entendido como repetibilidade dos dados. Quando as medidas dessas variáveis são repetidas continuamente, cada novo dado é gerado sob grande previsibilidade, produzindo pouco ou nenhum conteúdo de informação novo.

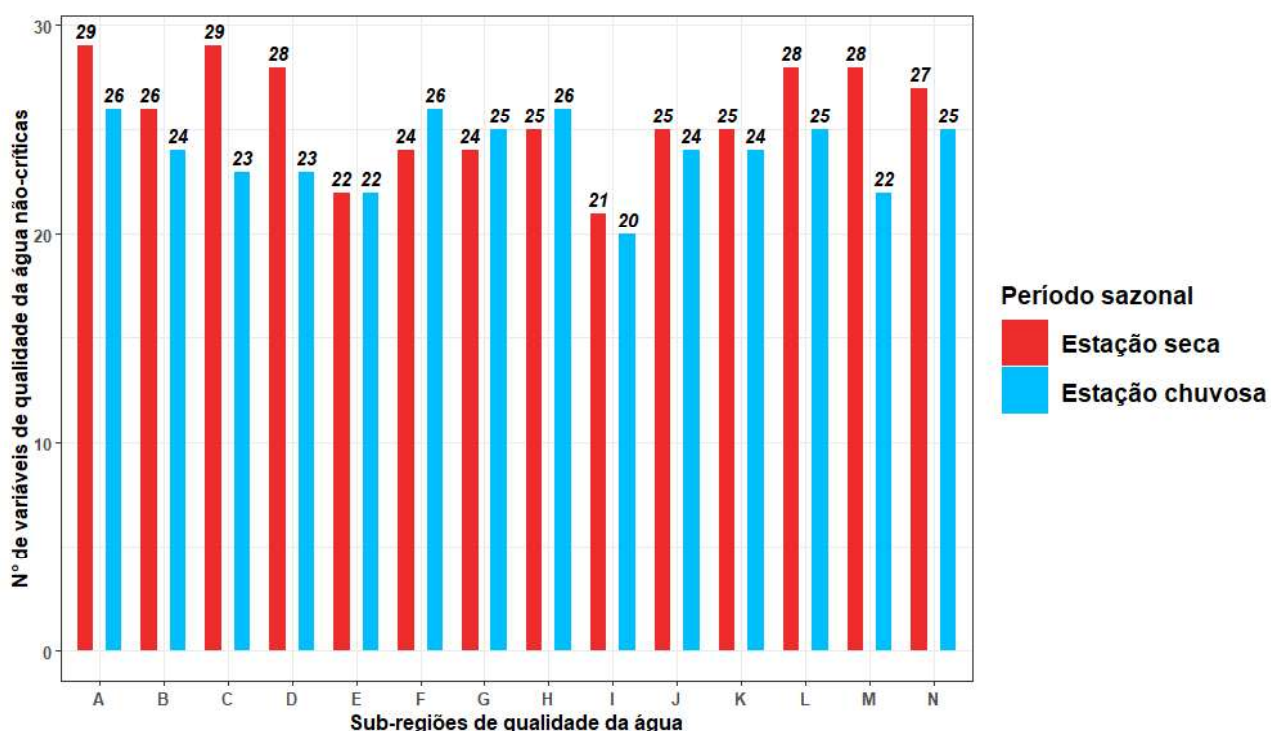
Esse fato, associado ao baixo risco ao meio ambiente evidenciado pelo exercício de confrontar as medidas das variáveis com os limites admissíveis definidos pelas normas ambientais, as torna pouco interessantes para os objetivos das redes emergenciais. No caso da bacia do Rio Doce, as variáveis “não-críticas” selecionadas a cada nova campanha agregam pouca informação útil em termos práticos pois não conseguem capturar efetivamente os efeitos adversos dos poluentes dispersados na bacia, ou por outro lado ilustram um cenário de avanço na recuperação.

Em quaisquer dos casos, o único ponto de consenso é que o método aplicado, ao indicar quais variáveis são “não-críticas” à qualidade da água, permite de fato reorientar os esforços para o *conjunto* de variáveis que por não comporem esse grupo, demandam maior interesse e, portanto, devem permanecer sendo monitoradas no mínimo nos mesmos padrões operacionais vigentes, ao passo que as outras oferecem oportunidade para ajustes no projeto das redes com o mínimo de risco envolvido.

7.1.4 Variabilidade espaço-temporal das “variáveis não-críticas”

Em todas as subregiões, as “variáveis não-críticas”, bem como sua quantidade, variaram entre as estações seca e chuvosa em diferentes proporções, exceto para a subregião E (Figura 15).

Figura 15 - Número de variáveis de qualidade da água “não-críticas” por subregião e período sazonal



Devido a mudanças nos padrões de precipitação e escoamento associado à dinâmica espacial de uso/ocupação do solo em bacias hidrográficas, a sazonalidade frequentemente impacta a quantidade e a qualidade da água (Rodrigues *et al.*, 2018; Schliemann; Grevstad; Brazeau, 2021). A influência da sazonalidade também foi observada por pesquisadores que investigaram a qualidade da água na bacia

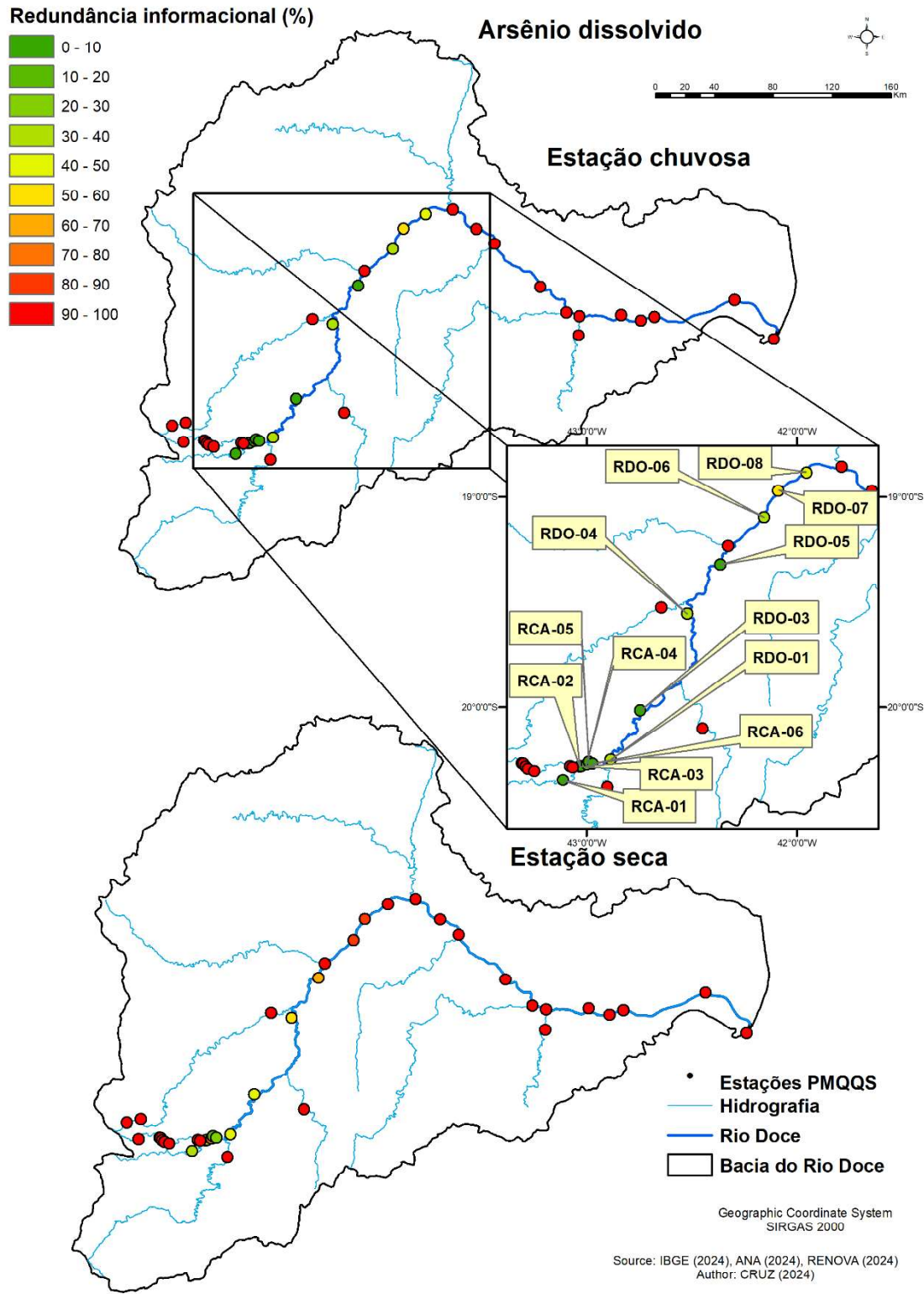
hidrográfica do Rio Doce, antes e depois do desastre do Fundão (Passos *et al.*, 2021; Nogueira *et al.*, 2021; Petrucio *et al.*, 2005).

A estação seca predominou sobre a chuvosa quanto ao número de variáveis “não-críticas” de qualidade da água. A poluição hídrica de fontes pontuais, principalmente de esgotos não tratados e tratados, permanece praticamente constante ao longo do ano, enquanto durante a estação seca, pequenos volumes precipitados e vazões reduzidas na rede fluvial, aumentam o efeito de cargas difusas na bacia hidrográfica e promovem menor erosão e ressuspensão de sedimentos nos rios (Oliveira; Quaresma, 2017; da Cunha Richard *et al.*, 2020). Consequentemente, durante a estação seca grande parte das variáveis de qualidade da água apresenta menor variabilidade (menor incerteza) e menor entropia quando comparadas à estação chuvosa.

As medidas de entropia variaram ao longo da bacia hidrográfica de acordo com a variável e a estação do ano. A título ilustrativo, as Figuras 16 e 17 demonstram a distribuição espacial da redundância informacional, respectivamente nas estações seca e chuvosa, para Arsênio dissolvido e Manganês total. O Manganês total foi considerado uma variável importante da qualidade da água em todas as subregiões, enquanto o Arsênio dissolvido foi identificado como uma “variável não-crítica” em 10 subregiões.

A redundância informacional do Arsênio dissolvido demonstra uma grande dispersão ao longo da bacia, tanto na estação seca quanto na chuvosa, variando de 0 a 100%. Zonas de menor redundância são muito bem delimitadas no trecho superior do Rio Doce com valores abaixo de 40%. As estações de monitoramento incluídas nessa área correspondem a todas as estações das subregiões E, G e I. A sub-região E corresponde à bacia hidrográfica do Rio Carmo, historicamente explorada pela mineração de ouro (Daus *et al.*, 2005; Borba *et al.*, 2000; Silva *et al.*, 2018). Essa atividade, especialmente nos séculos passados, quando a mineração era rudimentar e sem preocupações voltadas aos impactos ambientais, promoveu a liberação de Arsênio do solo e dos minerais, representando um sério risco aos ecossistemas e à saúde humana devido às suas características de toxicidade (Alonso *et al.*, 2020; Barcelos *et al.*, 2020).

Figura 16 - Distribuição espacial da redundância informacional de Arsênio dissolvido nas estações seca e chuvosa na bacia do Rio Doce



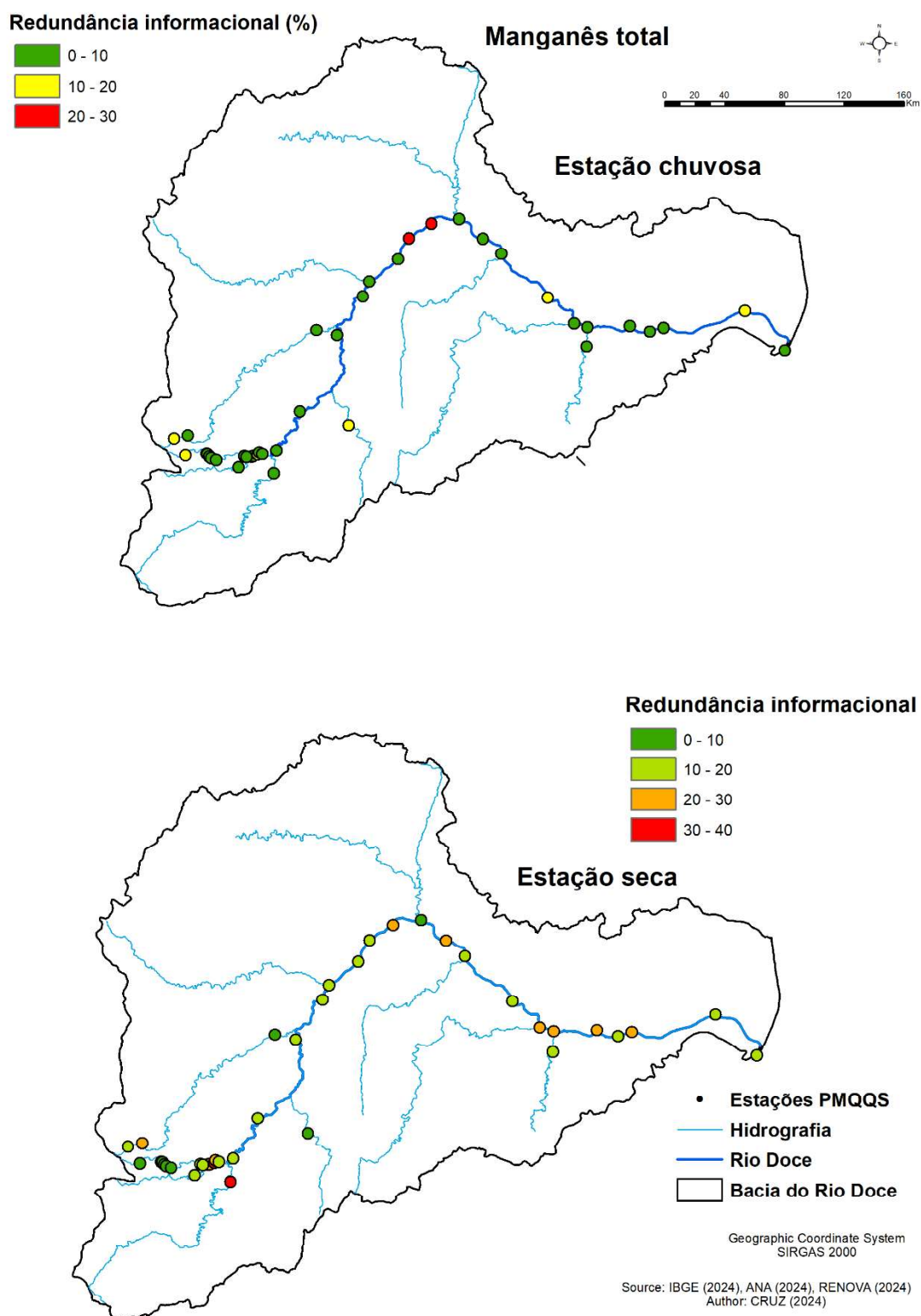
As subregiões G e I estão localizadas no alto Rio Doce, provavelmente influenciadas pela mineração de ouro na bacia hidrográfica do Carmo, uma vez que concentrações de Arsênio dissolvido em desconformidade com as normas ambientais foram registradas antes e depois do desastre de Fundão nesta localidade (IGAM, 2020b). Comparando-se os mapas da Figura 16, a redundância diminui da estação seca para a chuvosa no trecho médio da bacia (subregião J). Em menor frequência, níveis perigosos de Arsênio também foram identificados na água para consumo humano na região do médio do Rio Doce, especialmente em Governador Valadares (Teixeira *et al.*, 2020).

O arsênio dissolvido foi indicado como variável não crítica em todas as subregiões, exceto nas quatro subregiões mencionadas (E, G, I e J).

Os mapas da Figura 17 mostram que, devido à presença e variabilidade do Manganês total em toda a bacia, a distribuição espacial da redundância é bastante homogênea e semelhante entre as estações seca e chuvosa. Baixos níveis de redundância predominam ao longo do ano, variando de 0 a 34,4% na estação seca e de 1,7 a 27,1% durante a estação chuvosa. Esse comportamento reflete uma variabilidade mais expressiva do Manganês total, tanto em termos temporais quanto espaciais, na bacia hidrográfica do Rio Doce.

Estudos recentes relatam medidas de Manganês total em altas concentrações e até mesmo violando padrões ambientais em toda a bacia hidrográfica do Rio Doce (Carvalho *et al.*, 2018; IGAM, 2020b). Embora haja registros de excedências de Manganês no período pré-desastre na bacia do Rio Doce, muitos autores tem atribuído a sua maior dispersão espacial e níveis de concentração atuais ao desastre de Fundão, uma vez que o Manganês é parte da composição dos rejeitos de mineração de Ferro e tende a ser gradualmente remobilizado dos sedimentos depositados nos canais dos rios para as águas superficiais por processos químicos e físico-químicos, representando um risco de médio e longo prazo para a população e os ecossistemas aquáticos (Duarte; Neves, 2023; Duarte *et al.*, 2021; Moreira *et al.*, 2021; Carvalho *et al.*, 2018).

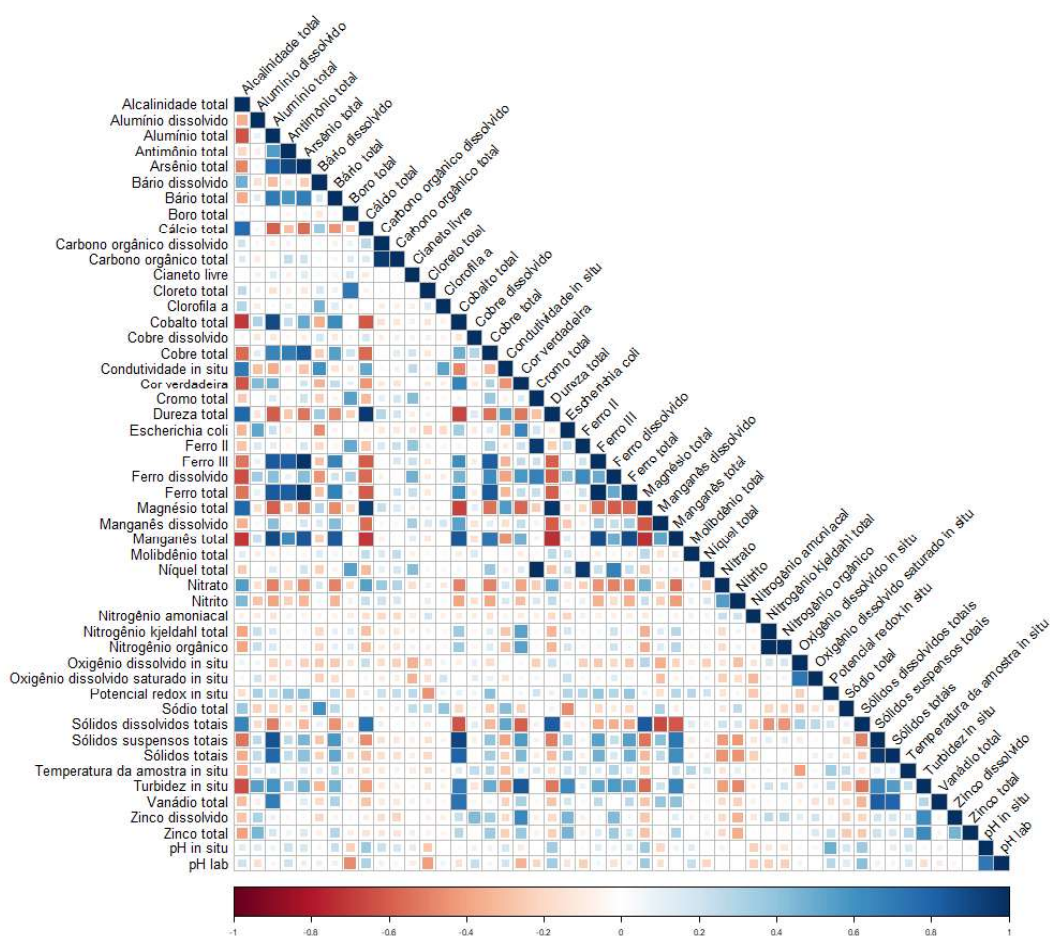
Figura 17 - Distribuição espacial da redundância informacional de Manganês total nas estações seca e chuvosa na bacia do Rio Doce



7.1.5 Validação da seleção de “variáveis não-críticas”

A análise de correlação demonstrou a existência de forte multicolinearidade entre as variáveis de qualidade da água em todas as subregiões. A Figura 18 ilustra a correlação baseada no coeficiente de Pearson entre as variáveis na subregião A (ver Apêndice - Figuras A2 a A14 para outras subregiões).

Figura 18 - Matriz de correlação de variáveis de qualidade da água na subregião A



O conjunto total de variáveis eliminadas, consideradas as correlações aos pares, variou de 9 (subregião F) a 40 variáveis (subregião N) (Apêndice - Tabela A6). As variáveis cuja variabilidade era nula ou desprezível e não puderam ser padronizadas também foram excluídas da análise e, portanto, o conjunto total de variáveis investigadas na ACP variou de 9 a 46 (Apêndice - Tabela A7).

A variância explicada cumulativa das duas primeiras componentes principais variou de 21,5 (subregião E) a 67,6% (subregião N), com um valor médio entre as subregiões de 29,9%. Os autovalores das componentes em todas as subregiões foram superiores a 1, sendo portanto representativas segundo o critério de Kaiser (Ferré, 1995).

Todas as subregiões de qualidade da água incluíram em alguma proporção “variáveis não-críticas” investigadas. A subregião N apresentou apenas uma “variável não-crítica” investigada na ACP, enquanto a subregião M apresentou 11 “variáveis não-críticas” (Tabela 8).

Tabela 8 - “Variáveis não-críticas” de qualidade da água investigadas pela ACP

A	B	C	D	E	F	G
Antimônio total	Arsênio dissolvido	Berílio total	Boro total	Antimônio total	Antimônio total	Antimônio dissolvido
Cianeto livre	Boro total	Boro total	Cromo total	Cobalto dissolvido	Boro total	Boro total
Clorofila a	Cianeto livre	Cianeto livre	Feoftina	DBO	Clorofila a	Clorofila a
Cromo total	Ferro II	Clorofila a	Molibdênio total	Feoftina	Feoftina	Cobalto dissolvido
Molibdênio total	Molibdênio total	Nitrogênio amoniacal	Nitrogênio amoniacal	Nitrogênio amoniacal	Fósforo total	Níquel total
Nitrogênio amoniacal	Nitrogênio amoniacal	Nitrogênio kjeldahl total	Nitrogênio kjeldahl total	Polifosfato	Nitrogênio amoniacal	Nitrogênio amoniacal
Nitrogênio kjeldahl total	Nitrogênio kjeldahl total	Sulfetos totais	Polifosfato	Sólidos sedimentáveis	-	Selênio total
Vanádio total	Sólidos sedimentáveis	-	Sólidos sedimentáveis	-	-	-
-	Sulfetos totais	-	Sulfetos totais	-	-	-
H	I	J	K	L	M	N
Arsênio dissolvido	Boro total	Antimônio dissolvido	Berílio total	Antimônio total	Antimônio total	Chumbo total
Cianeto livre	Chumbo total	Berílio total	Cianeto livre	Cianeto livre	Arsênio dissolvido	-
Clorofila a	Níquel total	Boro total	Clorofila a	Clorofila a	Berílio total	-
DBO	-	Cobalto dissolvido	Nitrogênio amoniacal	Nitrogênio amoniacal	Boro total	-
Feoftina	-	Feoftina	Polifosfato	Polifosfato	Molibdênio total	-
Nitrogênio amoniacal	-	Nitrogênio amoniacal	-	Sulfetos totais	Níquel total	-
-	-	Polifosfato	-	-	Nitrogênio amoniacal	-
-	-	Selênio total	-	-	Nitrogênio kjeldahl total	-
-	-	Sólidos sedimentáveis	-	-	Polifosfato	-
-	-	Sulfetos totais	-	-	Sulfetos totais	-
-	-	-	-	-	Vanádio dissolvido	-

O conjunto de variáveis com maior contribuição para a primeira componente principal (CP1) nas subregiões A, B, C, E, F, H, L, M e N para CP1 não incluiu “variáveis não - críticas”. As variáveis com percentual de contribuição individual superior à média de cada subregião foram consideradas de maior contribuição. As subregiões G, I e K incluíram uma e as subregiões D e J incluíram duas “variáveis não-críticas” nesse rol, todas destacando-se em relação a primeira componente (Figura 19) (Tabela 9).

Figura 19 - Contribuição de variáveis de qualidade da água para CP1 e CP2 (%), na subregião D

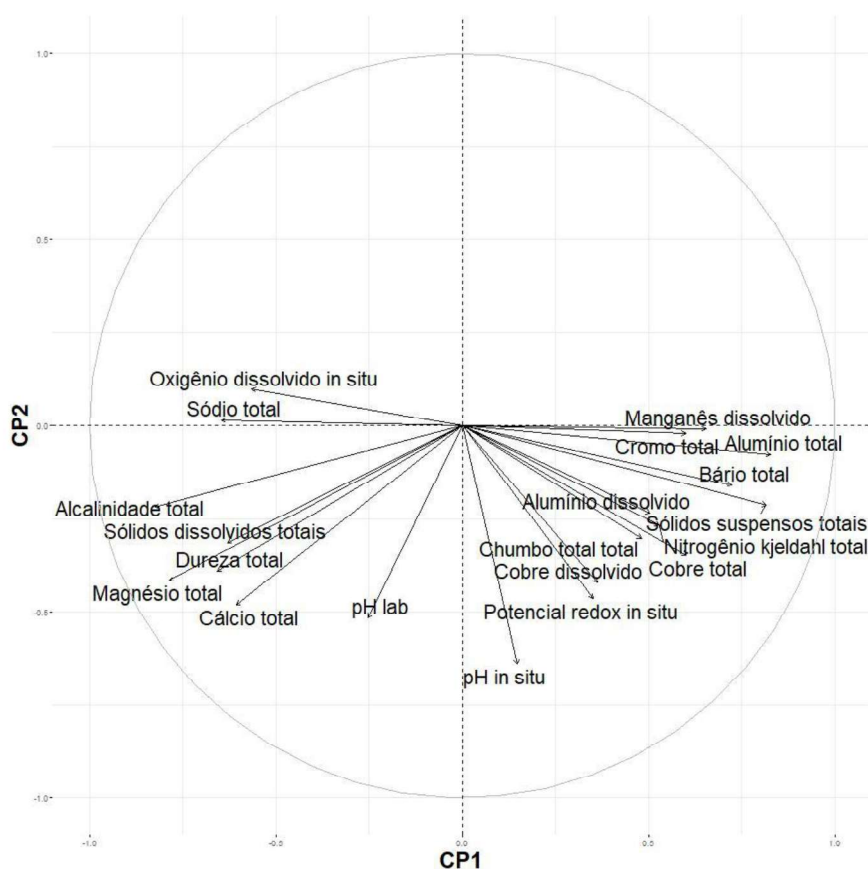


Tabela 9 - Variáveis de qualidade da água “não-críticas” listadas dentre as de maior contribuição para CP1

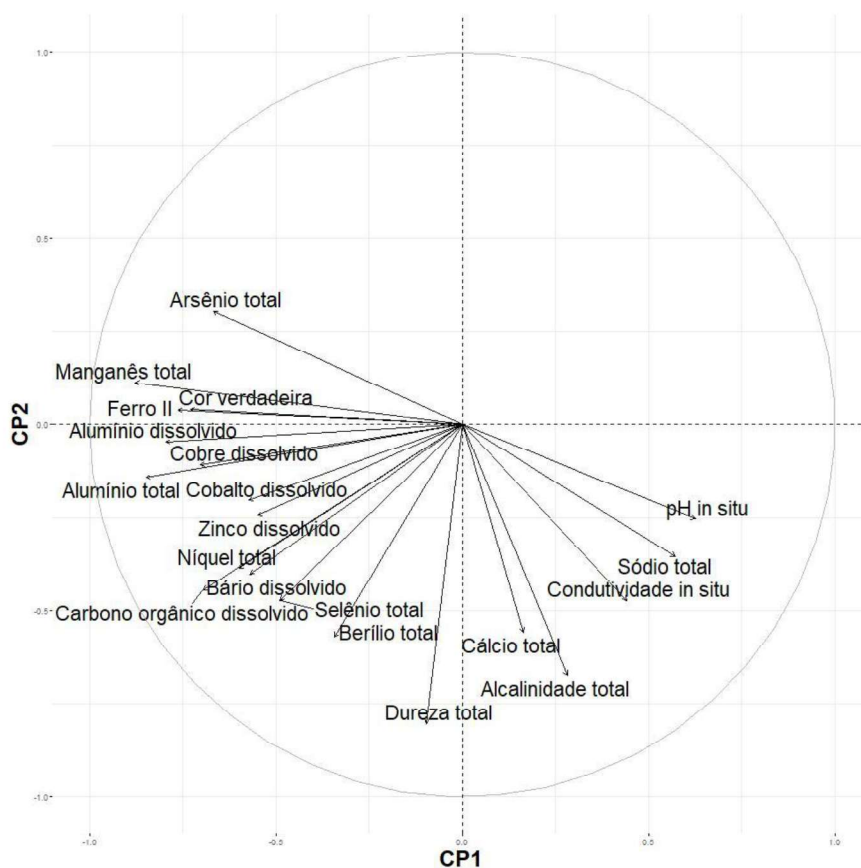
Variável “não crítica” de qualidade da água	Subregião
Cromo total	D
Nitrogênio Kjeldahl total	D
Chumbo total	I
Níquel total	G
Cobalto dissolvido	J
Selênio total	J
Clorofila a	K

Da mesma forma, também foram identificadas “variáveis não-críticas” entre aquelas de maior contribuição para a segunda componente principal - CP2 (Tabela 10), exceto nas subregiões E, F e I. Nas demais subregiões, as “variáveis não-críticas” que contribuíram expressivamente para CP2 variaram de 1 (subregião A, G, H, M e N) a 3 variáveis (subregião B) (Tabela 10) (Figura 20) (Apêndice - Figura A15).

Tabela 10 - Variáveis de qualidade da água “não-críticas” listadas dentre as de maior contribuição para CP2

Variável “não crítica” de qualidade da água	Subregião
Nitrogênio amoniacal	A, C, K
Cianeto livre	B, H
Nitrogênio Kjeldahl total	B
Ferro II	B
Sulfetos totais	C, L
Sólidos sedimentáveis	D
Chumbo total	N
Polifosfato	D, K, L, M
Antimônio dissolvido	G
Berílio total	J, K
Selênio total	J

Figura 20 - Contribuição de variáveis de qualidade da água para CP1 e CP2 (%), na subregião D



A ACP demonstrou que, na maioria dos casos, as “variáveis não-críticas” da qualidade da água, propostas por este estudo, tiveram menor relevância para explicar a variabilidade da qualidade da água da bacia hidrográfica do Rio Doce. A eventual ocorrência de “variáveis não-críticas” entre aquelas com maior contribuição para as duas primeiras componentes principais deveu-se à presença de *outliers* (Figuras 21 e 22). *Outliers* quando presentes em conjuntos de dados tratados por métodos estatísticos, frequentemente adicionam variância aos dados (Hair Jr *et al.*, 2006). *Outliers* imputaram um elevado grau de importância relativa a “variáveis não-críticas” na ACP, o que não é consistente com as características dessas variáveis.

Figura 21 - Box-Whisker de “variáveis não-críticas” de qualidade da água com maior contribuição para CP1 por subregião¹

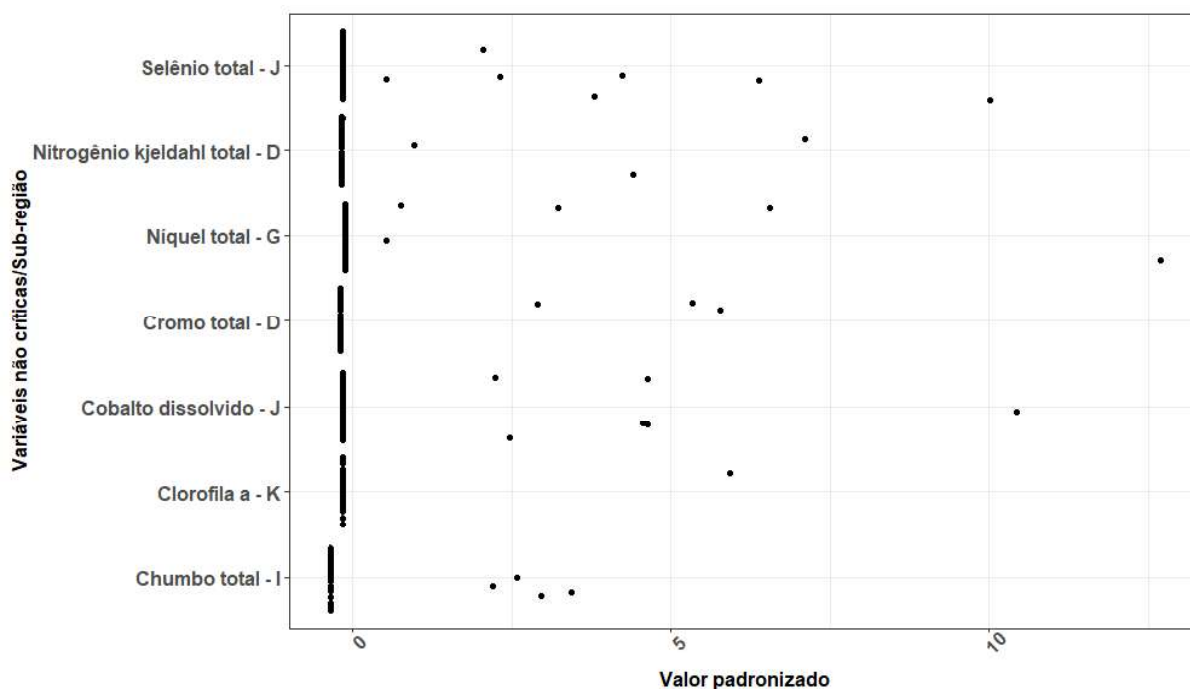
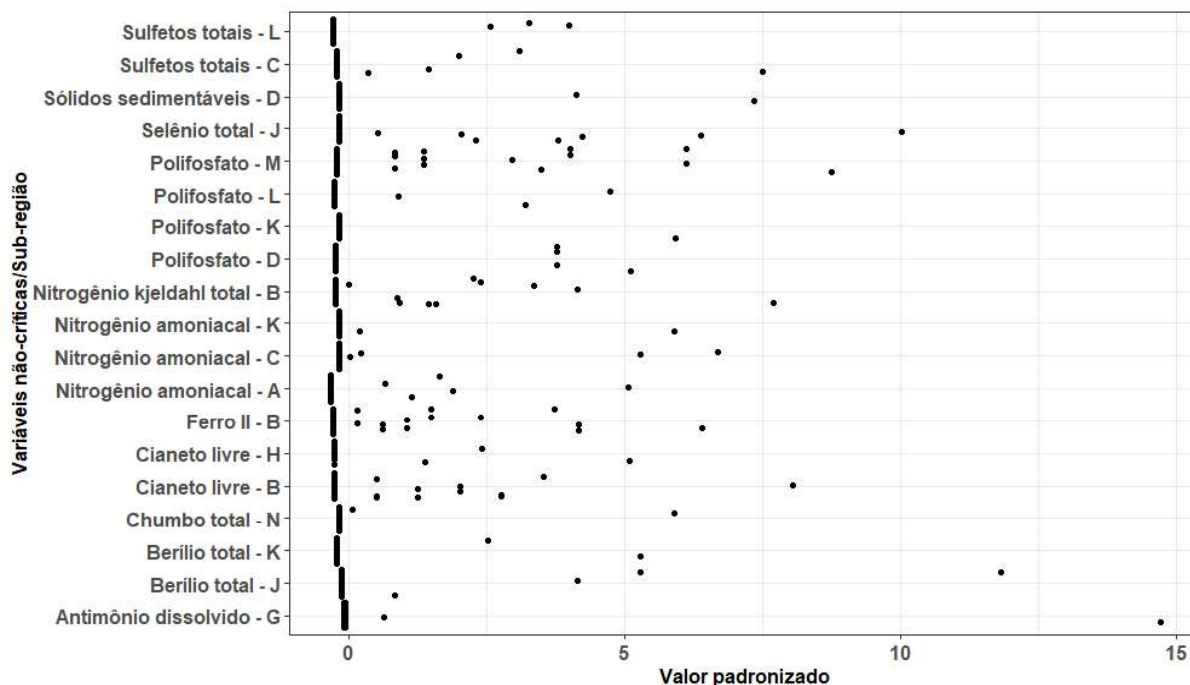


Figura 22 - Box-Whisker de “variáveis não-críticas” de qualidade da água com maior contribuição para CP2 por subregião²



¹ Variáveis foram padronizadas para melhorar a visualização. Letras representam as subregiões de qualidade da água

² Variáveis foram padronizadas para melhorar a visualização. Letras representam as subregiões de qualidade da água

7.2 Seleção da melhor configuração espacial de estações de monitoramento da qualidade da água

7.2.1 Subregiões de qualidade da água investigadas e análise exploratória de EWQI

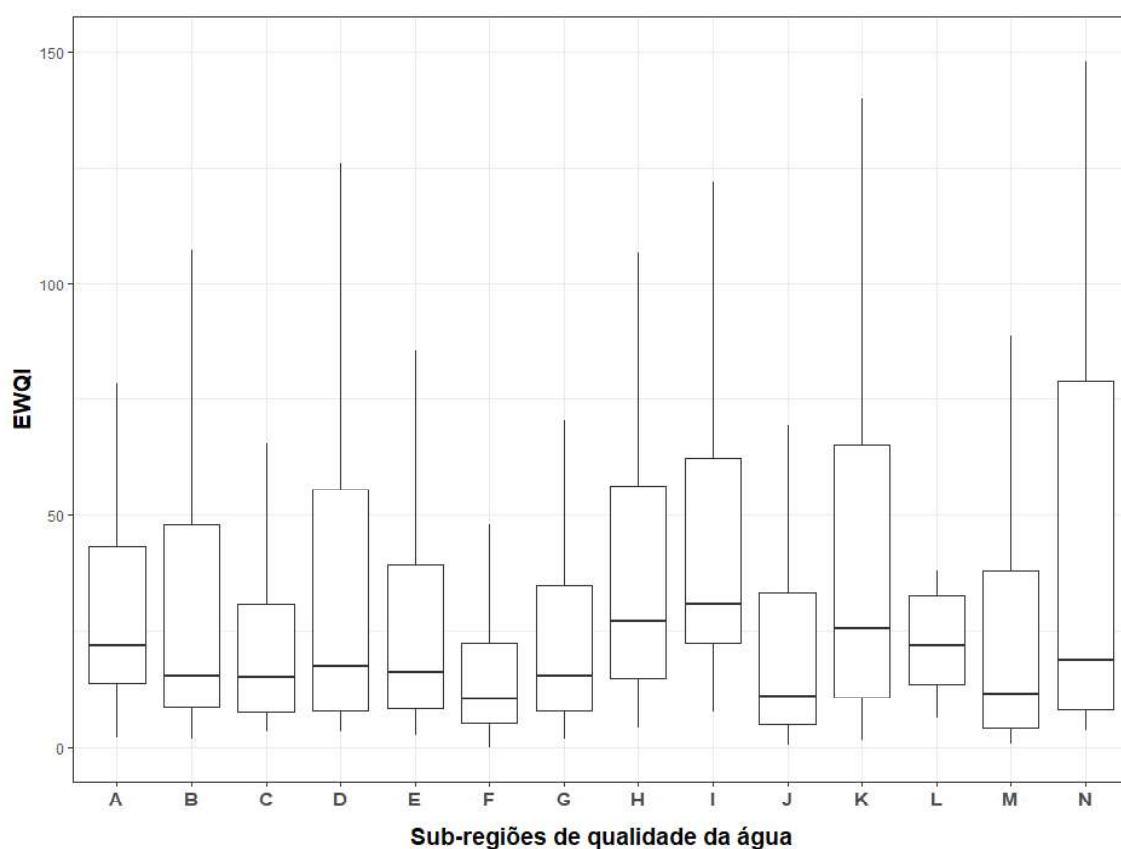
Em todas as subregiões há presença de valores do EWQI baixos, associados ao nível “Excelente qualidade da água”. A subregião com menor EWQI foi a F. Por outro lado, quanto ao máximo valor do EWQI, destacou-se a subregião K, com valor 1083,76, representativo da classe “Qualidade da água extremamente ruim” (Tabela 11).

Tabela 11 - Estatísticas descritivas de EWQI ao longo da bacia do Rio Doce

Sub-regiões	Mínimo	Máximo	Média	Mediana	Desvio padrão	Coeficiente de assimetria
A	2,01	364,48	62,87	28,24	79,6	1,95
B	1,80	892,35	67,68	19,89	122,1	3,78
C	3,26	1351,87	143,25	22,91	284,6	2,67
D	3,38	337,65	58,93	23,50	74,8	1,83
E	2,61	405,45	36,86	17,40	51,6	3,40
F	0,04	182,97	20,55	10,43	29,1	3,26
G	1,72	243,76	29,18	15,40	34,7	2,64
H	4,03	691,23	63,06	28,06	116,4	4,31
I	7,62	175,58	53,36	37,26	44,0	1,17
J	0,48	419,31	37,97	12,45	63,8	3,28
K	1,45	1083,76	171,58	47,65	277,6	1,94
L	6,10	833,05	133,53	28,07	210,0	1,80
M	0,66	814,99	47,34	15,49	84,6	4,47
N	3,53	445,82	58,58	22,72	84,5	2,79

De maneira geral o desvio padrão de valores do EWQI é expressivo. Isso sugere considerável dispersão do índice em todas as subregiões. Adicionalmente examinando o comportamento geral do coeficiente de assimetria pode-se inferir também que esse espalhamento não se dá de maneira uniforme. Ao contrário, os valores positivos e superiores a uma unidade sugerem forte assimetria à direita no comportamento distributivo dos valores do EWQI, reforçado pela análise dos gráficos Box-Whisker (Figura 23) e da discrepância entre média e mediana em cada subregião.

Figura 23 - Box-Whisker de EWQI na bacia do Rio Doce



A não aderência dos dados à distribuição normal, hipotetizada, confirmou-se pelo teste de Shapiro-Wilk (Tabela 12), visto que em todas as subregiões a hipótese nula de distribuição normal dos dados foi rejeitada. Esse fato toma importância nas etapas posteriores realizadas, pois norteou a escolha de testes não-paramétricos, recomendados para dados de distribuições não-simétricas (Maity, 2018), para avaliar outras hipóteses sobre o comportamento do EWQI.

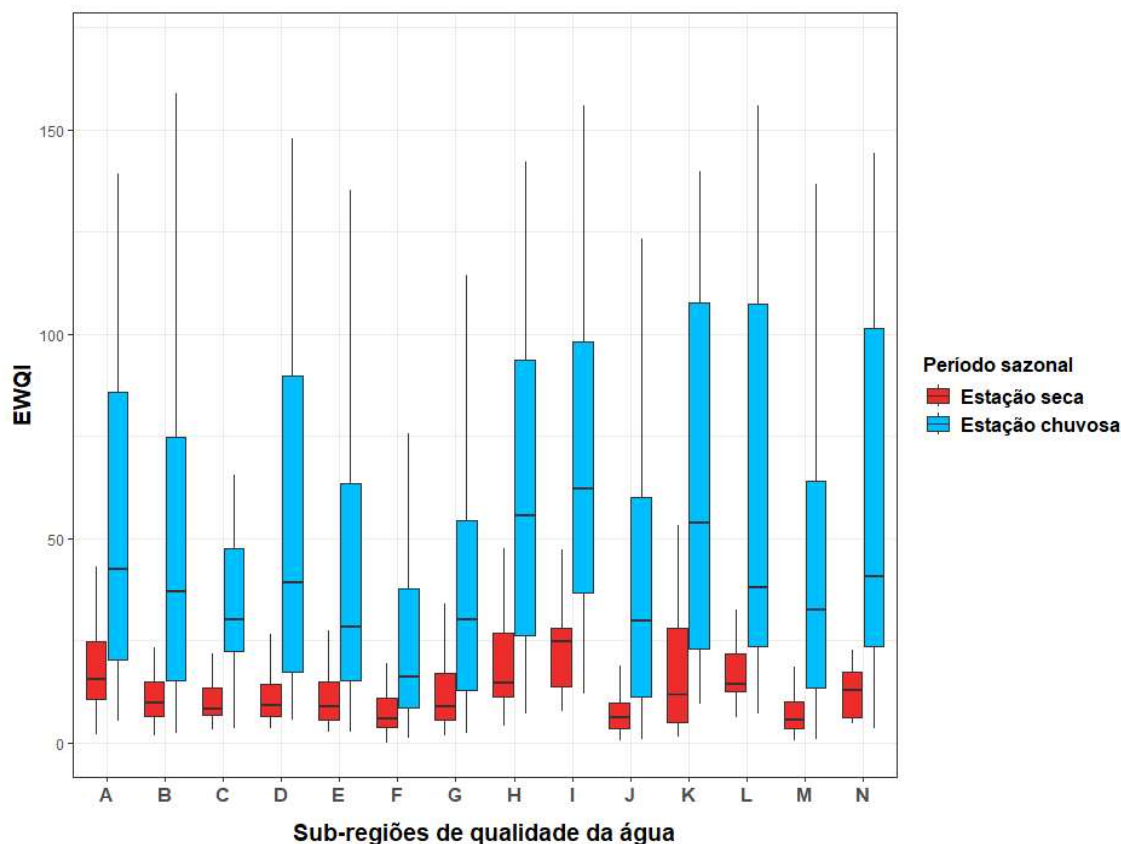
Tabela 12 - *p*-valor dos testes de normalidade de Shapiro-Wilk para séries temporais do EWQI em todas as subregiões

Subregiões	<i>p</i>-valor
A	0,0000003569
B	0,0000000000
C	0,0000000000
D	0,0000000003
E	0,0000000000
F	0,0000000000
G	0,0000000000
H	0,0000000001
I	0,0001426763
J	0,0000000000
K	0,0000000280
L	0,0000000304
M	0,0000000000
N	0,0000000297

7.2.2 Variabilidade espaço-temporal da qualidade da água

A sazonalidade exerce grande influência sobre a qualidade da água em toda a bacia hidrográfica do Rio Doce. Em todas as subregiões a estação seca promove menor dispersão do EWQI com predominância de valores em patamares inferiores. A estação seca, de forma geral, detém até 75% dos valores do EWQI em níveis inferiores ou próximos a 50 unidades (Figura 24). Esse comportamento sugere haver predominância de níveis de melhor qualidade da água nesse período. Ao contrário a estação chuvosa em geral produz valores do EWQI mais expressivos que a estação seca, assim como há maior dispersão, possivelmente como efeito direto da dinâmica espacial das chuvas na bacia hidrográfica (Passos *et al.*, 2021).

Figura 24 - Box-Whisker de EWQI na bacia do Rio Doce por período sazonal



A observação da distribuição das classes de qualidade da água por período sazonal, corrobora esse aspecto (Figura 25). O comprimento das barras está associada à frequência com que cada classe se apresentou ao longo do período analisado. Na estação seca a classe de qualidade excelente é bem mais frequente que as demais, embora tenham ocorrido eventualmente registro das demais classes.

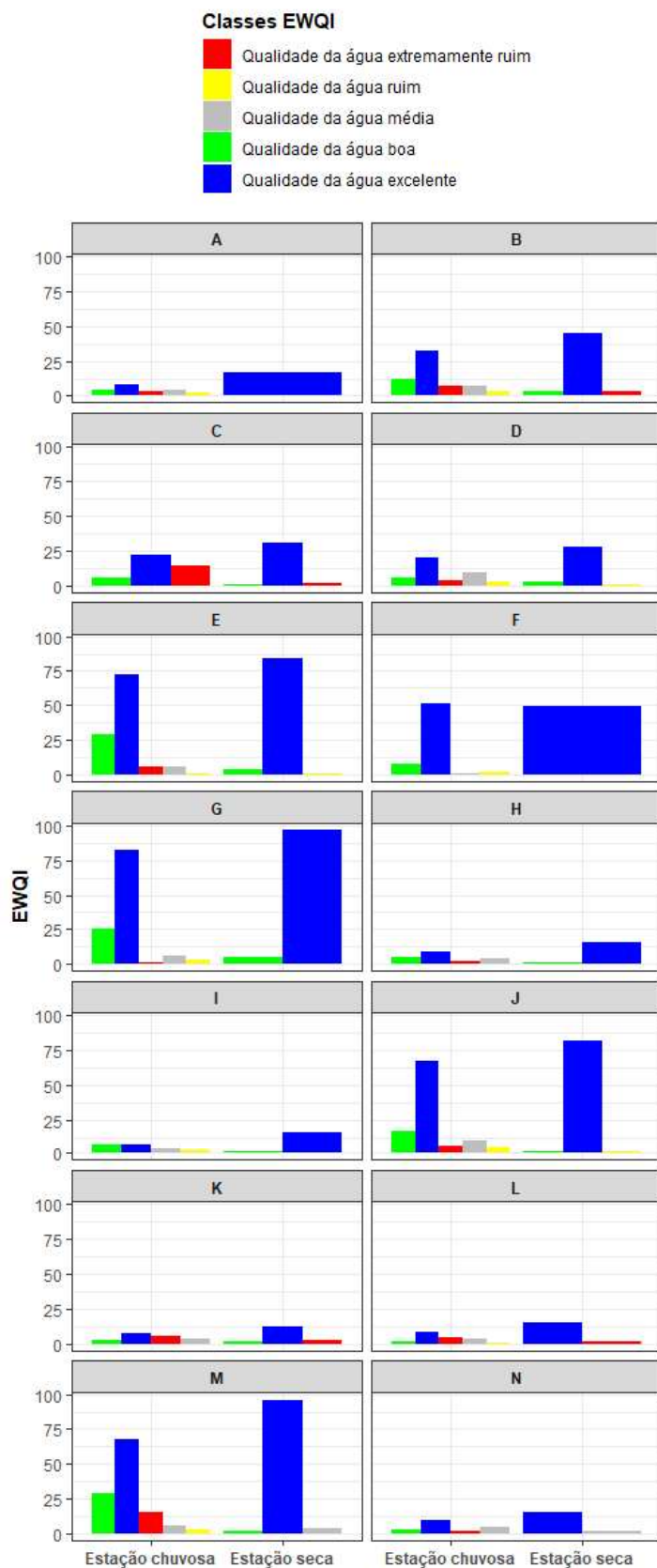
Durante a estação chuvosa, por outro lado, embora haja menor número absoluto de registros da classe de qualidade da água excelente, comparativamente à estação seca, há maior representativa das demais classes, o que confirma a maior dispersão de EWQI demonstrada na Figura 24. Além disso, esse período é marcado por valores de EWQI mais expressivos, representativos dos piores níveis de qualidade da água previstos na interpretação do índice.

Esse comportamento está associado ao menor efeito da ressuspensão de sedimentos e menor influência de fontes difusas de poluição durante a estação seca, em virtude da característica redução da precipitação. Já na estação chuvosa por outro lado, além da influência dos processos mencionados há maior impacto da poluição de origem

antrópica das áreas com baixa cobertura vegetal e proteção dos solos, pela lixiviação (Oliveira; Quaresma, 2017).

Além disso, a presença de grande volume de rejeitos depositados no leito e margem dos rios, por onde a pluma de rejeitos passou, aumenta a possibilidade de remobilização de poluentes específicos associados ao desastre de Fundão, como metais e metaloides, pelo aumento da vazão nas áreas atingidas, durante a estação chuvosa (Andrades *et al.*, 2021; da Cunha Richard *et al.*, 2020).

Figura 25 - Distribuição das classes de qualidade da água definidas para EWQI por período sazonal na bacia do Rio Doce



Outros autores também evidenciaram o efeito da sazonalidade e a influência dos rejeitos de mineração sobre o comportamento de variáveis da qualidade da água na bacia do Rio Doce. Santos *et al.* (2023) ao estudar o comportamento sazonal da qualidade da água na bacia do Rio Gualaxo do Norte, utilizando ACP concluíram que durante a estação seca a baixa capacidade de diluição dos rios da região maximiza os efeitos da poluição orgânica, derivada principalmente de efluentes domésticos, enquanto que durante a estação chuvosa, as chuvas intensas promovem aumento de Turbidez e Sólidos totais, como efeito direto da presença de depósitos de rejeitos de mineração ao longo de grande parte dos rios da bacia.

Ao comparar médias de variáveis de qualidade da água de 2017 a 2018, entre as estações seca e chuvosa, utilizando dados de 63 estações operadas pelo IGAM, na bacia do Rio Doce, Passos *et al.* (2021) observaram que tanto variáveis associadas às diferentes formas de sedimentos na água, como Sólidos suspensos totais, Sólidos dissolvidos totais, Sólidos totais e Turbidez; quanto metais como, Alumínio, Níquel e Zinco apresentaram aumento durante o período chuvoso. Os autores atribuem esse comportamento ao aumento do escoamento superficial, pelo aumento da ocorrência das chuvas, no primeiro caso, e no segundo, à presença dos metais associados a partículas argilosas que compõem o sedimento, cuja remobilização para a água é facilitada com o aumento do escoamento.

A aplicação do teste de Mann-Whitney, a 5% de significância estatística, para comparação de EWQI entre as estações seca e chuvosa confirma que a sazonalidade exerce forte influência no comportamento temporal da qualidade da água na bacia, aspecto também confirmado por Fraga *et al.* (2021). A hipótese nula de igualdade das medianas foi rejeitada na aplicação do teste em todas as subregiões de qualidade da água, considerando as séries temporais de EWQI. Portanto, pode-se assumir que há diferenças estatisticamente significativas entre a qualidade da água entre as estações seca e chuvosa na bacia (Tabela 13).

Tabela 13 - p -valor de teste de Mann Whitney aplicado sobre séries temporais de EWQI por período sazonal

Subregiões	p -valor
A	0,000145947
B	0,000004842
C	0,000004996
D	0,000011246
E	0,000000027
F	0,000000000
G	0,000000020
H	0,000050657
I	0,004364295
J	0,000000001
K	0,030925146
L	0,004827955
M	0,000000005
N	0,000093878

7.2.3 Variáveis de maior contribuição à EWQI

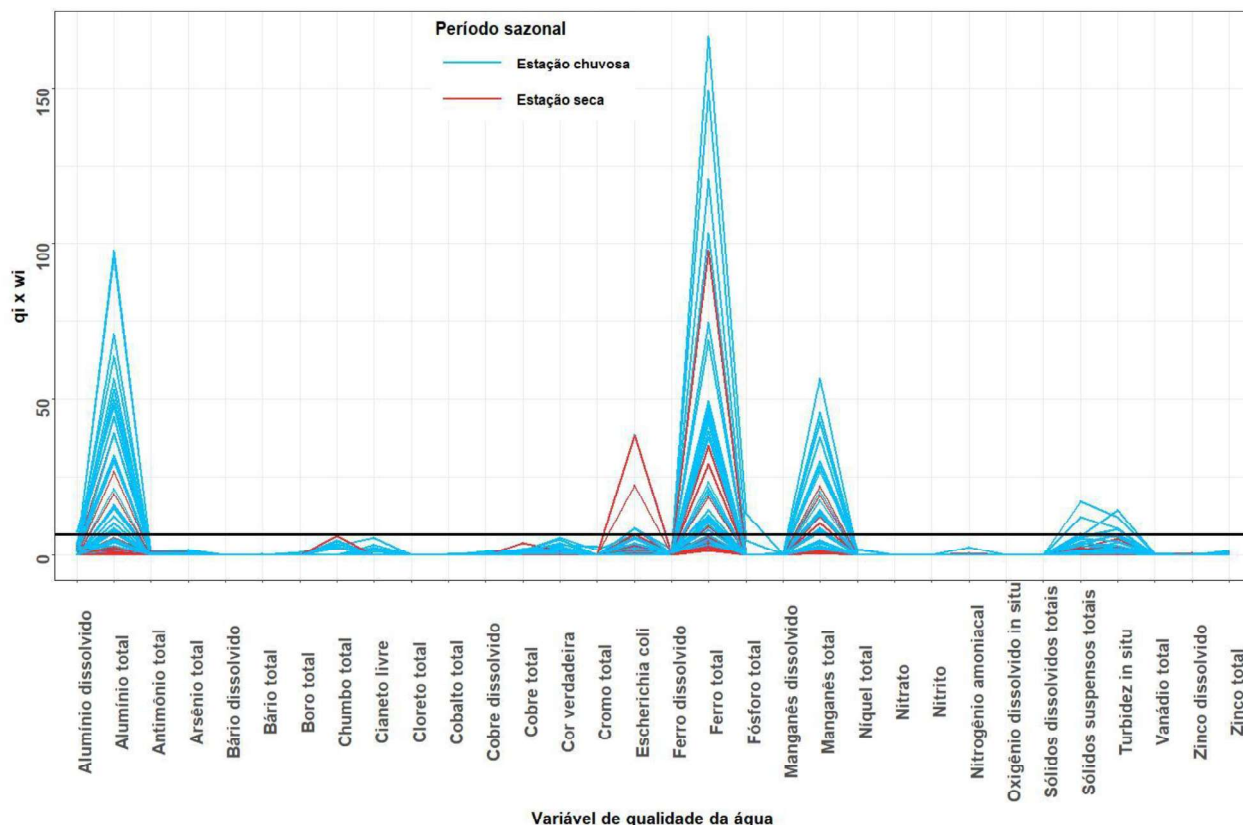
O valor de EWQI determinado por estação a cada campanha é o somatório dos termos $q_i w_i$ atribuídos a cada variável que compõe o índice, como descrito na seção 6.1.2.1. Portanto, esse termo pode indicar claramente quais variáveis possuem maior contribuição para o índice geral (Figuras 26 a 30). Foram elaborados gráficos sub-regionais onde cada linha representa uma campanha de monitoramento realizada por estação que, por sua vez, relaciona cada variável integrante de EWQI ao valor do termo $q_i w_i$.

Além disso, de forma a maximizar a interpretação dos gráficos, as linhas assumiram cores específicas atribuídas aos distintos períodos sazonais da bacia, estação seca e chuvosa, e foi adicionado a cada gráfico uma linha constante representativa do percentil 95 referente à distribuição de $q_i w_i$ por sub-região.

De forma geral na bacia, as variáveis de qualidade da água que contribuem mais expressivamente para EWQI, alternando-se nos postos de maior importância, são Ferro total, Alumínio total, Alumínio dissolvido e Manganês total (Figura 26). De maneira geral variáveis que apresentam medidas maiores ou iguais aos limites estabelecidos como referência, tem maior peso no valor total do EWQI de cada campanha (Li; Qian; Wu, 2010).

Cabe ressaltar que as variáveis Ferro total e Alumínio total não possuem limites estabelecidos pela resolução CONAMA n° 357/2005 (BRASIL, 2005) e resolução conjunta COPAM/CNRH-MG n° 08/2022 (COPAM-CERH, 2022). Para que pudessem integrar o índice, nesses casos, foram tomados os mesmos limites admissíveis preconizados para as formas dissolvidas de ambos os elementos.

Figura 26 - Contribuição de variáveis de qualidade da água para EWQI na subregião D*



***A linha preta representa o percentil 95% de $q_i \times w_i$ de todas as campanhas da subregião analisada.**

Ferro, Manganês e Alumínio têm sido reportados como algumas das variáveis de qualidade da água com maior frequência de não-conformidades com os padrões ambientais na bacia do Rio Doce, segundo estudos baseados em dados anteriores ao desastre de Fundão, especialmente nas bacias localizadas na região de cabeceiras (bacias dos Rios Gualaxo do Norte, Carmo, Piracicaba e Piranga), que também estão circunscritas no chamado Quadrilátero Ferrífero (Jordão *et al.*, 1996; Wild *et al.*, 2023).

Esse cenário relaciona-se tanto a fatores de ordem natural quanto antropogênica. A composição geoquímica decorrente do processo de evolução geológica,

caracterizada pela formação de depósitos minerais no Quadrilátero Ferrífero, onde estão incluídos metais e metaloides, contribuíram para a elevada concentração desses elementos nos solos e sedimentos na região do alto Doce (Deschamps *et al.*, 2002; Guevara *et al.*, 2018; Kessler; Müller, 1988; Spier *et al.*, 2007). No quadrilátero os depósitos de minério de ferro frequentemente incluem Manganês e Alumínio em associação com oxi-hidóxidos de Ferro (Costa *et al.*, 2010).

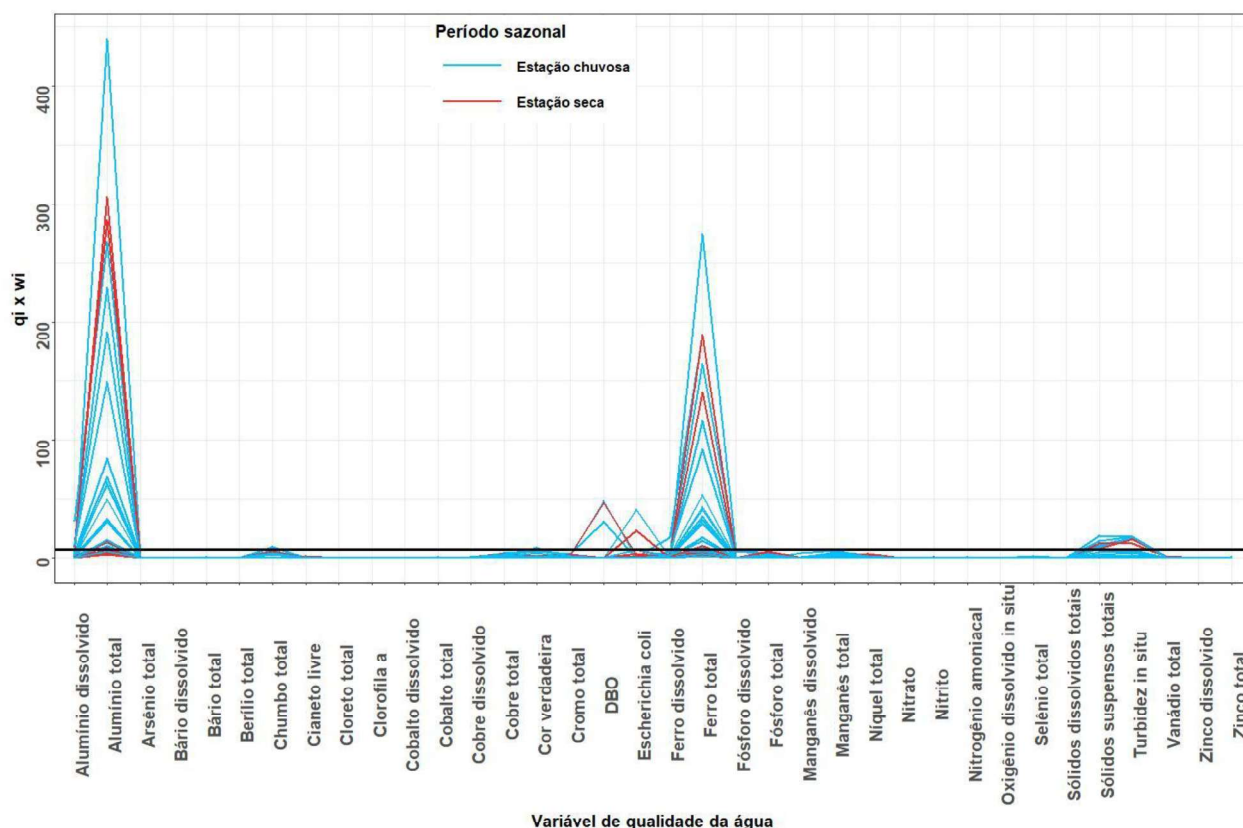
Adicionalmente, a mineração que se instalou e vem sendo desenvolvida no Quadrilátero ferrífero há séculos e que, marca profundamente a exploração econômica da bacia do Rio Doce, promoveu ainda mais o enriquecimento de solos, sedimentos e água da região em metais e metaloides (Silva *et al.*, 2013; de Vicq *et al.*, 2015; Paraguassú *et al.*, 2019; Yamamoto *et al.*, 2023; Hatje *et al.*, 2017).

Por outro lado, estudos recentes sobre a constituição física, mineralógica e geoquímica dos rejeitos anteriormente armazenados na barragem de Fundão, além de outros visando caracterizar a lama, produto da mistura de rejeitos e sedimentos de fundo depositados nas calhas e margem da rede hidrográfica da bacia do Rio Doce, dão conta da forte presença também de Ferro, Alumínio e Manganês (Duarte; Neves, 2023; Figueiredo *et al.*, 2020; Silva *et al.*, 2018; Segura *et al.*, 2016).

De fato, esses elementos não possuem uma única origem na bacia do Rio Doce e, da mesma forma há dificuldade em se estabelecer o grau de influência de cada fonte potencial na sua presença nos mananciais da bacia atualmente, embora alguns autores apontem forte influência do desastre nos padrões atuais de distribuição desses elementos nos mananciais impactos pelo desastre na bacia (Duarte; Neves, 2023; Yamamoto *et al.*, 2023).

O segundo grupo de variáveis de maior importância na composição do EWQI ao longo de todas as subregiões estudadas não possui relação direta com os impactos decorrentes do desastre de Fundão. São elas: *E. coli* e DBO (Figura 27). Muitos estudos sugerem haver forte correlação entre essas variáveis, muitas vezes interpretadas conjuntamente em estudos de qualidade da água como medida das condições sanitárias de rios (poluição orgânica) e fontes de poluição difusa em bacias hidrográficas (de Souza Pereira *et al.*, 2019; Yang *et al.*, 2020; Bortoletto *et al.*, 2015).

Figura 27 - Contribuição de variáveis de qualidade da água para EWQI na subregião L*



***A linha preta representa o percentil 95% de $q_i \times w_i$ de todas as campanhas da subregião analisada.**

De fato, a bacia do Rio Doce apresenta um significativo déficit quanto as suas condições sanitárias, especialmente quando são analisados seus dados estatísticos referentes ao esgoto doméstico, o que corrobora a importância das variáveis acima mencionadas na concepção dos valores do EWQI. Aproximadamente 600.000 pessoas não são atendidas por sistemas coletivos de coleta de esgoto na bacia do Rio Doce ($\cong 16,3\%$ da população total da bacia) (ENGEORPS, 2021).

Mesmo a parcela da população que tem seu esgoto coletado nem sempre é atendida por ETE's. Esse grupo é formado por aproximadamente 1.856.000 habitantes. Em resumo, aproximadamente 270.000.000 m³ de esgoto domésticos são lançados *in natura* anualmente na rede hidrográfica da bacia do Rio Doce, poluindo os mananciais e impactando a biota aquática (ENGEORPS, 2021; ANA, 2023a).

Vários autores ao investigarem variáveis de qualidade da água, em diversas regiões da bacia do Rio Doce também evidenciaram padrões dessas variáveis associando-as

as condições sanitárias deficientes (Oliveira Serrano *et al.*, 2020; Fraga *et al.*, 2021). Em pesquisa recente realizada na bacia do Rio Gualaxo do Norte, em que foram analisadas variáveis de qualidade da água a partir de amostras coletadas em áreas afetadas, não afetadas pelos rejeitos da barragem de Fundão e tributários, dos Santos *et al.* (2023) observaram que as maiores concentrações de *E. coli* e DBO ocorreram em tributários do Rio Gualaxo do Norte não impactados pelos rejeitos.

Os autores corroboram as precárias condições de saneamento da bacia do Rio Doce como promotoras do aumento da concentração dessas variáveis, ao mencionarem que isso deve-se à disposição incorreta de esgotos domésticos não tratados diretamente nos mananciais da região.

O terceiro grupo de variáveis que se destaca no EWQI, de forma geral na bacia, é constituído pelas variáveis Turbidez, Sólidos suspensos totais e Fósforo total (Figura 28).

A importância atribuída a esse conjunto pode estar relacionada a intensa fragmentação da paisagem e à baixa proteção exercida pela cobertura vegetal na bacia do Rio Doce, que favorece o transporte de solos, sedimentos e nutrientes aos rios, decorrentes da erosão e lixiviação dos solos, especialmente durante chuvas intensas (Oliveira; Quaresma, 2017; Fraga *et al.*, 2021). Segundo dados de ENGECORPS (2021) 58,4% da área total da bacia do Rio Doce possui forte susceptibilidade erosiva e 7,5% são classificadas nesse mesmo critério como muito fortemente susceptíveis à erosão.

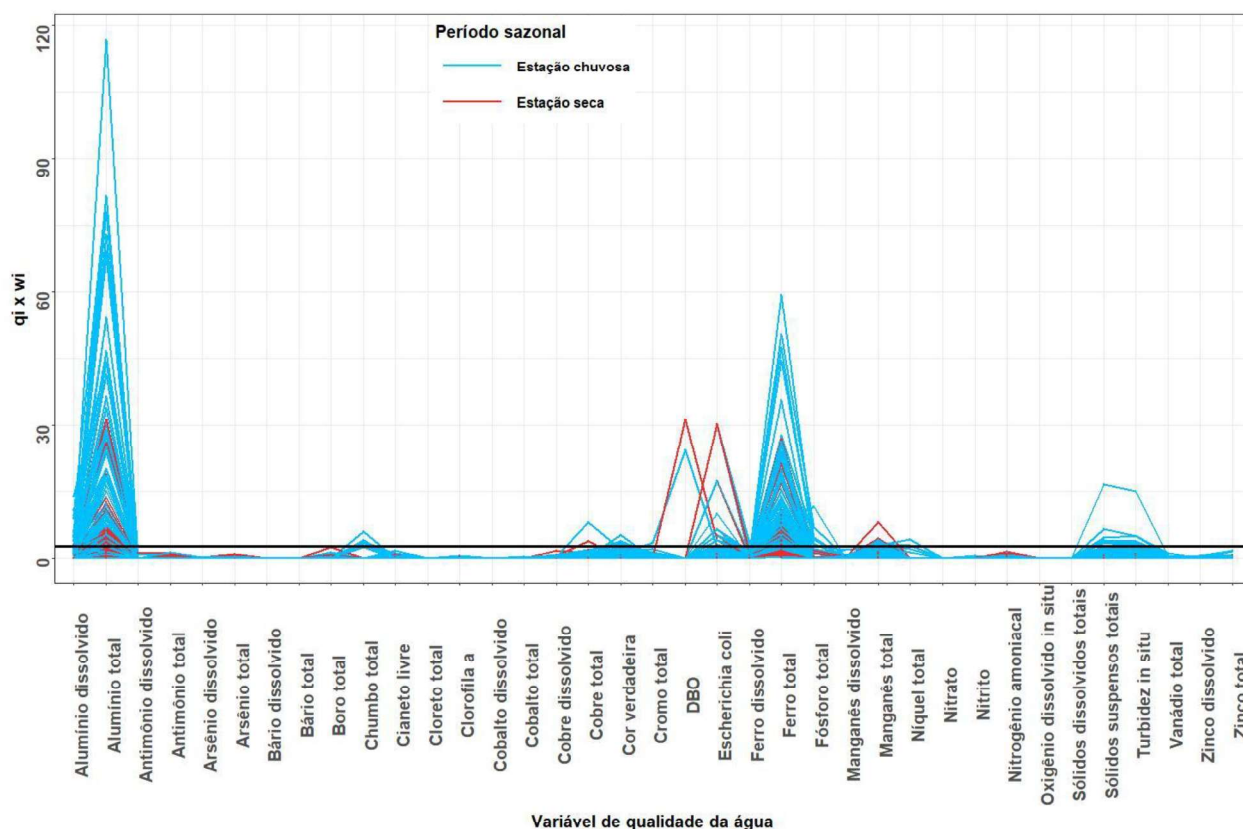
Contudo, particularmente no caso da Turbidez e Sólidos suspensos totais, inúmeras pesquisas têm demonstrado um aumento expressivo da carga de sólidos em suspensão nas áreas atingidas pela passagem da onda de rejeitos da barragem de Fundão, especialmente durante a estação chuvosa, atribuindo, portanto, a importância dessas variáveis na concepção do EWQI aos efeitos diretos da deposição de rejeitos de mineração ao longo da rede hidrográfica impactada (Kütter *et al.*, 2023; Miller, Ghisolfi; Barroso, 2023; dos Santos *et al.*, 2023).

Ainda assim, deve-se considerar que esses resultados foram influenciados pelo comportamento das chuvas nos últimos anos. A estação seca tem se prolongado e

intensificado com frequência na bacia especialmente durante os anos hidrológicos 2013/2014 a 2014/2015, com grande redução no índice pluviométrico em todas as regiões da bacia. Apesar desse período de maior concentração de secas severas corresponderem a era pré-desastre, o ano 2018/2019 também foi marcado por índices pluviométricos abaixo da média (Da Silva *et al.*, 2022b),

Esse fato pode ter atuado reduzindo a importância da turbidez e sólidos na concepção de EWQI para algumas regiões da bacia. Da Silva *et al.* (2022b) corroboram com essa hipótese ao afirmar que a turbidez tem demonstrado melhorias ao longo da bacia do Rio Doce, desde a estação seca de 2018, considerado o horizonte de análise do período pós-desastre.

Figura 28 - Contribuição de variáveis de qualidade da água para EWQI na subregião G*



***A linha preta representa o percentil 95% de $q_i \times w_i$ de todas as campanhas da subregião analisada.**

Para diversos autores a deposição de rejeitos de mineração oriundos da barragem de Fundão faz com que todos os locais impactados estejam suscetíveis à remobilização decorrente do aumento da vazão, por sua vez, decorrente de eventos de chuva. Com

efeito esse processo tende a aumentar a carga de Sólidos em suspensão totais e Turbidez, especialmente durante as cheias anuais (Yamamoto *et al.*, 2023; Hatje *et al.*, 2017; da Cunha Richard *et al.*, 2020; Santos *et al.*, 2023).

Se por um lado Turbidez e Sólidos suspensos totais apresentam forte correlação com a sazonalidade, decorrente de chuvas intensas ou da estação chuvosa, por outro, durante a estação seca os níveis dessas variáveis, segundo alguns autores (dos Santos *et al.*, 2023; Da Silva *et al.*, 2022b), já se aproximam dos padrões registrados no período pré-desastre em algumas regiões da bacia.

Entretanto, é necessário que as pesquisas que avaliam esse processo atentem para dois aspectos importantes. Em primeiro lugar, o período compreendido entre os anos de 2013 e 2017 foi marcado por secas severas e prolongadas na bacia do Rio Doce. Com efeito as vazões máximas registradas alcançaram níveis menores que às médias históricas, o que por sua vez, pode ter afetado os níveis de Turbidez e Sólidos suspensos totais registrados nesse período (de Jesus *et al.*, 2020).

A outra importante ressalva que deve ser feita, nesse contexto, é que dados de monitoramento realizado pelo IGAM na bacia do Rio Doce até 2015 podem não conter informação suficiente para se traçar um paralelo consistente entre o comportamento de variáveis de qualidade da água pré e pós-desastre, uma vez que frequentemente a instituição realizava apenas duas campanhas anuais de monitoramento, uma na estação seca e outra na chuvosa, com grande probabilidade de não conseguir capturar o comportamento das variáveis nas condições de maior vazão na bacia. Portanto, possivelmente as condições de referência de muitas variáveis podem estar sendo subestimadas (Wild *et al.*, 2023).

Quanto à variável Fósforo total, investigações de longa data também têm relacionado suas anomalias, e eventual excedência, a maior influência de poluição orgânica difusa e pontual ligadas ao lançamento de esgotos sem tratamento prévio nos mananciais da bacia, o que muitas vezes tem colaborado para os eventos de floração de algas historicamente observados (Jardim *et al.*, 2014; Petrucio *et al.*, 2005; Medeiros *et al.*, 2012; Figueiredo *et al.*, 2014). Contudo, pesquisas recentes têm destacado que modificações na geoquímica dos sedimentos da bacia decorrentes da lama de rejeitos

de Fundão pode estar aumentando sua biodisponibilidade, com sérios riscos aos ecossistemas aquáticos.

O rompimento da barragem de Fundão promoveu uma intensa descarga de rejeitos e nutrientes no sistema fluvial e zona costeira/estuarina. Níveis de Fósforo total observados no baixo Rio Doce são até cinco vezes maiores que na plataforma continental, durante a estação chuvosa. Em termos das frações, também nas águas interiores do baixo Doce são encontrados percentuais elevados de Fósforo inorgânico dissolvido em relação ao Fósforo total, com a maior proporção média observada durante a estação seca (~42%) (Cagnin *et al.*, 2022).

Além disso, muitos autores têm alertado para a possibilidade de efeitos crônicos de longo prazo da presença excessiva de oxi-hidróxidos de Ferro e Fósforo decorrentes do desastre de Fundão, tanto no baixo Rio Doce, quanto na zona costeira/estuarina, com efeitos nocivos ao equilíbrio das cadeias tróficas e saúde dos ecossistemas aquáticos pelas perturbações nos ciclos biogeoquímicos de macronutrientes.

Óxidos-hidróxidos de Ferro atuam na regulação de Fósforo no solo através dos processos de adsorção/dessorção em complexos associados a minerais como a geotita. Esse mecanismo também ocorre nos solos de regiões estuarinas, contudo, devido às condições menos estáveis relacionadas ao pH, potencial redox e decomposição de matéria orgânica, o equilíbrio desse mecanismo pode ser quebrado facilitando assim a transferência do Fósforo da fase imóvel nos complexos à fase livre na água (Queiroz *et al.*, 2021a; Costa *et al.*, 2021).

Embora o conteúdo de Ferro associado a minerais como hematita e geotita, tenha decrescido entre 2015 e 2017, em solos estuarinos impactados pelos rejeitos de Fundão, a concentração média de Fósforo dissolvido na água se mantém ainda em patamares elevados, superiores tanto aos níveis estabelecidos pelas normas ambientais brasileiras, quanto à linha de referência do período pré-desastre (Queiroz *et al.*, 2021a; Costa *et al.*, 2021).

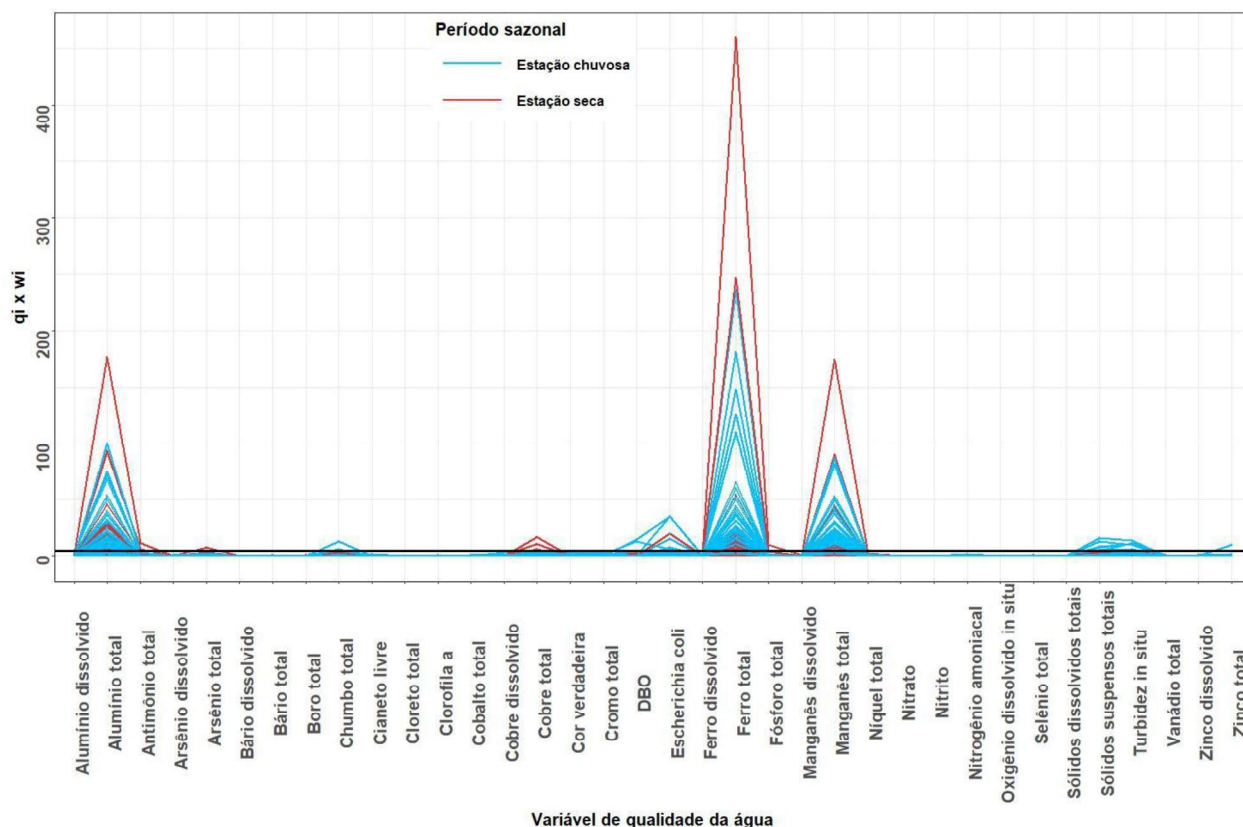
Esse fato corrobora a hipótese de que os rejeitos atuaram como transportadores de Fósforo para a região estuarina e fornece evidências de que os mecanismos que promovem a mobilidade de Fósforo dos sedimentos à água estão em curso, e podem

favorecer o aumento da produtividade primária, com consequente proliferação de algas nessa região (Queiroz *et al.*, 2021a; Costa *et al.*, 2021).

O último grupo de variáveis consideradas de maior contribuição ao EWQI não apresenta comportamento uniforme ao longo de todas as sub-regiões, mas destacam-se em algumas e em outras não. Além disso elas possuem contribuição relativamente menor ao EWQI, ilustrado pelos picos mais baixos nos gráficos, quando comparadas as que compõem os grupos já apresentados.

Arsênio total e Antimônio total destacam-se na subregião B (Figura 29); Boro total, Boro dissolvido, Sólidos totais dissolvidos e Cloreto na subregião N; Clorofila a possui importância na subregião M (Figura 30); Cádmio total na subregião J; Cianeto livre nas subregiões E e J; Cobre total nas subregiões B, E, F, G e M; e Chumbo total nas subregiões B, E, G, H, J, L e M. A contribuição das variáveis de qualidade da águas ao EWQI nas demais subregiões pode ser consultada no Apêndice, Figuras A16 a A24.

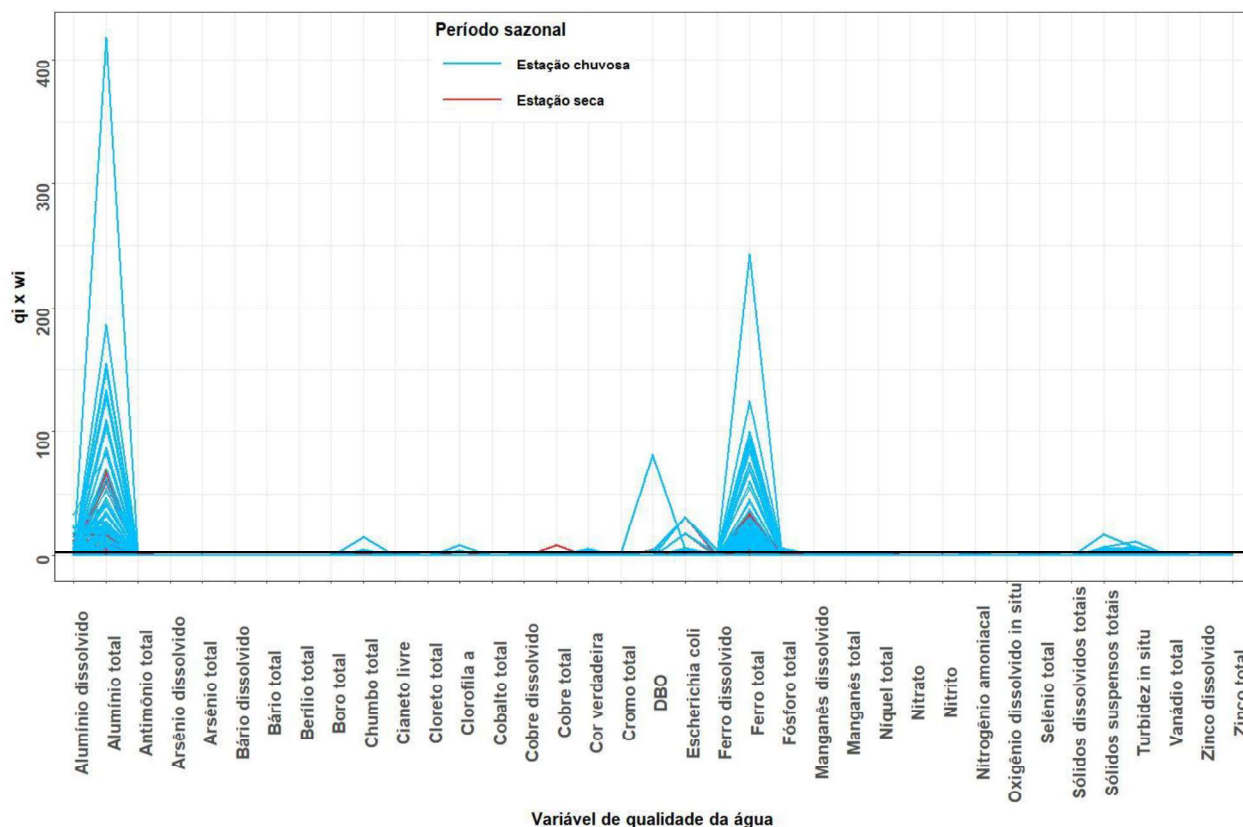
Figura 29 - Contribuição de variáveis de qualidade da água para EWQI na subregião B*



***A linha preta representa o percentil 95% de $q_i \times w_i$ de todas as campanhas da subregião analisada.**

A maioria dessas variáveis é classificada como metais e metaloides, cujas propriedades oferecem risco ao meio ambiente, ecossistemas aquáticos e saúde humana, em cenários de exposição aguda e crônica. Alguns pesquisadores afirmam que parte considerável desses elementos, não são integrantes da constituição dos rejeitos, mas, foram remobilizados de depósitos de sedimento fluviais e solos enriquecidos por metais e metaloides, derivados da atividade minerária historicamente desenvolvida na região do Quadrilátero Ferrífero, pela pluma de rejeitos (da Cunha Richard *et al.*, 2020; Silva *et al.*, 2018; Duarte *et al.*, 2021).

Figura 30 - Contribuição de variáveis de qualidade da água para EWQI na subregião M*



***A linha preta representa o percentil 95% de $q_i \times w_i$ de todas as campanhas da subregião analisada.**

A análise realizada para identificação das variáveis com maior contribuição ao EWQI ao longo das subregiões investigadas é consistente com outras pesquisas realizadas sobre os impactos do desastre de Fundão sobre a qualidade da água, uma vez que

as mesmas variáveis desde 2015 frequentemente são observadas por essas investigações em concentrações elevadas, muitas vezes superiores aos limites considerados seguros pela legislação brasileira e outros organismos internacionais (Martins; Takahashi, 2022; Sá *et al.*, 2021; de Carvalho *et al.*, 2018).

Além disso, os resultados apontam que a estação chuvosa concentra a maior densidade das maiores contribuições de cada EWQI determinado em todas as sub-regiões, fato igualmente evidenciado por outros pesquisadores, que associam esse comportamento também à ressuspensão da lama depositada pela passagem da pluma de rejeitos e, por consequência, remobilização de metais e metaloides para a água (de Carvalho *et al.*, 2018).

Presume-se que os casos em que esse padrão não é evidenciado estão relacionados a chuvas isoladas ocorridas na estação seca em período próximo à data em que as campanhas de monitoramento foram realizadas, de forma que o efeito desses episódios sobre as variáveis ainda pôde ser capturado pelas amostras. A exceção coube às variáveis Cobre total e Cádmio total cujos resultados sugerem possuir concentrações mais elevadas durante a estação seca.

Alguns esforços já têm sido feitos na investigação de potenciais e efetivos impactos sobre organismos aquáticos decorrentes da exposição a elementos como os mencionados e destacam os elevados riscos a que estão expostos esses organismos e a população da bacia, especialmente em relação ao consumo de peixes e outros organismos aquáticos.

Outrossim, respostas fisiológicas, comportamentais e genotóxicas têm sido observadas em organismos bioindicadores face a exposição a agentes estressores ambientais, como os metais ou a própria carga excessiva de sedimentos decorrente do desastre (Mendes *et al.*, 2020; Gabriel *et al.*, 2020; Yamamoto *et al.*, 2022; Denis Abessa *et al.*, 2023; Andrades *et al.*, 2021).

Adicionalmente, a despeito de todos os diversos impactos e efeitos agudos associados aos rejeitos da barragem de Fundão já evidenciados, e dos efeitos crônicos decorrentes da presença excessiva de nutrientes descritos, muitos pesquisadores também têm alertado para a possibilidade de efeitos crônicos

decorrentes da deposição de sedimentos enriquecidos por metais e metaloides ao longo da bacia.

A composição dos rejeitos predominantemente formados por óxi-hidróxidos de Ferro, como caulinita e geotita, favorece a formação de complexos com metais e metaloides, através de processos de adsorção. Entretanto, a estabilidade desses complexos está relacionada às condições ambientais. Alterações de pH, temperatura e condições oxirredutoras podem promover a remobilização dos metais tornando-os facilmente hidrossolúveis e biodisponíveis (Queiroz *et al.*, 2018; Queiroz *et al.*, 2021b; Gabriel *et al.*, 2021).

Tendo em vista a enorme carga de sedimentos depositados ao longo da bacia do Rio Doce contendo minerais ferruginosos, e a forte presença de metais e metaloides, cria-se um cenário de risco para a população, recursos hídricos e ecossistemas aquáticos no interior da bacia, especialmente no baixo Rio Doce, e zona costeira/estuarina, pois as condições desses ambientes pode favorecer os mecanismos biogeoquímicos que governam a mobilidade dos metais entre água, sedimentos e cadeias tróficas (Queiroz *et al.*, 2018; Queiroz *et al.*, 2021b; Gabriel *et al.*, 2021).

Dessa forma, torna-se ainda mais importante a manutenção dos esforços de monitoramento de água e sedimentos na bacia do Rio Doce, pois deles depende não somente a avaliação da recuperação socioambiental da bacia, mas o acompanhamento da dinâmica espaço-temporal dos poluentes e a construção de eventuais estratégias necessárias ao enfrentamento de condições de contaminação crônica cujos efeitos podem se perpetuar por período indeterminado.

7.2.4 Configuração espacial de melhor desempenho para estações da rede do PMQQS

As subregiões de qualidade da água A, H, I, K, L e N não foram investigadas quanto à configuração espacial da rede por serem compostas por uma única estação cada. As subregiões F, G, J, K, L e M tiveram uma estação removida da análise cada. Optou-se pela remoção dessas estações por situarem-se em locais não impactados pela pluma de rejeitos, compreendendo, portanto, fonte de geração de dados das condições de referência em análises comparativas da evolução temporal da qualidade

da água na bacia. As estações descartadas foram RCR-01, RGU-01, RMA-01, RMH-01, RPG-01, RSA-01, RCA-01, RPC-01 e RSG-01.

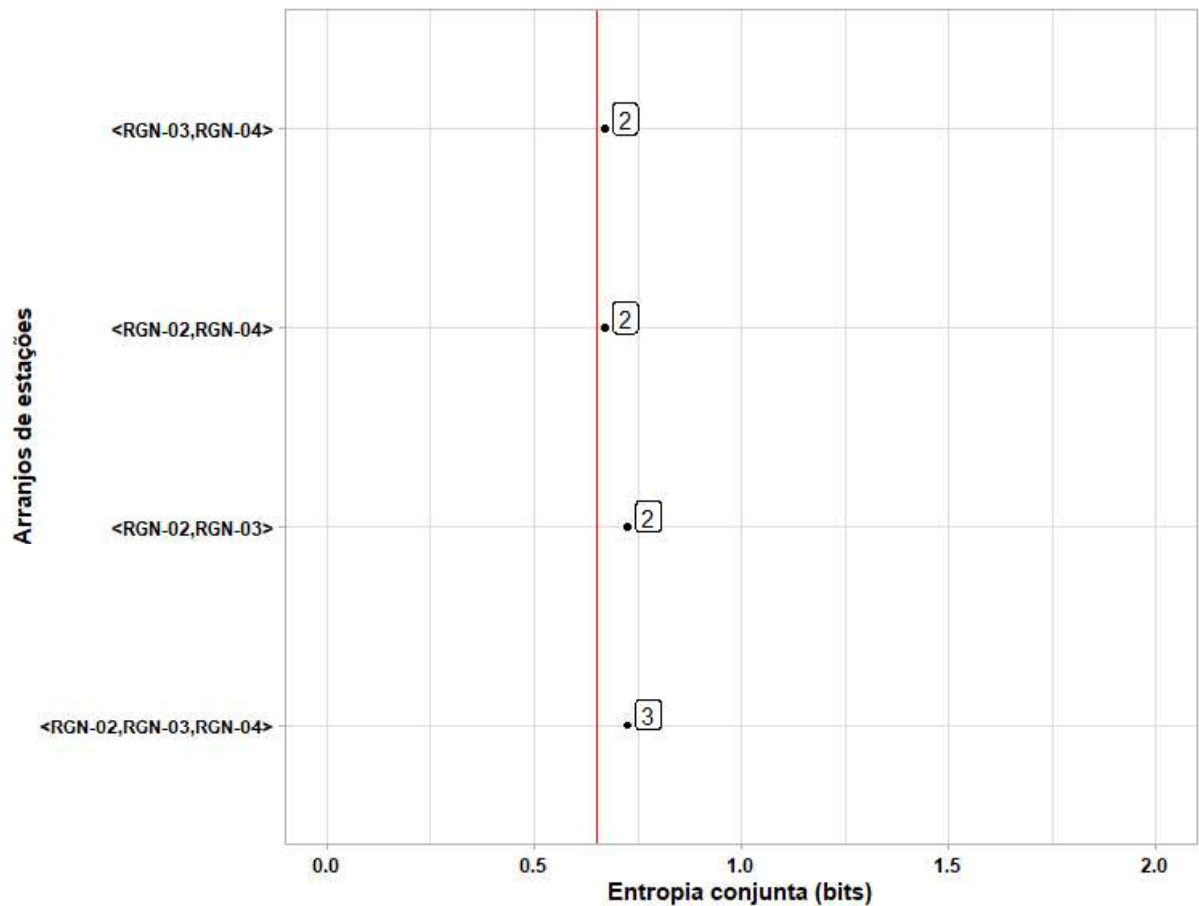
A análise de sensibilidade orientou a discretização das séries temporais dos valores do EWQI (Tabela 14). Cada série temporal das estações por sub-região, foi então, reclassificada, a partir desses critérios. As figuras A24 a A32 (Apêndice) ilustram a análise de sensibilidade realizada para determinação da largura dos *bins*.

Tabela 14 - Largura de *bins* utilizada na discretização dos valores do EWQI em todas as subregiões

Subregião	Largura do <i>bin</i>
B	163,4
C	377,6
D	92,7
E	90,6
G	95,1
J	103,8
M	123,7

A subregião B é formada pelo agrupamento das estações RGN-02, RGN-03 e RGN-04. Três diferentes possibilidades de arranjo alcançaram níveis de entropia conjunta superiores ao limiar de 90% definido como critério para seleção do melhor conjunto de estações. Apesar disso, a configuração baseada nas estações RGN-02 e RGN-03 apresentou valor sensivelmente superior, representando assim a solução ótima global. Nesse caso, a estação RGN-04 poderia ser descontinuada ou, na melhor hipótese realocada com objetivo de produzir informações menos redundantes e de maior valor para a rede (Figura 31).

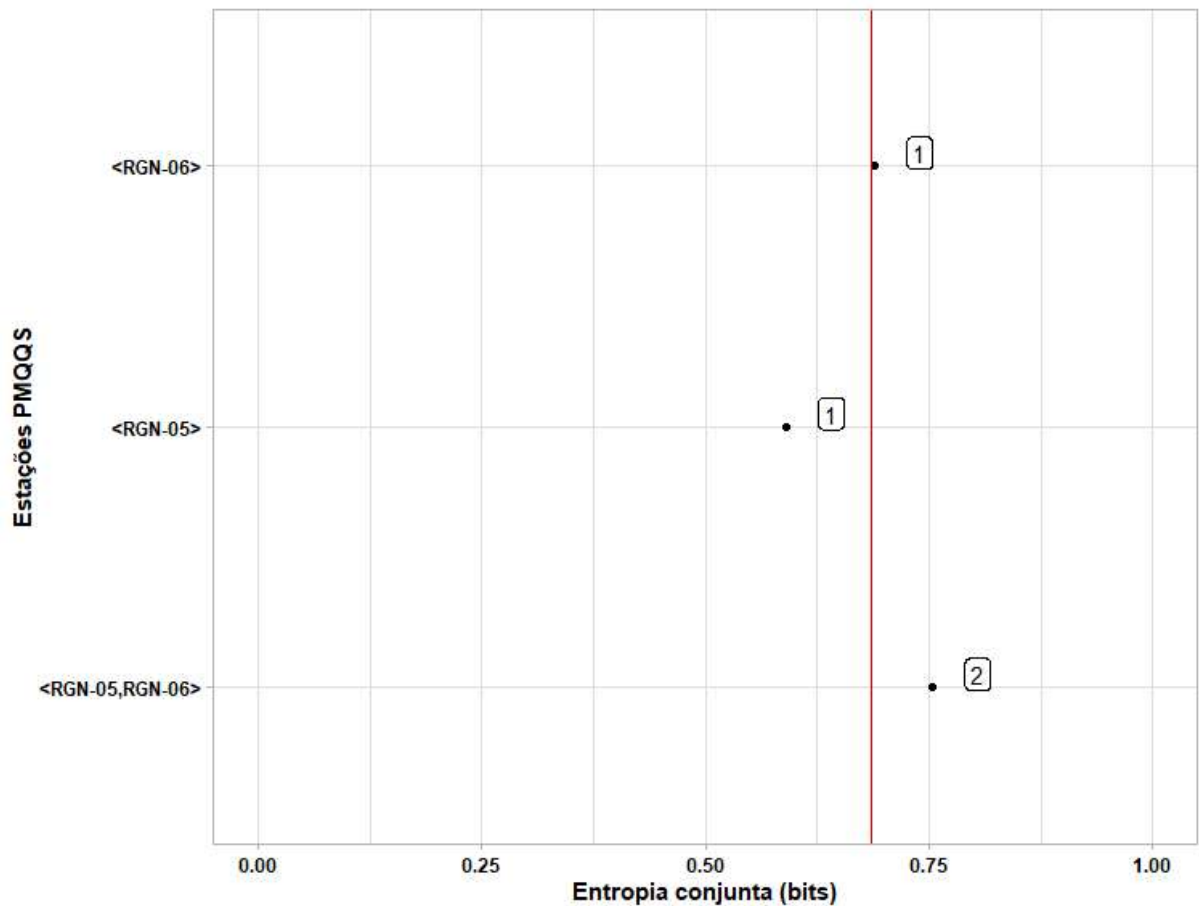
Figura 31 - Cenários para otimização da rede PMQSS na subregião B*, método FMOHBE/ função monoobjetivo 1



* A linha vermelha representa o valor de referência para seleção do arranjo ótimo da rede. Os números nas caixas representam o número de estações de cada cenário.

A investigação da subregião C demonstrou que a estação RGN-06 individualmente consegue capturar mais de 90% da informação total dessa sub-região. A estação RGN-05, por sua vez, mostrou-se menos informativa isoladamente. Nesse caso, a solução ótima global corresponde à manutenção da RGN-06 na rede e descontinuação ou realocação da estação RGN-05 (Figura 32).

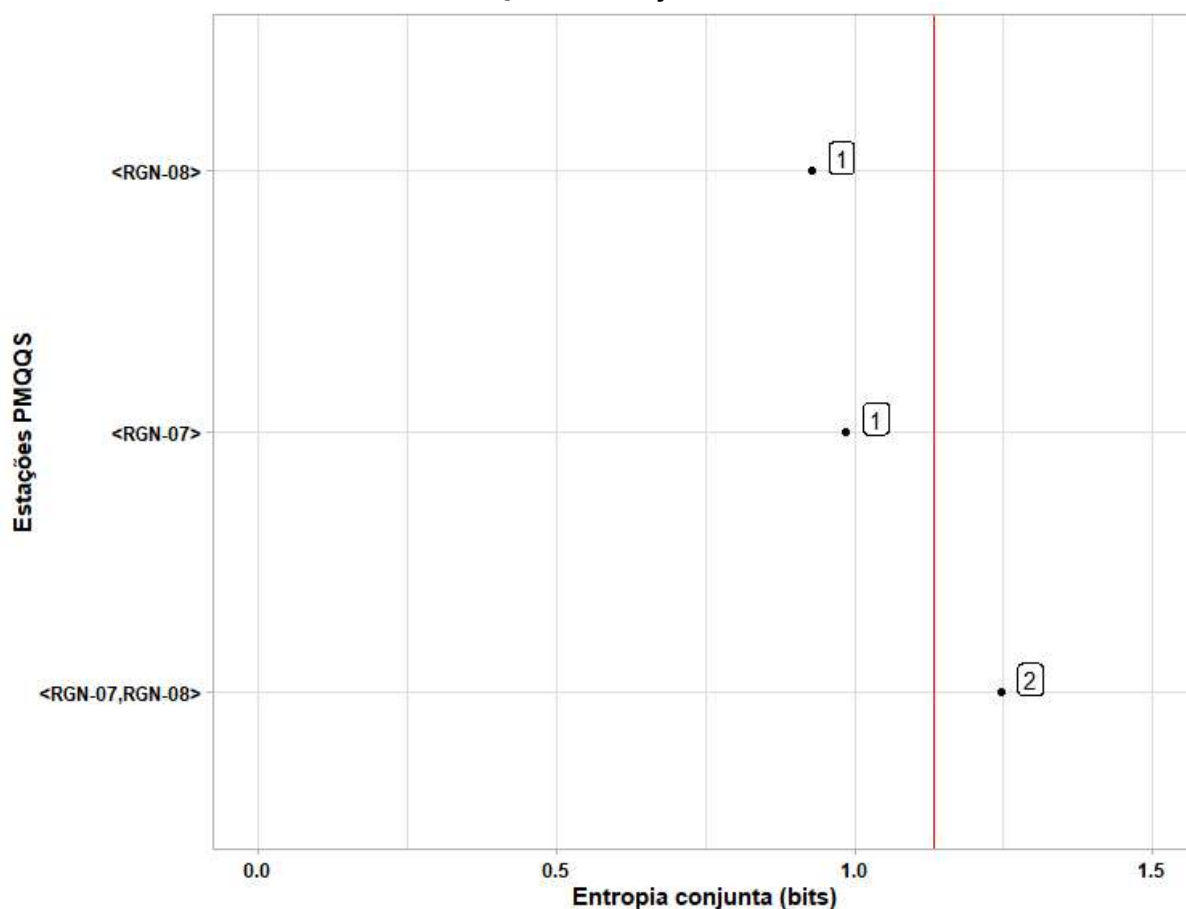
Figura 32 - Cenários para otimização da rede PMQQS na subregião C*, método FMOHBE/ função monoobjetivo 1



* A linha vermelha representa o valor de referência para seleção do arranjo ótimo da rede. Os números nas caixas representam o número de estações de cada cenário.

A subregião D, formada pelas estações RGN-07 e RGN-08, não apresentou um cenário ótimo que permitisse a identificação de estações redundantes. Individualmente nenhuma estação apresentou nível de entropia marginal superior a 90% da entropia conjunta das estações. Logo, essa subregião não oferece oportunidade para ajustes (Figura 33).

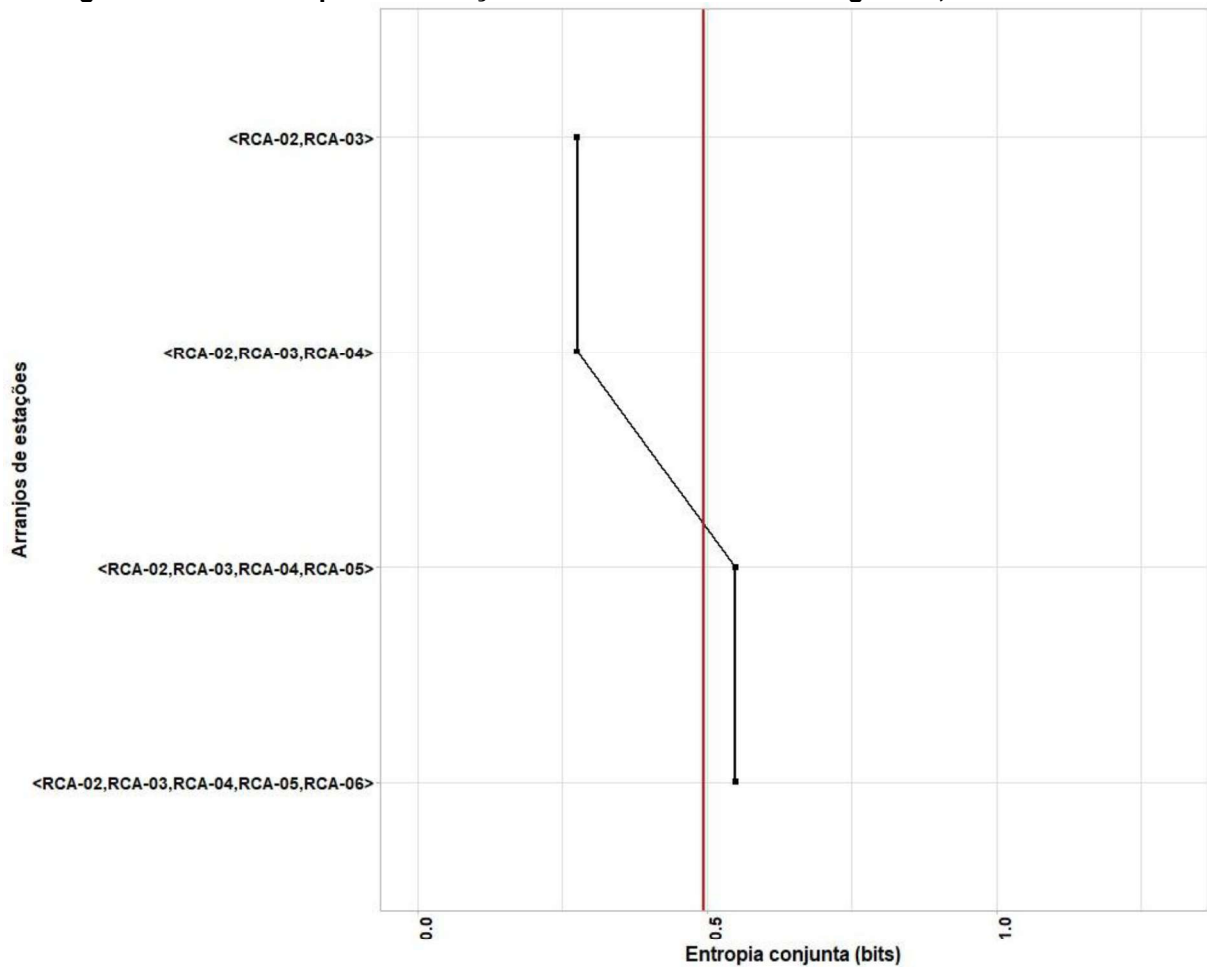
Figura 33 - Cenários para otimização da rede PMQQS na subregião D*, método FMOHBE/ função monoobjetivo 1



* A linha vermelha representa o valor de referência para seleção do arranjo ótimo da rede. Os números nas caixas representam o número de estações de cada cenário.

O método MGAMO obteve êxito em encontrar uma única solução ótima global para a subregião E, formada pelas estações RCA-02, RCA-03, RCA-04, RCA-05 e RCA-06. A única configuração espacial das estações que atendeu à restrição estabelecida compreendeu o conjunto das estações RCA-02, RCA-03, RCA-04 e RCA-05, constituindo, portanto, uma solução em que apenas uma estação deveria ser descontinuada a fim de alcançar a configuração de melhor desempenho para a rede (Figura 34).

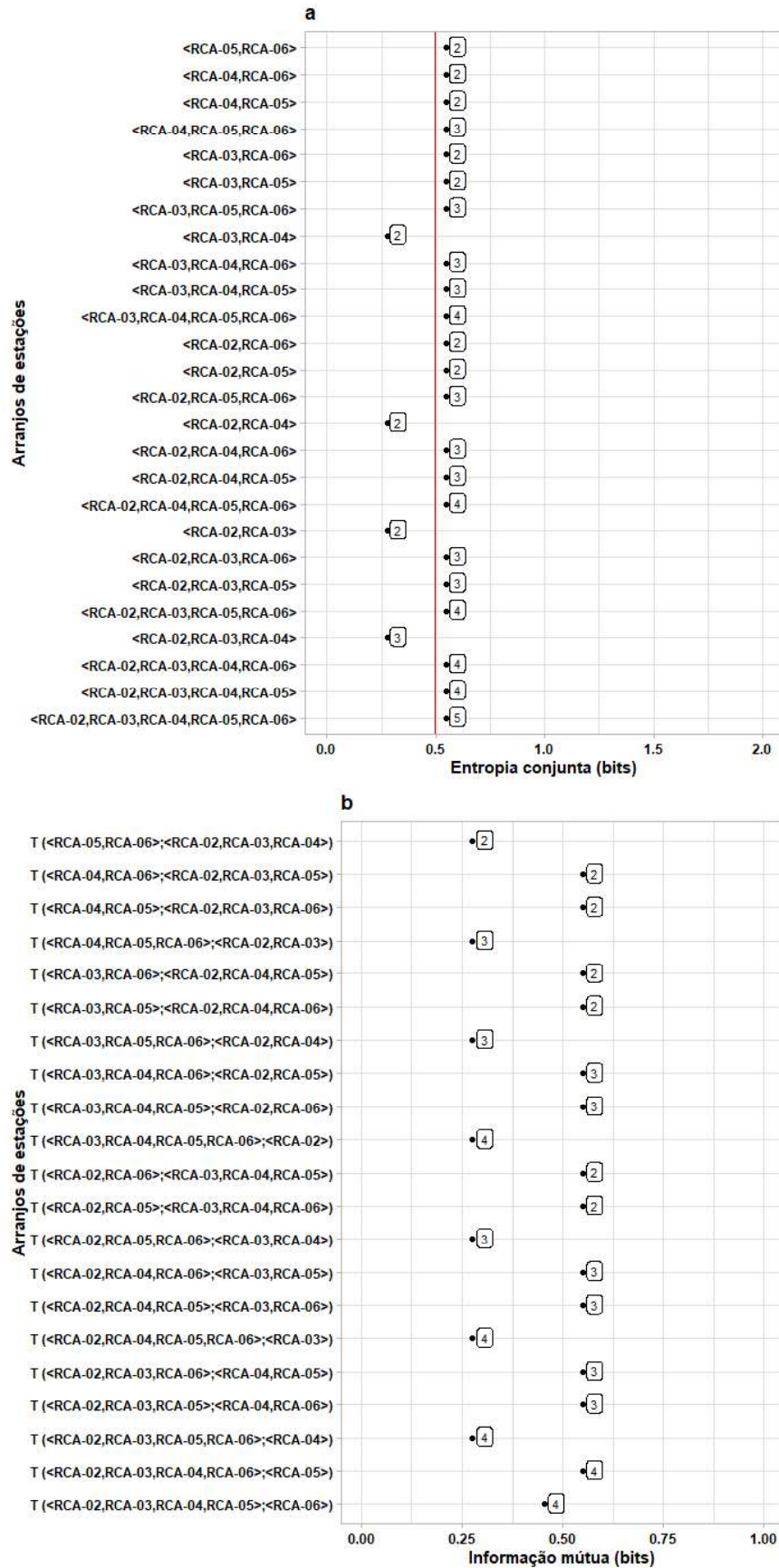
Figura 34 - Cenários para otimização da rede PMQQS na subregião E*, método MGAMO



* A linha vermelha representa o valor de referência para seleção do arranjo ótimo da rede.

Na investigação do método FMOHBE, na mesma sub-região, os resultados foram completamente divergentes. Inicialmente a implementação do algoritmo de otimização tomando os critérios estabelecidos para solução ótima reportaram não apenas uma, mas 22 diferentes arranjos que atenderam à restrição relativa à entropia conjunta (Figura 35a). Dessa forma, o método passou à otimização da próxima função monoobjetivo hierarquicamente definida, onde a solução ótima deve maximizar a transinformação (Figura 35b).

Figura 35 - Cenários para otimização da rede PMQQS na subregião E, método FMOHBE/ função monoobjetivo 1 (a), 2 e 3 (b)



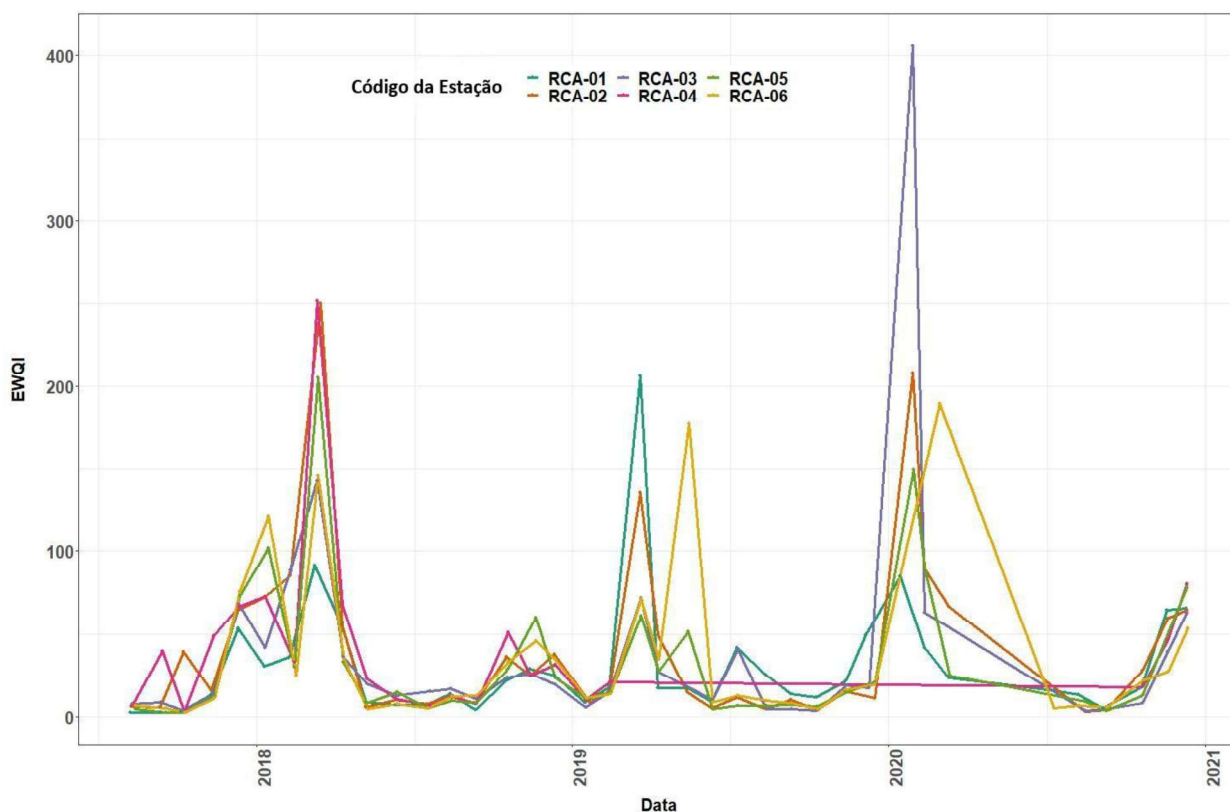
*** A linha vermelha representa o valor de referência para seleção do arranjo ótimo da rede. Os números nas caixas representam o número de estações de cada cenário.**

De forma semelhante, a segunda função monoobjetivo não obteve êxito em encontrar uma solução ótima global. Ao contrário, mais uma vez ela retornou vários cenários candidatos ao arranjo ideal, nenhum prevalecendo sobre os demais (Figura 35b). Ao total 13 cenários potenciais apresentaram o melhor desempenho com o mesmo valor de transformação. Dessa forma, a terceira e última função monoobjetivo hierarquizada passou a ser utilizada pelo método FMOHBE, objetivando minimizar o número total de estações (Figura 35b).

O algoritmo de otimização, em relação à função monoobjetivo 3, igualmente não retornou uma solução ótima global, uma vez que 7 configurações elegíveis à melhor configuração formadas por pares de estações foram obtidas. Dessa forma, o método FMOHBE isoladamente não foi suficiente para encontrar a solução representativa do conjunto ótimo de estações na sub-região, necessitando de critérios adicionais para esse fim.

Optou-se então por assumir critérios relacionados à distribuição espacial e o comportamento das séries temporais como acessórios para seleção do cenário de melhor desempenho. Das séries temporais plotadas pode-se notar que há relativa superposição das linhas que representam a evolução temporal do EWQI. Isso sugere que uma proporção significativa da informação capturada pelas estações é redundante, isto é, várias das estações dessa subregião estão capturando as mesmas informações concomitantemente (Figura 36).

Figura 36 - Séries temporais das estações da subregião E



Dentre todas as alternativas elencadas pelo método FMOHBE como elegíveis as que possuem a melhor relação do número de estações, ou seja, aquelas que conjuntamente agregam o maior conteúdo informativo, tem a maior capacidade de transferência de informação e constituem arranjos com o menor número de estações são os arranjos formados por duas estações. Essas soluções são bastante atrativas, uma vez que, em tese permitiriam a descontinuação ou realocação de no mínimo 3 estações do conjunto atual.

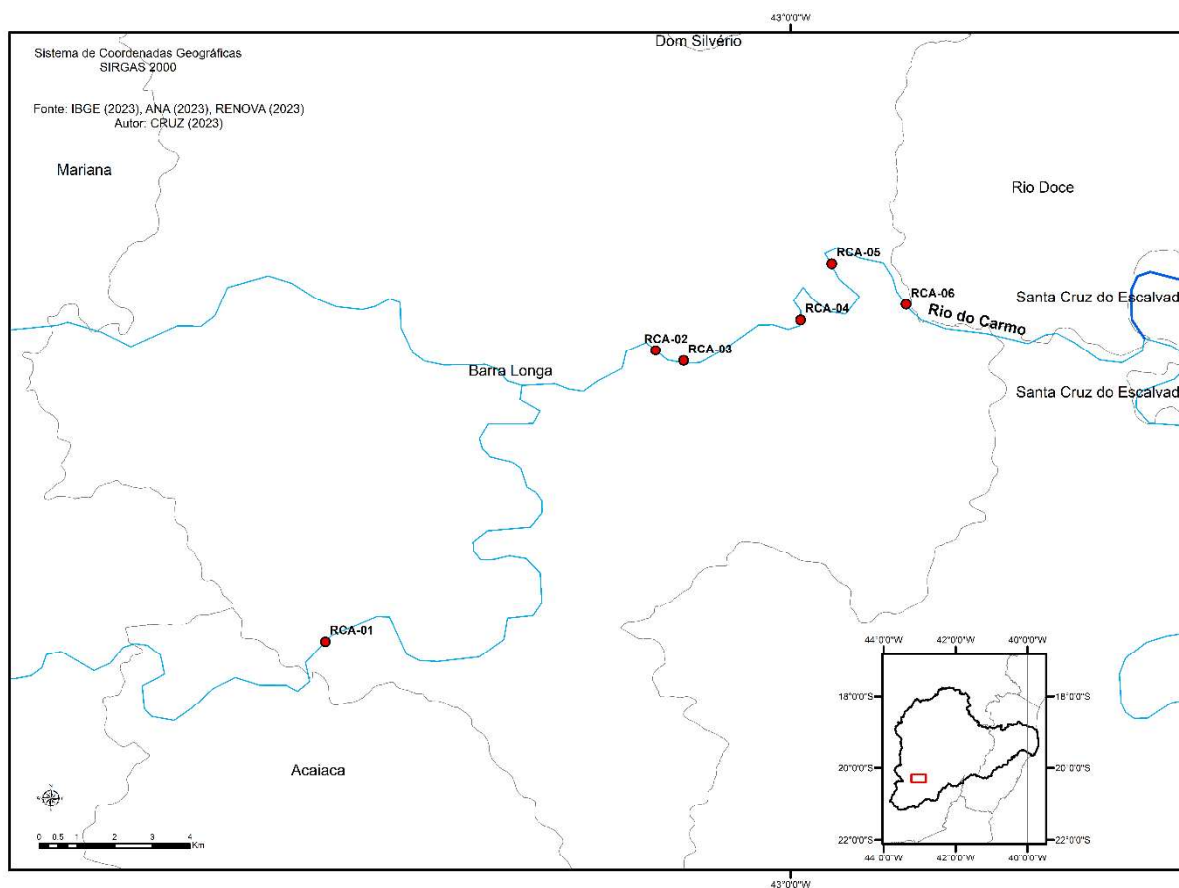
Nessa perspectiva, a solução aparentemente mais apropriada seria manter as estações RCA-02 e RCA-06, espacialmente distribuídas nos extremos da sub-região, e descontinuar as demais. Além disso, ao aplicar o teste de Kruskal-Wallis a 0,05 de significância estatística, tomando as séries temporais de EWQI não discretizados das estações RCA-03 a RCA-05, conclui-se não haver diferenças estatisticamente significativas entre as estações (p -valor=0,61), corroborando a redundância aparentemente sugerida no comportamento temporal de EWQI associado às estações.

Entretanto, ao analisar a disposição espacial das estações nesse cenário nota-se que haveria uma região muito grande entre as estações não cobertas pelo monitoramento. Dessa forma, ao adicionar no mínimo mais uma estação ao par, apesar de possivelmente não haver ganho de informação no cômputo total, haveria maior densidade espacial de estações. Essa solução embora um pouco mais redundante apresenta impacto positivo a rede, vide a cobertura que promove na eventual ocorrência de falhas no monitoramento periódico, conforme destacado por (Foroozand; Weijs, 2021)

As estações RCA-02 e RCA-03 são relativamente próximas, umas das outras, com grande chance de terem elevado grau de similaridade quanto às variáveis monitoradas. Nesse caso, pelo fato da estação RCA-02 estar mais próxima da confluência do Rio Gualaxo do Norte, seria conveniente mantê-la uma vez que ela em tese pode registrar o efeito desse manancial na mistura com as águas do Rio do Carmo. As estações RCA-04, RCA-05 e RCA-06 estão no fim desse trecho do Rio do Carmo aproximadamente equidistantes entre si. Como não há nenhuma confluência importante nesse trecho, eliminar a estação mais ao centro parece ser coerente (Figura 37).

Dessa forma o arranjo ideal proposto para essa subregião pelo método FMOHBE com critérios adicionais é compreendido pelas estações RCA-02, RCA-04 e RCA-06.

Figura 37 – Distribuição espacial das estações situadas na subregião E



Ao comparar as alternativas derivadas da aplicação de ambos os métodos de otimização investigados para a subregião E, nota-se que, embora, ambas possuam o mesmo grau de entropia conjunta, em termos de transinformação a alternativa escolhida pelo método FMOHBE com critérios adicionais se saiu melhor, representando um arranjo com maior potencial de transferência de informação. Adicionalmente, o método FMOHBE também se saiu melhor quanto ao número de estações, uma vez que a solução ideal indicada pelo método é constituída de 3 estações, ao passo que a solução de MGAMO possui 5 estações.

A não obtenção de uma única solução ótima global, gerando a necessidade de incorporar critérios adicionais deve-se à forte similaridade no comportamento de EWQI entre o conjunto de estações. Apesar de ter-se chegado a uma solução satisfatória pelo uso de critérios adicionais, a investigação dessa subregião demonstrou que redundância em proporção elevada entre conjuntos de estações impõe uma limitação ao método FMOHBE que, como demonstrado, pode não

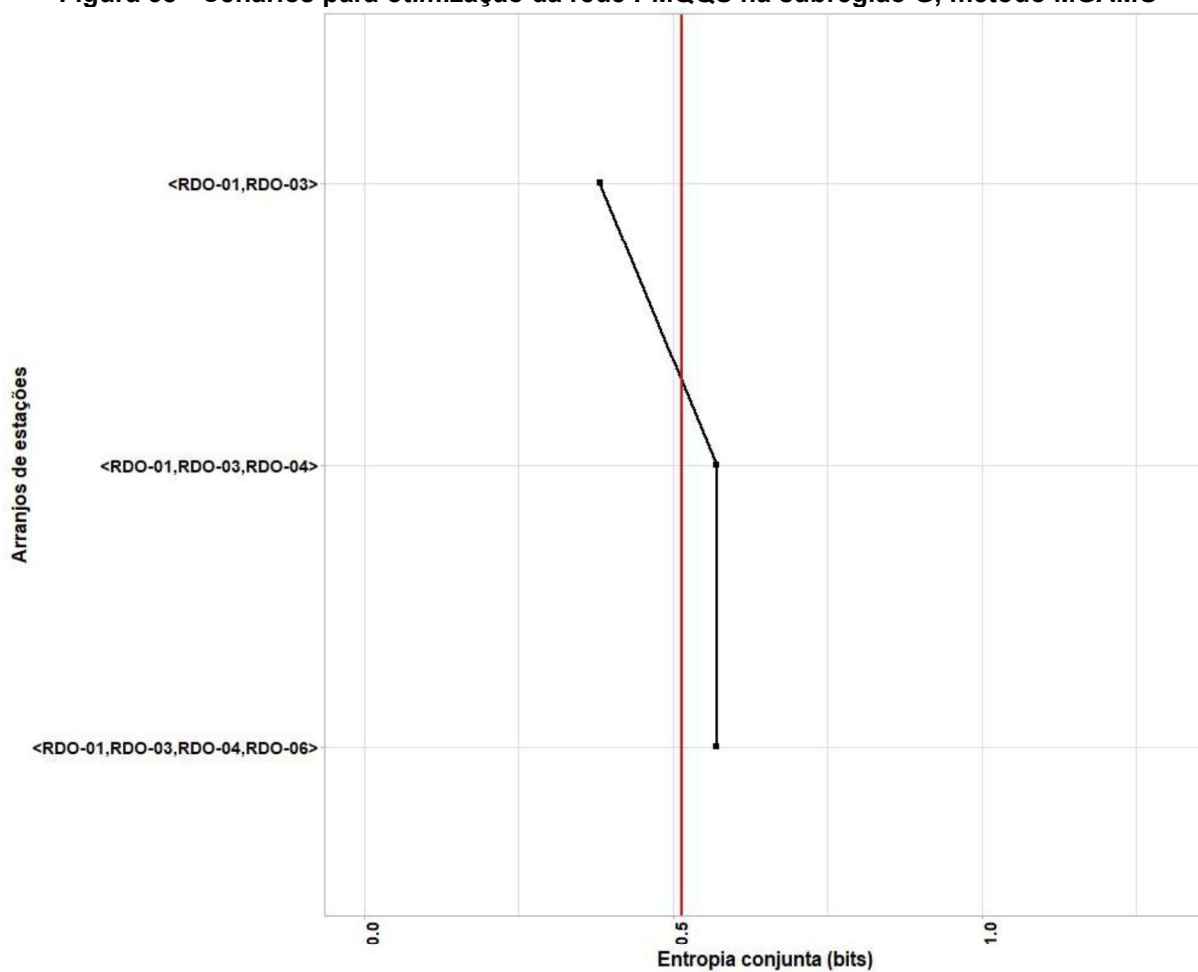
encontrar a solução ótima global de maneira objetiva e despida de viés, como idealmente deve ser.

Não obstante, a investigação da subregião E revelou que o método MGAMO ao adicionar progressivamente uma nova estação a cada rodada da implementação do algoritmo guloso acaba não conseguindo avaliar alternativas de arranjos de estações que não estejam geograficamente próximas. Esse fato por um lado adiciona viés à análise, uma vez que o sentido escolhido para a investigação, isto é, de montante à jusante ou de jusante à montante, passa a ser determinante para a seleção da configuração espacial ótima das estações. Por outro lado, também gera propensão à obtenção de soluções ótimas globais menos parcimoniosas em termos de quantitativo de estações, caso as estações mais informativas não estejam nos primeiros postos de acordo com o sentido em que se dará a agregação das estações, como demonstrado.

A investigação da subregião G pelo método MGAMO indicou a solução ótima global após a segunda rodada de otimização. O cenário mais promissor compreendeu o arranjo formado pelas estações RDO-01, RDO-03 e RDO-04 (Figura 38). FMOHBE ao investigar a mesma subregião necessitou a implementação do algoritmo de otimização de todas as funções monoobjetivo. A otimização da primeira função revelou que 4 cenários distintos atenderam à restrição inserida como condição para solução ideal (Figura 39a). A otimização da segunda função, por sua vez, restringiu as soluções a 3, entretanto, todas foram associadas ao mesmo conteúdo de informação transferido, ilustrado pela transinformação. A terceira função, contudo, indicou uma única solução ótima global formada pelas estações RDO-01 e RDO-04 (Figura 39b).

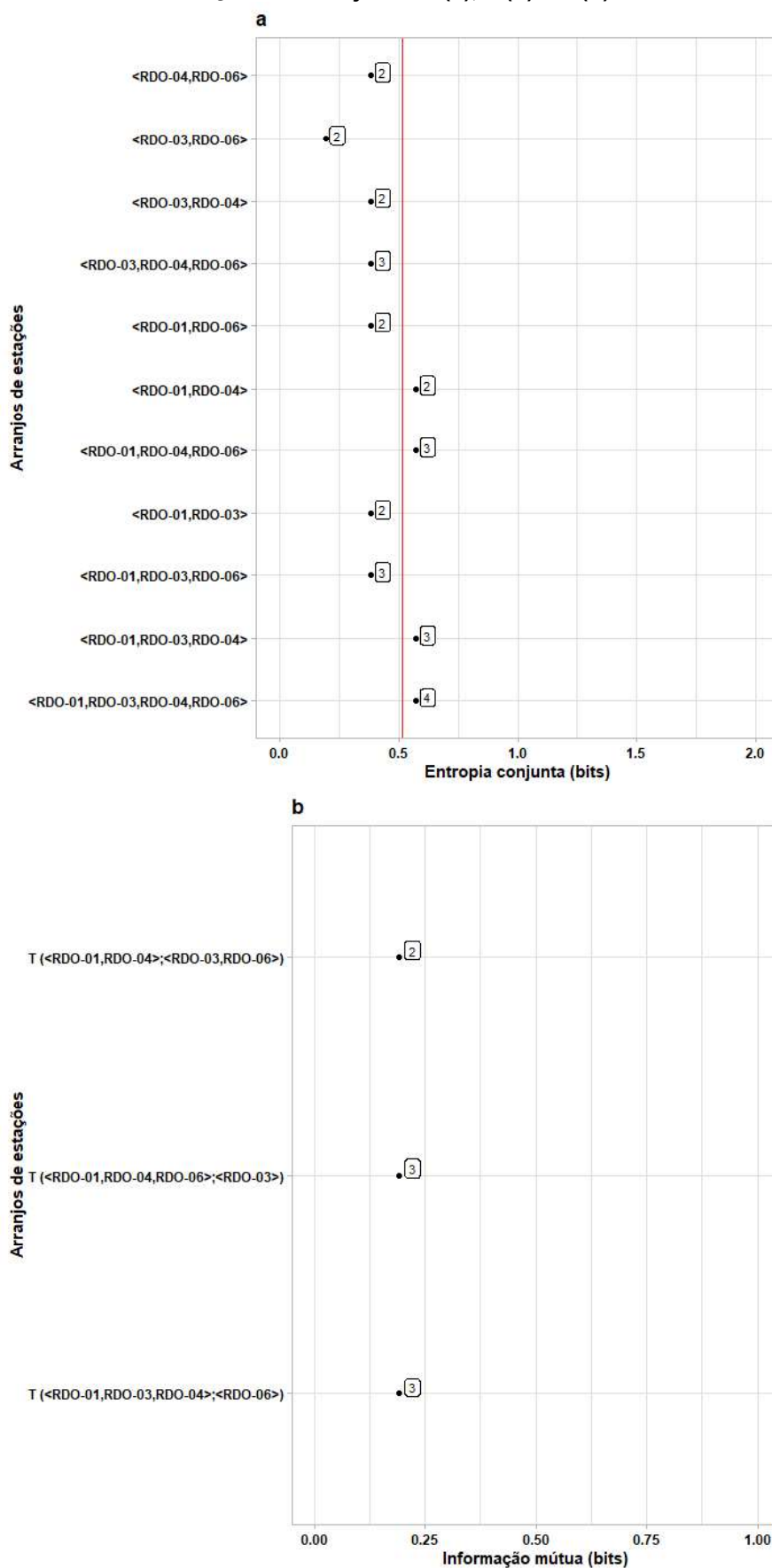
Novamente o método FMOHBE foi mais eficiente e indicou a solução mais plausível, pois embora ambas as alternativas apresentadas pelos métodos possuam o mesmo conteúdo de informação relativa à entropia conjunta e transinformação, a solução ótima global de FMOHBE é mais parcimoniosa, contendo apenas duas estações (Figura 39). Nesse caso, as estações RDO-01 e RDO-06 poderiam ser descontinuadas ou realocadas, sem alterações no conteúdo de informação capturado conjuntamente pelo arranjo ótimo de estações indicado por FMOHBE.

Figura 38 - Cenários para otimização da rede PMQQS na subregião G, método MGAMO



* A linha vermelha representa o valor de referência para seleção do arranjo ótimo da rede.

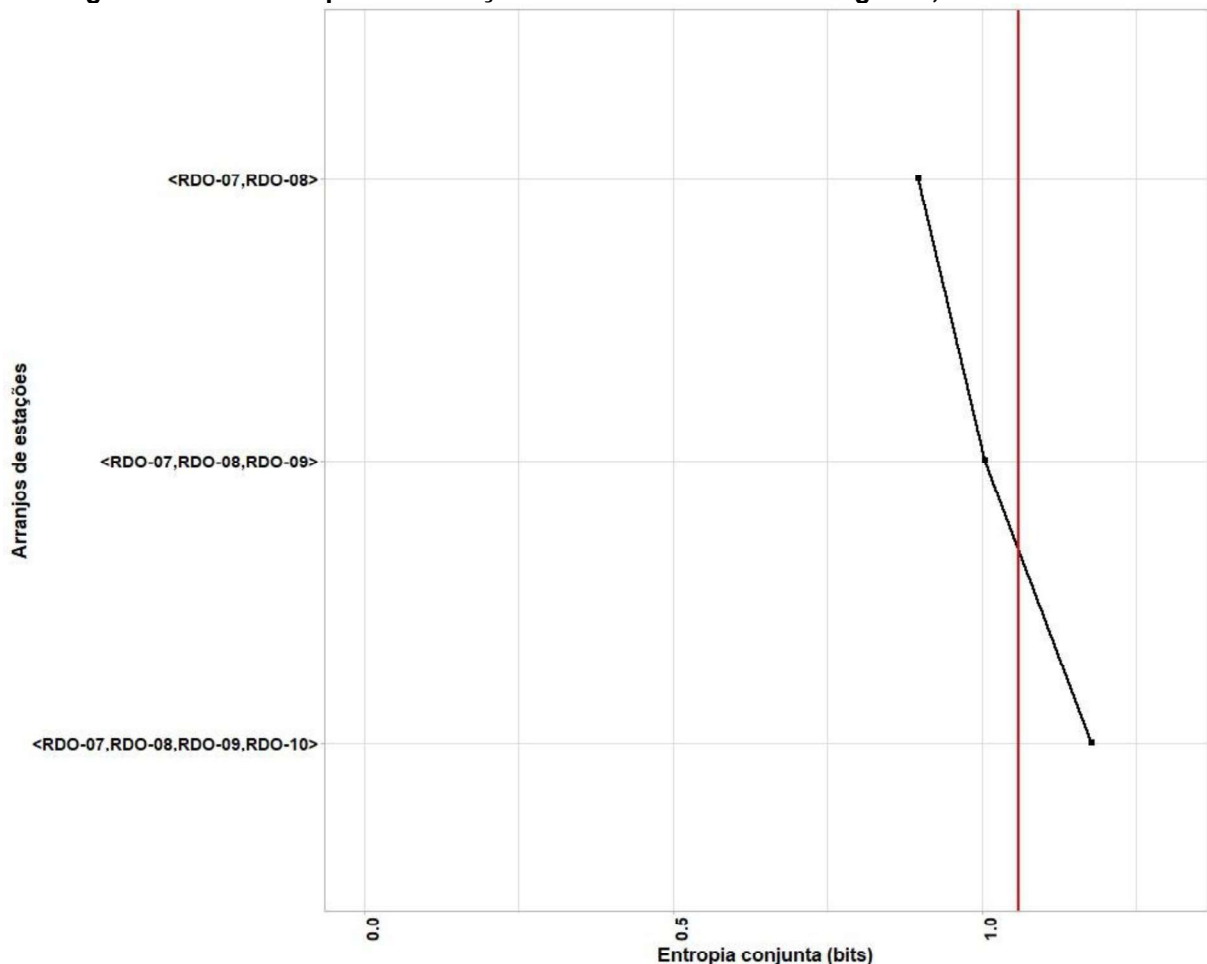
Figura 39 - Cenários para otimização da rede PMQQS na subregião G, método FMOHBE/ função monoobjetivo 1 (a), 2 (b) e 3 (b)



* A linha vermelha representa o valor de referência para seleção do arranjo ótimo da rede. Os números nas caixas representam o número de estações de cada cenário.

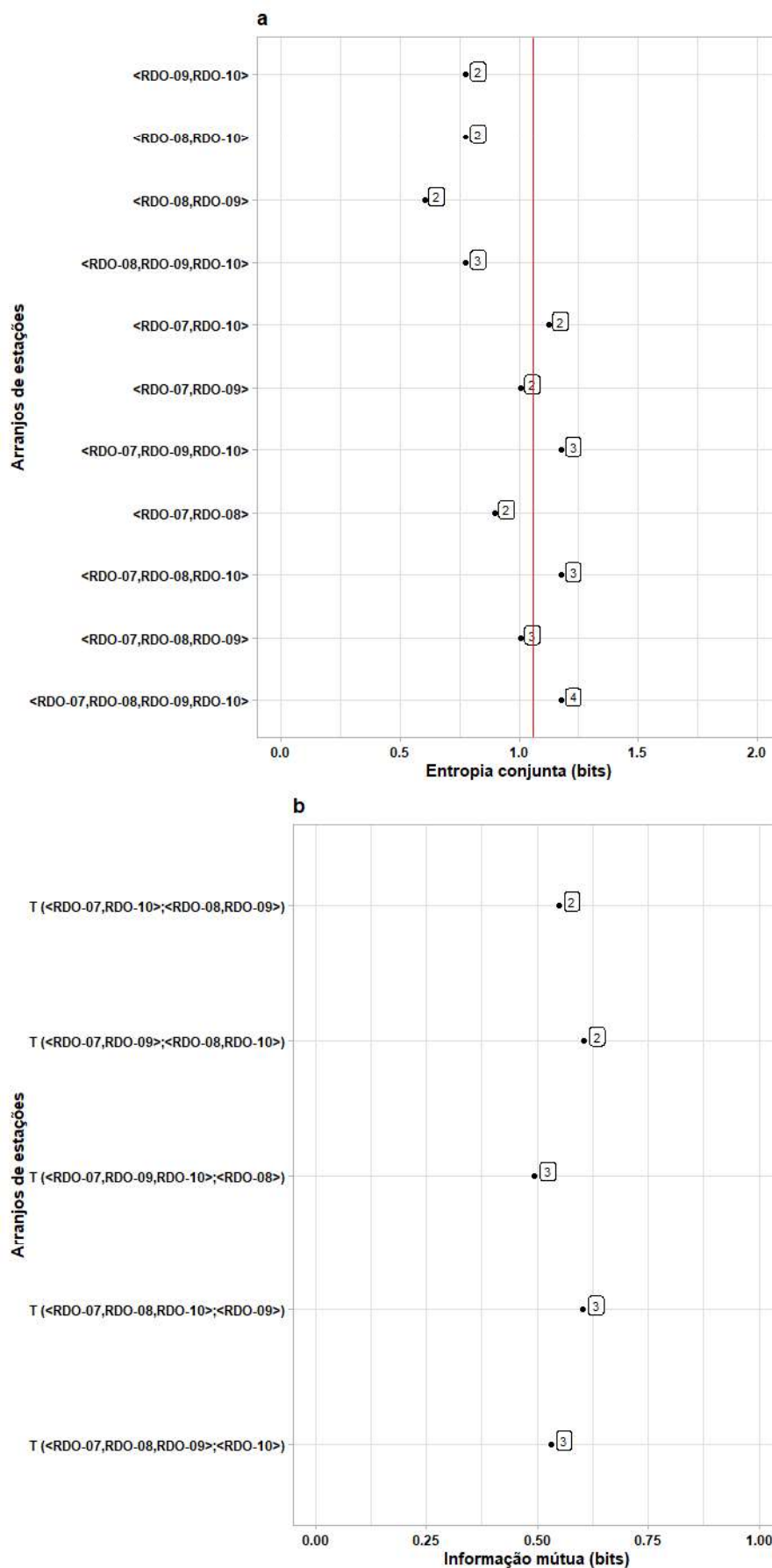
A investigação do método MGAMO não obteve êxito na busca de uma única solução ótima global, na subregião J (Figura 40). Nenhuma das alternativas atendeu ao critério de seleção estabelecido como restrição à otimização. FMOHBE, entretanto, após a implementação do algoritmo de otimização sobre as função monoobjetivo 1 e 2 indicou o arranjo formado pelas estações RDO-07, RDO-08 e RDO-10 como solução ótima global (Figuras 41a e 41b).

Figura 40 - Cenários para otimização da rede PMQQS na subregião J, método MGAMO



* A linha vermelha representa o valor de referência para seleção do arranjo ótimo da rede.

Figura 41 - Cenários para otimização da rede PMQQS na subregião J, método FMOHBE/ função monoobjetivo 1 (a), 2 e 3 (b)

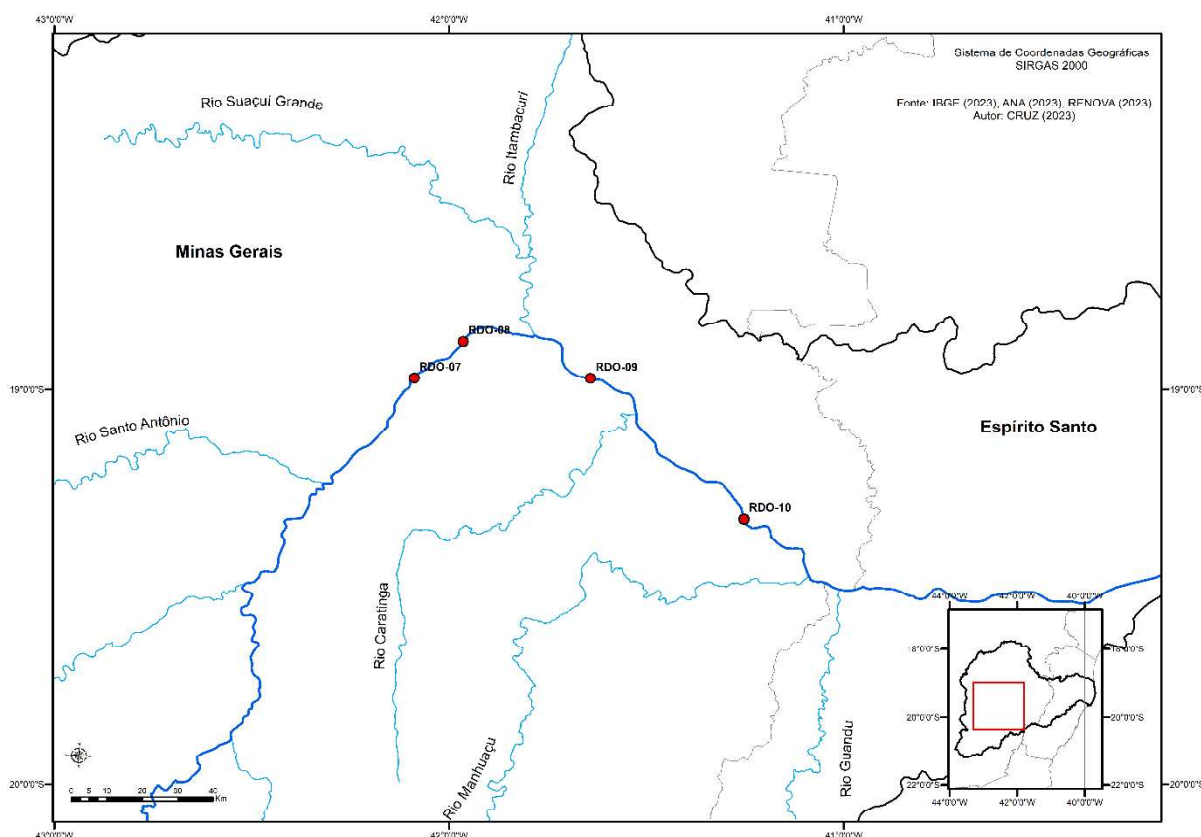


*** A linha vermelha representa o valor de referência para seleção do arranjo ótimo da rede. Os números nas caixas representam o número de estações de cada cenário.**

Apesar de ambas as alternativas elegíveis a melhor configuração de rede, apresentadas por FMOHBE, possuírem a mesma entropia conjunta e o mesmo número de estações, o arranjo composto pelas estações RDO-07, RDO-08 e RDO-10 proveu a alternativa com o melhor desempenho na transferência de informações e, portanto melhor capacidade preditiva, fato esse destacado por Singh (2013) como um dos principais objetivos na operação de redes de monitoramento da qualidade da água, logo após a maximização da informação conjunta.

Embora, a investigação dos métodos para a subregião J não tenha demandado nenhum critério adicional, a título de exercício foi realizada uma análise espacial do impacto da solução considerada ideal para a configuração da rede, a fim de investigar mais a fundo se as estratégias baseadas em otimização considerando funções objetivo associadas a determinação de diferentes formas da entropia de Shannon são consistentes. Ao descontinuar a estação RDO-09 um trecho relativamente extenso do Rio Doce estaria descoberto de monitoramento (Figura 42). Esse fato, isoladamente, poderia motivar a decisão pela permanência da estação RDO-09 na rede dessa sub-região, a despeito da solução de melhor conteúdo informacional selecionada.

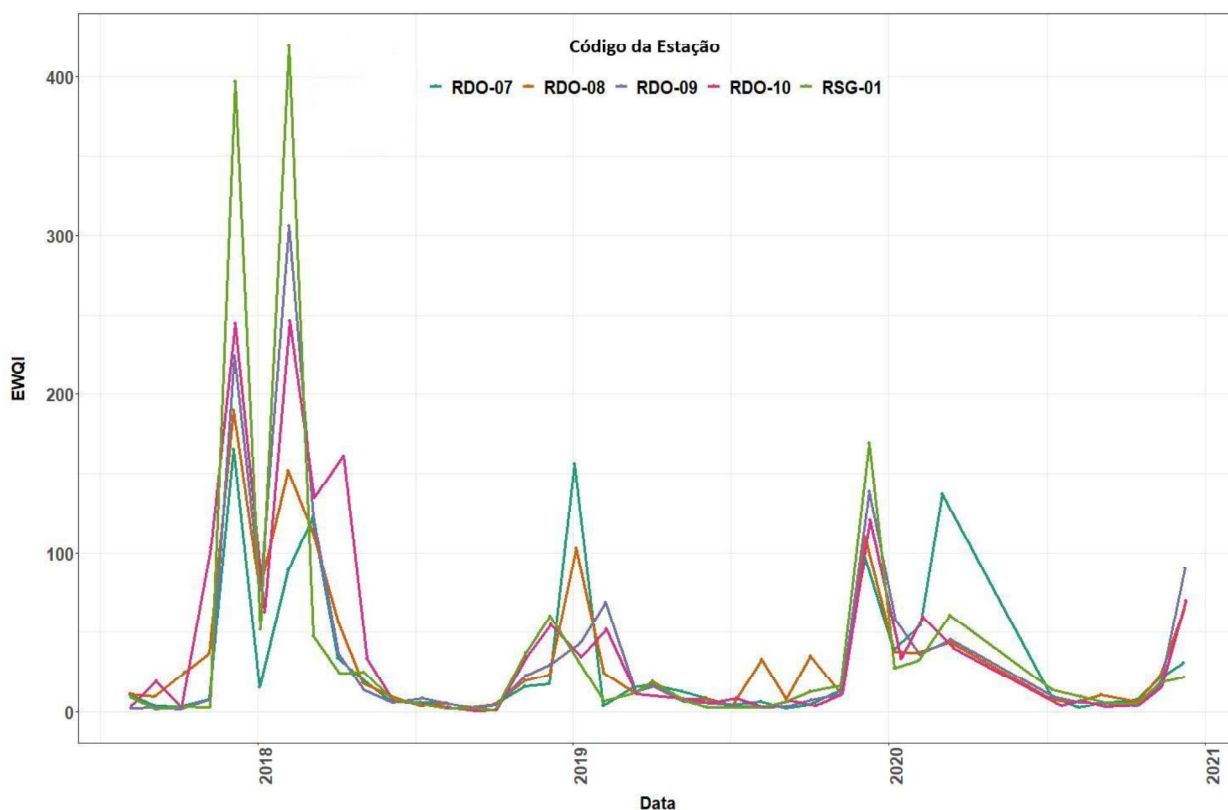
Figura 42 - Distribuição espacial das estações situadas na subregião J



Apesar desse contexto, de fato a operação dessa estação não oferece ganho líquido de informação, uma vez que: i) conforme pode-se observar do gráfico de séries temporais de EWQI há grande superposição entre o comportamento temporal de EWQI dessa estação com as demais, sugerindo haver forte redundância entre elas (Figura 43); ii) O teste de Mann-Whitney a 5% de significância estatística aplicado sobre as séries temporais não discretizadas de EWQI para comparação entre as estações RDO-08/RDO-09 e RDO-09/RDO-10 revelou a não existência de elementos suficientes para rejeição da hipótese nula de igualdade das medianas (p -valor [RDO-08/RDO-09] = 0.2638; p -valor [RDO-09/RDO-10] = 0.7951), confirmando, não haver diferenças estatisticamente significativas entre as estações.

Isso confirma a elevada redundância da estação RDO-09 vinculado ao registro periódico de informações repetidas, igualmente capturadas pelas estações vizinhas. Dessa forma, talvez a alternativa mais plausível de forma a manter um nível razoável de densidade espacial das estações nessa subregião seja realocar a estação ao invés de descontinuar-la.

Figura 43 - Séries temporais das estações da subregião J



O exercício realizado, confirma o desempenho superior de FMOHBE como ferramenta de apoio a seleção dos arranjos de melhor desempenho no redimensionamento de redes de monitoramento, comparativamente ao método MGAMO. Sem embargo, ele também revela que a análise espacial é bem-vinda como critério complementar na seleção do arranjo das redes com melhor desempenho.

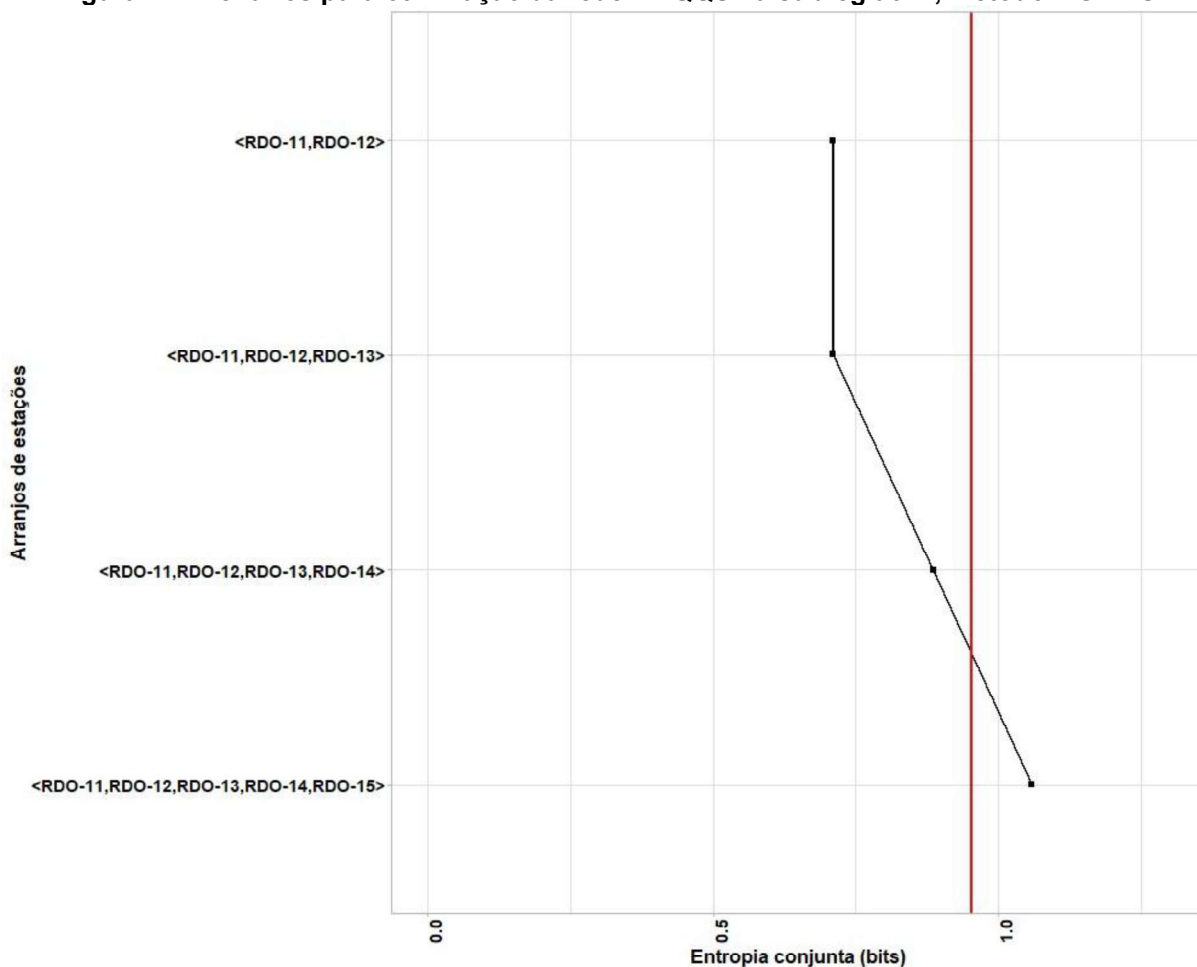
Como demonstrado, a análise espacial das estações, como critério único de avaliação de desempenho, através de adensamento ou redistribuição arbitrária de estações de maneira a tornar seu arranjo mais uniforme espacialmente não fornece garantias de ganho de informação útil para o gerenciamento da qualidade da água.

Em outras palavras, manter uma estação em operação apenas como forma de garantir que uma grande porção da rede hidrográfica não fique desprovida de monitoramento, caso da estação RDO-09 na subregião J, as vezes pode representar apenas o atendimento a suposta necessidade de cobertura de monitoramento vindo de uma percepção intuitiva, que parece basear-se em critérios de ordem técnica quando na verdade é permeada de subjetividade e sujeita a juízos pessoais.

Ao contrário, métodos de otimização baseados na análise de formas variadas de medida de informação, como o método FMOHBE, são vantajosos pois baseiam-se exclusivamente em critérios de análise da informação extraída diretamente dos dados obtidos pela operação da rede, sendo, portanto, desprovidos de viés.

Quanto à subregião M, a otimização por MGAMO não indicou nenhuma alternativa que atendesse ao critério para seleção dos cenários ideais para o redimensionamento da rede nessa subregião (Figura 44). O método FMOHBE, por outro lado demonstrou haver 4 alternativas potenciais de arranjos com igual entropia conjunta na sub-região, logo, demandando a segunda etapa de otimização visando maximizar a transinformação (Figuras 45a e 45b). Da mesma forma, não foi encontrada uma única solução ótima global nessa etapa e nem na otimização da última função objetivo, visto que os cenários mais promissores, apresentaram o mesmo número de estações, arranjos de pares de estações (Figuras 45b).

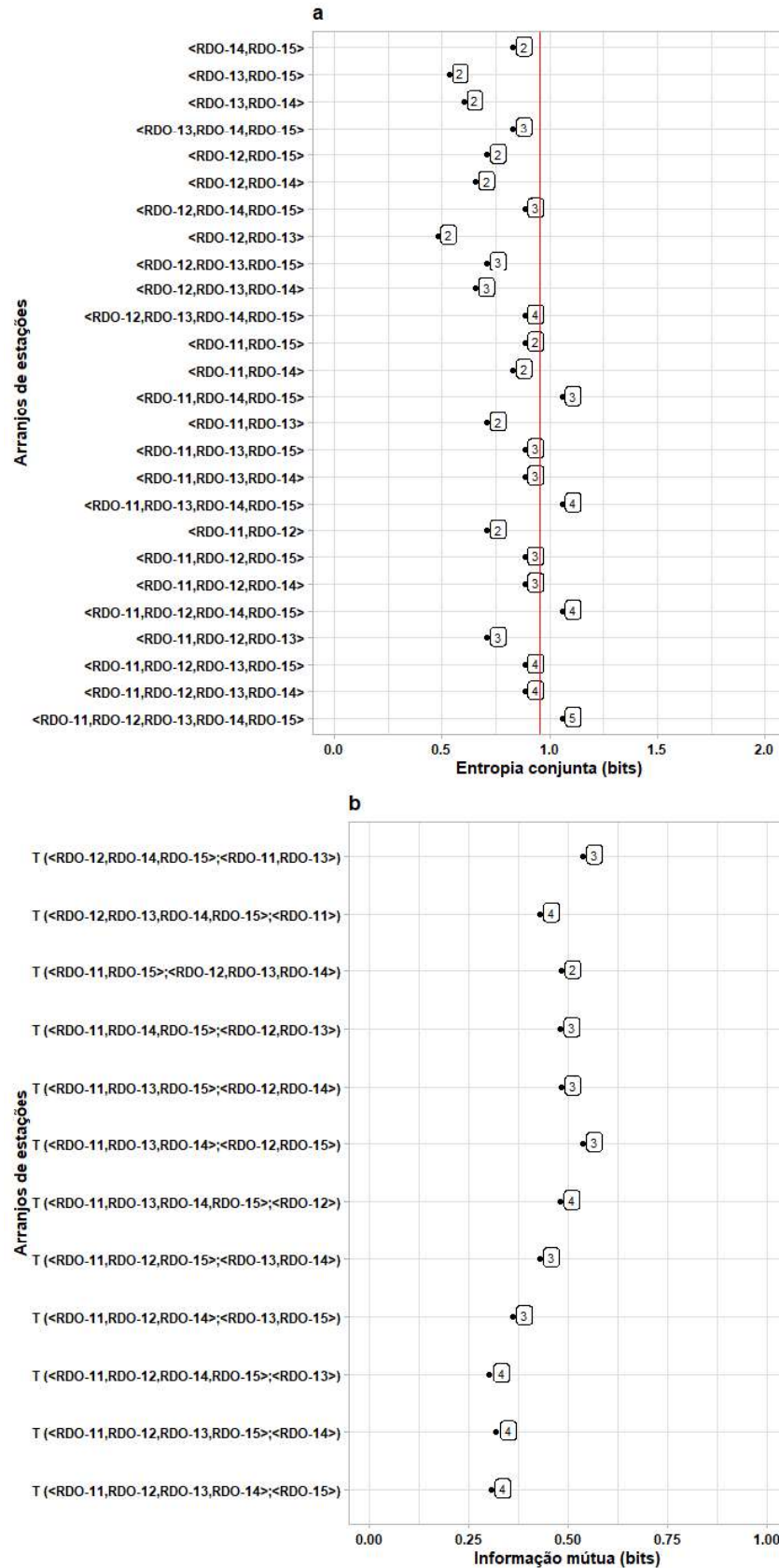
Figura 44 - Cenários para otimização da rede PMQQS na subregião M, método MGAMO



*** A linha vermelha representa o valor de referência para seleção do arranjo ótimo da rede.**

Procedeu-se então, tal como na subregião E, a análise espacial da distribuição das estações e das séries temporais da subregião como critério auxiliar para a seleção da melhor configuração da rede. A alternativa referente às estações RDO-11, RDO-13 e RDO-14 parece ter maior potencial para descarte, uma vez que em termos espaciais a proximidade entre as estações RDO-13 e RDO-14 pode colaborar para ambas as estações capturem as mesmas informações de qualidade da água (Figura 46). Essa hipótese ainda é fortemente corroborada pelo padrão apresentado no gráfico de séries temporais que demonstra superposição das linhas (Figura 47).

Figura 45 - Cenários para otimização da rede PMQQS na subregião M, método FMOHBE/ / função monoobjetivo 1 (a), 2 e 3 (b)



* A linha vermelha representa o valor de referência para seleção do arranjo ótimo da rede. Os números nas caixas representam o número de estações de cada cenário.

Figura 46 - Distribuição espacial das estações situadas na subregião M

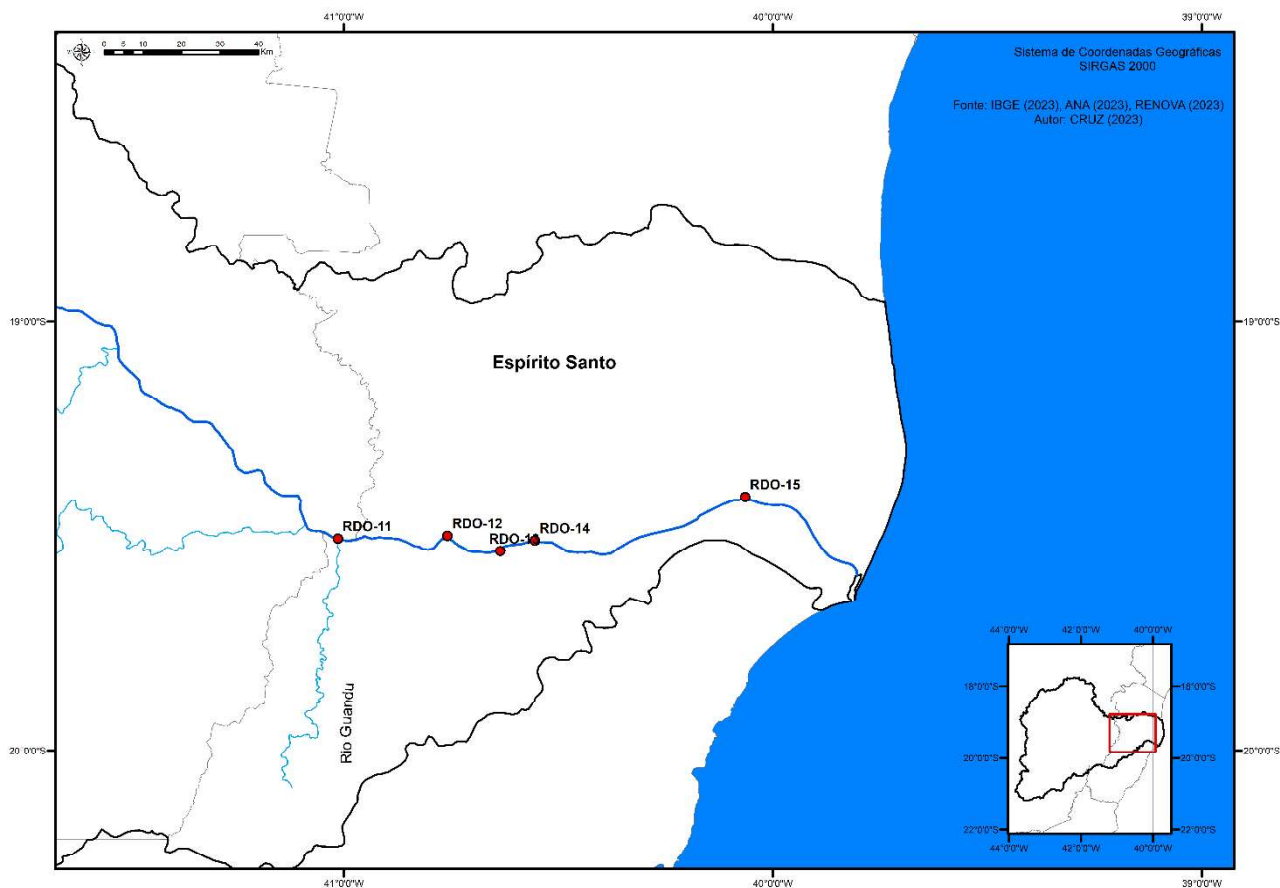
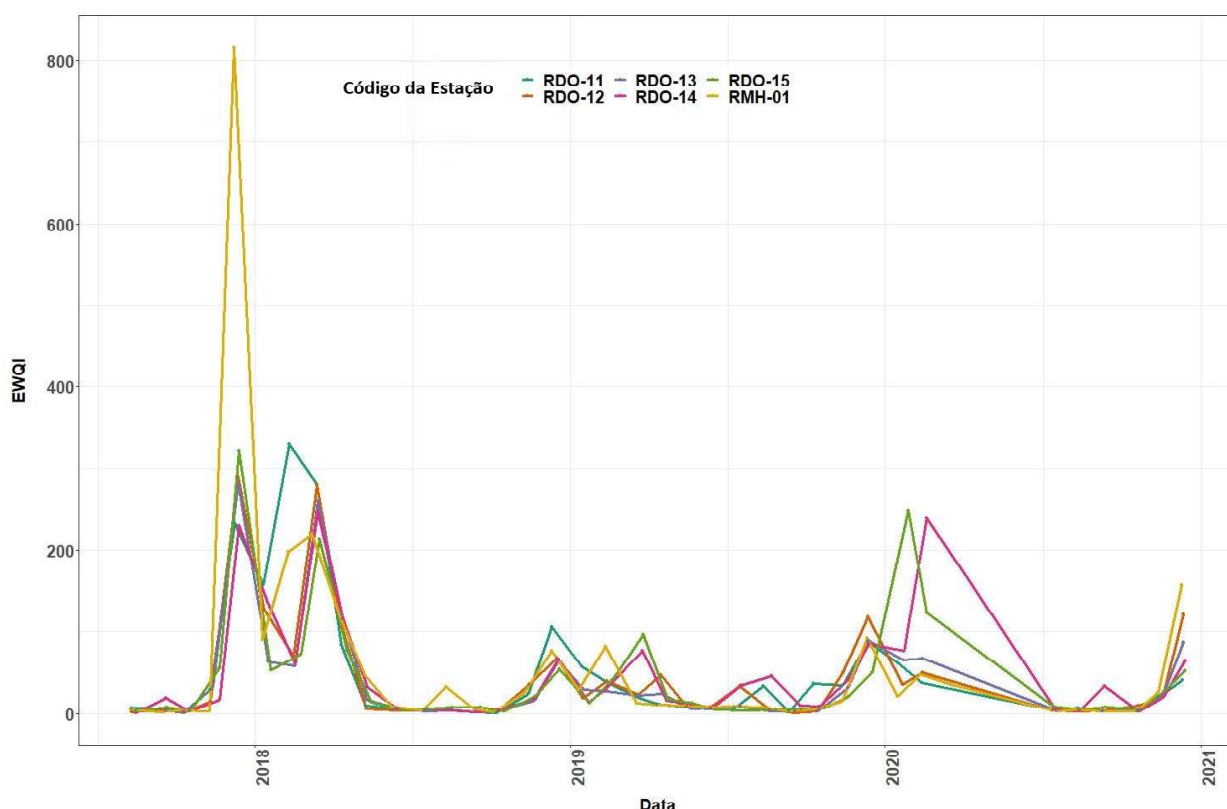


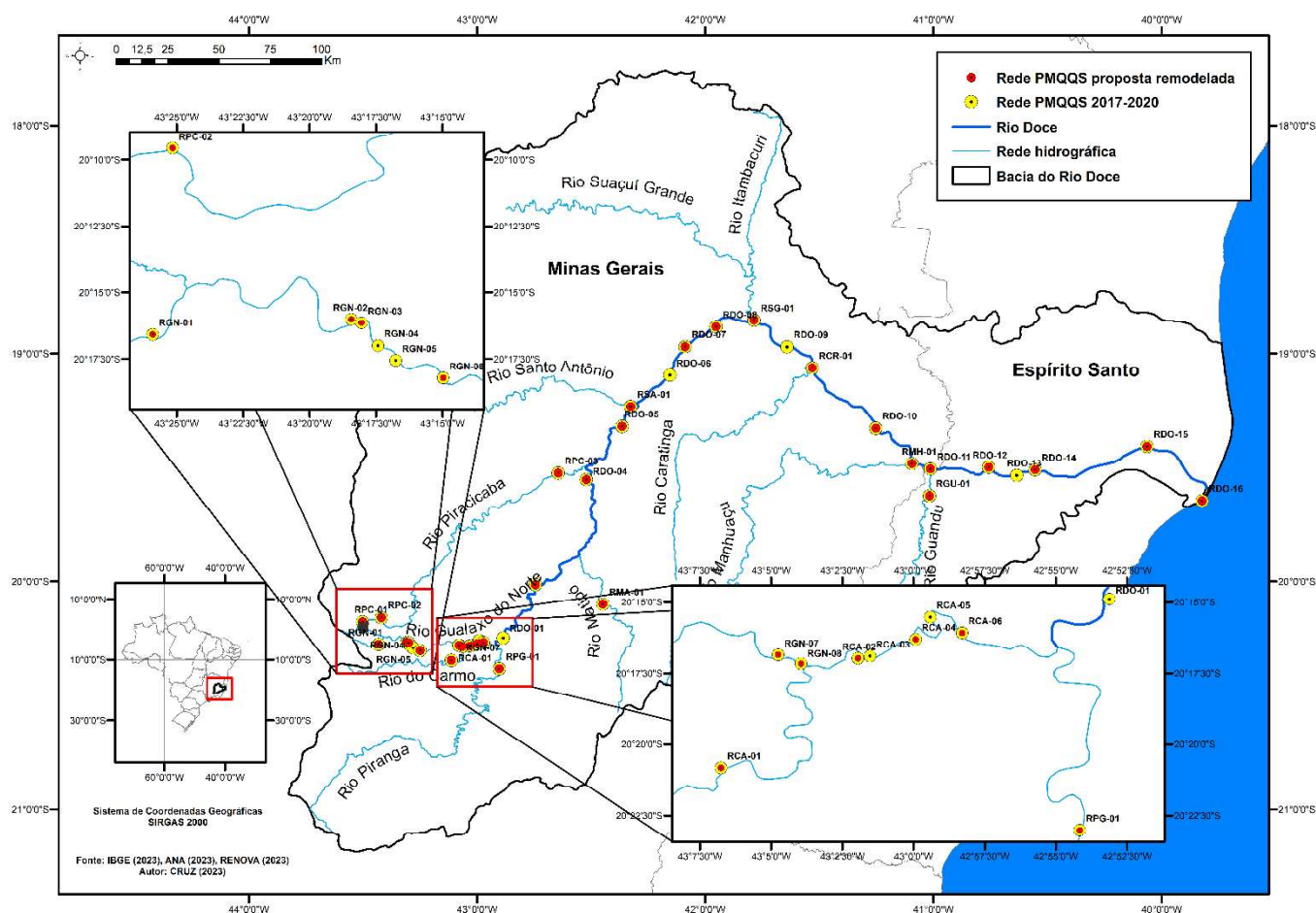
Figura 47 - Séries temporais das estações da subregião M



Para avaliar a consistência da hipótese foi aplicado o teste de Mann-Whitney sobre as séries temporais de EWQI não discretizado referente às estações RDO-13 e RDO-14. Os resultados confirmam a não existência de diferenças estatisticamente significativas entre a qualidade da água sumarizada por EWQI entre elas (p -valor = 0.249), corroborando a existência de redundância e reforçando que a decisão pela configuração formada pelas estações RDO-12, RDO-14 e RDO-15, deve ter prioridade no redimensionamento da rede, enquanto a estação RDO-13 poderia ser realocada ou descontinuada.

No cômputo geral, a integração dos resultados mais promissores de cada subregião sugere que a configuração de melhor desempenho para a rede PMQQS na bacia do Rio Doce é constituída por 29 estações de monitoramento (Figura 48).

Figura 48 – Rede de monitoramento do PMQS proposta após otimização



Os resultados demonstram haver oportunidade para redução de 20,5% das estações na bacia, sem prejuízos efetivos à informação gerada. As estações com elevada redundância identificadas alternativamente realocadas e não descontinuadas, teriam efeito positivo na rede pela manutenção da densidade espacial de estações em operação. O êxito dessa estratégia poderia ser investigado pelos mesmos critérios de avaliação de desempenho da configuração da rede aplicados, centrados no conteúdo de informação útil produzida, em uma nova atualização programática do projeto da rede.

Apesar de o método FMOHBE de maneira geral ter apresentado melhor desempenho que MGAMO o contínuo exercício realizado nas aplicações às subregiões da bacia do Rio Doce revelaram que ele ainda carece de aprimoramento, vide a dificuldade demonstrada na busca em alguns contextos.

Esse fato foi observado de forma clara nas subregiões “E” e “M” em que a abordagem de otimização de funções monoobjetivo em cascata não foi suficiente para encontrar uma solução ótima global, necessitando de critérios auxiliares. O principal fator comum a elas refere-se ao número mais elevado de estações, cada uma delas possui 5 estações de monitoramento, enquanto as demais subregiões investigadas têm estações variando de 2 a 4.

Embora a inclusão de critérios adicionais tenha permitido selecionar uma alternativa considerada factível para ambas as sub-regiões, no esforço de otimização, essa não é a solução pragmática e isenta de viés almejada. O cenário ideal desejável deveria vir da simples aplicação do método FMOHBE. Ao analisar cuidadosamente os cenários de não-convergência do método FMOHBE foi possível estabelecer uma hipótese e duas possíveis alternativas de contorno a esse problema.

A hipótese tem como premissa a semelhança ou redundância entre o conteúdo informacional de duas ou mais estações nas subregiões “E” e “M” em nível excessivamente elevado. Ainda que cada cenário tenha sido analisado sob duas diferentes formas de informação e primando pela solução mais parcimoniosa, a semelhança excessivamente expressiva da informação capturada pelas estações redundantes não permite ao método diferenciar as pequenas nuances que caracterizariam os diferentes arranjos de estações para alguns cenários.

A primeira alternativa de contorno do problema passa pela busca de outra estratégia para definição da largura dos *bins* de EWQI, visto que a discretização realizada *a posteriori* pode ter suavizado excessivamente as características estatísticas do índice (Sreeparvathy; Srinivas, 2020), eliminando assim as pequenas diferenças que caracterizariam estações redundantes.

Como demonstrado na seção 6.1.2.2, vários cenários foram simulados para a análise da sensibilidade da relação largura dos *bins* vs versus entropia conjunta. Os valores que marcaram os limites inferiores para início das simulações foram definidos com base nos critérios de Sturges e Scott (Scott, 1979; Sturges, 1926) com saltos de 10 unidades de EWQI por simulação. Como aprimoramento dessa metodologia poderiam ser aplicados novos critérios para definição do valor inicial de EWQI e simulados novos valores, com medição da entropia conjunta correspondente, acompanhados de

redução do valor simulado em cada salto, passando de 10 unidades para 5 ou até 2 unidades de EWQI.

Dessa forma, apesar de haver maior esforço computacional decorrente das simulações, haveria probabilidade de obtenção da largura do *bin* para discretização de EWQI com suavização mais atenuada, de forma que mesmo estações muito semelhantes poderiam revelar pequenas diferenças passíveis de captura pela implementação do método de otimização FMOHBE. Nessa alternativa seriam mantidas as mesmas funções monoobjetivo hierarquizadas propostas no método original.

A outra alternativa de contorno demandaria adequações nas funções monoobjetivo hierarquizadas. Um dos critérios utilizados adicionalmente na análise dos casos de subregiões onde o método FMOHBE não foi suficiente isoladamente compreendeu análises de padrões espaciais da configuração das redes.

A partir dos resultados do redimensionamento, é plausível supor que conjuntos de estações muito próximas ampliam a probabilidade de redundância, pela captura da mesma informação de qualidade da água. Essa diferença de posição poderia ser incorporada como critério específico adicional na abordagem de funções monoobjetivo hierarquizadas.

De forma a manter e ampliar a exploração do conceito de entropia seria conveniente utilizar a entropia espacial a esse fim. Variáveis aleatórias que mantêm estreito relacionamento com zonas ou regiões geográficas podem ter seu conteúdo informacional determinado como função de sua distribuição espacial. Essa categoria de entropia denomina-se entropia espacial (Singh, 2013; Batty, 1974) e dela podem derivar outras formas, como a redundância espacial, a entropia conjunta espacial e a transinformação espacial, por exemplo (Altieri; Cocchi; Roli, 2018; Li *et al.*, 2017).

A entropia espacial poderia ser determinada tanto utilizando como critério espacial a área de drenagem relacionada à posição de cada estação na rede ou alternativamente a área incremental coberta pelo monitoramento de cada estação de montante à jusante.

Ao implementar qualquer dessas duas alternativas haveria grande chance de se lograr êxito e dispensar o uso de critérios alternativos adicionais, como aqueles aplicados nessa pesquisa. Contudo, a segunda alternativa deve ser priorizada no futuro aprimoramento desse método, pois não demanda ajustes nas primeiras etapas que subsidiam a metodologia implementada ao passo que amplia o uso das formas de entropia que se busca otimizar nas funções objetivo, fato considerado positivo.

Outro ponto que merece atenção remete ao uso de prévia regionalização para fins de redimensionamento da rede completa. Ainda que esse procedimento não implique em segmentação da rede de monitoramento original, ou demande o desenvolvimento de mecanismos de gestão específicos para as estações situadas em cada sub-região, dado que a rede redimensionada corresponde apenas à agregação do arranjo de estações ideal proposto para cada sub-região, é importante realizar algumas considerações sobre essa abordagem. A primeira delas refere-se ao fato de que há inúmeras formas de regionalização que podem ser implementadas. A abordagem baseada na determinação de regiões homogêneas, considerando as informações de monitoramento de todas as variáveis, a partir do método de *cluster* apresentou resultados considerados satisfatórios, contudo, outros poderiam ter sido utilizados. Se considerar-se apenas os métodos hierárquicos, haveria possibilidade de utilizar diferentes critérios para determinação do grau de similaridade entre os grupos em detrimento da distância euclidiana, tais como a distância de Minkowsky ou o coeficiente de concordância de Jaccard, por exemplo (Mingoti, 2005).

Além disso, há diversos critérios para determinação do número de grupos ideal em métodos hierárquicos. A estatística *pseudo-F* e estatística *ccc* (*cubic clustering criterium*) são objetivos e utilizam equações específicas a esse fim (Mingoti, 2005). Outros, porém, são menos ortodoxos e oferecem mais liberdade ao pesquisador para analisar e tomar a decisão sobre a linha de corte do dendrograma e, por consequência, do número de grupos formados e da configuração das unidades que compõem cada grupo (Hair Jr *et al.*, 2006).

Para fora dos métodos hierárquicos há igualmente outros métodos não hierárquicos, como *K-means*, *Fuzzy*, *c-means* (Zhang; Wang; Wang, 2018; Hamed, 2019), e mesmo

métodos mais sofisticados baseados em redes neurais artificiais, como as redes de Kohonen (Simeonov *et al.*, 2010), seriam também alternativas a esse fim.

Em resumo, qualquer que fosse a escolha para a determinação das subregiões de qualidade da água não haveria garantias de que essa fosse a configuração correta ou ideal para o contexto da bacia, uma vez que esse fator está intimamente relacionado com a natureza de cada método. Dessa forma, apesar dos bons resultados atribuídos à abordagem de regionalização prévia e da menção da literatura aos efeitos positivos que ela pode ter no redimensionamento das redes (Khalil; Ouarda; St-Hilaire, 2011), estudos futuros poderiam investigar seu impacto comparando a configuração proposta da rede com alternativas que não adotem essa estratégia. Para isso, seria fundamental garantir a uniformidade dos critérios especificados nas funções monoobjetivo.

Para esse fim poderiam ser utilizados algoritmos de otimização inspirados na natureza, como os algoritmos genéticos, as colônias de abelhas e enxames de partículas, pois têm tido grande relevância atualmente no campo do redimensionamento de redes de monitoramento da qualidade da água (Ghorbani *et al.*, 2020; Zhu *et al.*, 2019; Vega-Rodríguez *et al.*, 2021).

Em decorrência da capacidade dos algoritmos inspirados na natureza de solucionar problemas complexos de otimização (Khishe; Mosavi, 2020), as funções objetivos de FMOHBE poderiam ser agrupadas nessa aplicação em uma única solução objetivo composta, considerando os mesmos critérios de maximização ou minimização, as formas de entropia e as restrições da abordagem já implementada. Essa estratégia preservaria condições semelhantes de aplicação em ambos os métodos e permitiria comparações mais justas dos resultados, posteriormente.

8 CONCLUSÕES

Os métodos investigados nessa pesquisa buscaram identificar oportunidades para o redimensionamento de redes emergenciais, visando aprimorar o desempenho dos programas de monitoramento e melhoria da relação custo-benefício na geração dos dados pela redução de redundâncias

A estratégia adotada de regionalização de grandes bacias permitiu que os diferentes contextos socioambientais da bacia, subjacentes à qualidade da água, fossem incorporados indiretamente na investigação. Além disso essa abordagem promoveu redução da complexidade dos cenários através da redução do número de estações e cenário analisados em cada subregião.

Apesar disso, esse procedimento deve ser realizado com cautela e atenção, uma vez que há grande diversidade de abordagens potencialmente aplicáveis para esse fim e a escolha do critério de regionalização impacta diretamente a configuração das subregiões. Em suma, não há garantias que ao adotar outros critérios para regionalização da rede emergencial analisada o método proposto produza o mesmo resultado quanto a seleção das “variáveis não-críticas” e redefinição da rede de monitoramento ideal.

O critério para seleção de “variáveis não críticas” demonstrou capacidade de capturar o conteúdo informativo de variáveis de qualidade da água cuja presença no meio ambiente pode ser da ordem de partes por bilhão. Portanto, a baixa proporção no meio ambiente não constituiu uma limitação ao método, revelando assim um amplo potencial de aplicação em bacias impactadas por desastres com diversidade de poluentes hídricos envolvidos.

A entropia da Shannon indicou o conjunto de variáveis que podem ser monitoradas em menor frequência sem prejuízo à informação útil à gestão produzida pelas redes emergenciais. As “variáveis não-críticas” na bacia do Rio Doce compreenderam de 32 a 50% do total de variáveis analisadas. Embora não seja recomendado a descontinuação imediata dessas variáveis, a simples redução na frequência de monitoramento pode impactar positivamente o gerenciamento das redes, pela redução dos custos.

O método de seleção de “variáveis não-críticas” também demonstrou capacidade de capturar as flutuações espaço-temporais na qualidade da água. Três aspectos relativos aos resultados do estudo de caso na bacia do Rio Doce corroboraram essa conclusão. O primeiro deles foi o fato de que em nenhuma subregião houve repetibilidade do mesmo conjunto de variáveis selecionadas, embora haja variáveis que demonstraram possuir baixa importância em grande parte das subregiões.

O segundo, compreendeu o fato de variáveis consideradas redundantes na maior da bacia mas que sabidamente tem *hot-spots* em zonas específicas, terem sido classificadas como “variáveis de baixo conteúdo de informação” em todas as áreas à exceção dessas subregiões, onde demonstraram ser bem mais informativas. O terceiro aspecto deveu-se a demonstração que o conjunto de variáveis não-críticas pode sofrer alterações decorrentes da sazonalidade.

Apesar desse método ter demonstrado capacidade para apoiar os tomadores de decisão e gestores das redes emergenciais ele possui uma limitação associada à presença de *outliers*. Como todos os métodos baseados na entropia de Shannon são em essência métodos estatísticos, deve-se ter sempre atenção quanto à presença de *outliers*, pois eles podem enviesar os resultados, subestimando a redundância de informações das variáveis e produzindo resultados inconsistentes.

Quanto à seleção da configuração de melhor desempenho para redes de monitoramento emergencial, a estratégia em integrar os dados de qualidade da água por meio do EWQI permitiu caracterizar o comportamento sazonal das classes de qualidade da água e das variáveis de maior contribuição. Cada índice regional calculado foi corroborado por outros estudos na bacia investigada.

O uso da entropia de Shannon nos problemas de otimização monoobjetivo hierárquico, logrou êxito na identificação das estações pouco informativas cujo elevado grau de redundância as habilitaria a descontinuação, nas situações de número reduzido de estações. Há potencial para redução de 20% das estações em operação na rede do PMQQS.

Alternativamente estações dispensáveis podem ser realocadas, ao invés de descontinuadas. Esse procedimento embora não impacte os custos de operação das

redes, pode oportunizar a geração de novas informações úteis, desde que os resultados dessa operação avaliados pela reaplicação dos métodos propostos, demonstrem esse fato. Além disso, ao realocar as estações redundantes mantém-se a mesma densidade espacial de estações nas redes analisadas

Em subregiões com número de estações mais expressivo, o método de otimização proposto apresentou dificuldades para encontrar uma solução ótima global., de forma que são necessários ajustes metodológicos para tornar o método mais eficiente, objetivo e isento de juízos pessoais de critérios auxiliares.

Por fim, considerando o contexto ainda presente e recorrente de desastres ambientais que trazem grande prejuízo aos sistemas hídricos, ecossistemas e população humana reitera-se o potencial dos métodos propostos em apoiar os gestores de programas de monitoramento emergencial de qualidade da água, na busca da melhor eficiência na operação das redes.

9 RECOMENDAÇÕES

O tratamento dos dados que apoiaram a seleção de “variáveis não-críticas” incluiu uma etapa de limpeza dos *outliers* das séries temporais das variáveis analisadas, como forma de eliminar ou no mínimo reduzir o seu efeito sobre a determinação da entropia de Shannon. Embora, tenha sido escolhido o método de Gini, devido aos resultados considerados satisfatórios, recomenda-se que em aplicações futuras sejam testados outros métodos, que podem mostrar-se também adequados aos dados analisados.

A capacidade demonstrada pelo método de seleção de “variáveis não-críticas” em capturar as flutuações espaço-temporais do conteúdo de informação de variáveis de qualidade da água pode prover aos gestores de programas emergenciais de qualidade da água elementos suficientes para a incorporação dessas flutuações nos critérios de monitoramento.

Dessa forma, recomenda-se a adoção de diferentes frequências de monitoramento nos procedimentos operacionais dos PMQA, considerando as flutuações sazonais do conteúdo de informação demonstradas pelo método aplicado. Esse desdobramento teria potencial para impactar positivamente o gerenciamento das redes, pois garantiria

o monitoramento alinhado aos períodos em que as variáveis são mais informativas. Além disso, após algum tempo haveria maior segurança para os gestores decidirem ou não pela descontinuação do monitoramento de variáveis redundantes, considerando seu conteúdo de informação.

Quanto à seleção da configuração de melhor desempenho das redes recomenda-se não realizar qualquer tipo de limpeza ou imputação de dados nas séries temporais das variáveis de qualidade da água, uma vez que esse procedimento poderia influenciar a determinação dos valores do EWQI e, por consequência, o processo de discretização das variáveis.

Como forma de superar a necessidade de critérios auxiliares de decisão sobre a configuração da rede redimensionada, recomenda-se que estudos futuros investiguem ajustes na metodologia apresentada. Isso tanto pode ser feito, a partir de alterações nos critérios de determinação da largura dos *bins* e, posterior, discretização de EWQI ou via ajuste do método de otimização pela inclusão de uma nova função monoobjetivo, preferencialmente calcada no uso da entropia espacial. Deve-se priorizar a segunda alternativa entre as opções, visto que além de aprofundar o uso de variadas formas de entropia como critério de seleção da rede ideal, ela dispensa a necessidade de ajuste nos procedimentos metodológicos que estão nas bases da metodologia de redimensionamento de redes emergenciais proposta.

Quanto ao redimensionamento da configuração espacial das redes, embora a rede proposta decorrente da aplicação do estudo de caso seja factível, vide o fato de que de maneiras diversas pôde-se identificar estações com elevada redundância, é importante que se possa aferir objetivamente a qualidade dos resultados da metodologia confrontando-os com outros obtidos de um procedimento de redimensionamento que não adote a regionalização da rede investigada como etapa prévia.

Além disso, apesar do ponto focal dessa pesquisa ter sido as redes emergenciais integrantes de programas de monitoramento emergencial da qualidade da água, é perfeitamente possível que aplicações dos critérios desenvolvidos para a seleção das variáveis de qualidade da água e redimensionamento da configuração espacial do conjunto de estações obtenham bons resultados em outras categorias de redes, como

as investigativas e regulatórias, por exemplo. Dessa forma, encoraja-se o uso da metodologia proposta para além do contexto vinculado as bacias impactadas por desastres ambientais.

REFERÊNCIAS

ABDI, H.; WILLIAMS, L. J. Principal component analysis. **Wiley Interdisciplinary Reviews: Computational Statistics**, v. 2, n. 4, p. 433–459, 2010.

AFSAR, B.; MIETTINEN, K.; RUIZ, F. **Assessing the performance of interactive multiobjective optimization methods** : A survey. v. 54, n. 4, 2021.

AFTAB, O., CHEUNG, P., KIM, A., THAKKAR, S., AND YEDDANAPUDI, N. **Information theory and the digital age**. Massachusetts, MA: Bandwagon: MIT, 2011. Disponível em: <<http://web.mit.edu/6.933/www/Fall2001/Shannon2.pdf>>.

AGÊNCIA NACIONAL DE ÁGUA (BRASIL) ANA. **Encarte Especial sobre a Bacia do Rio Doce Rompimento da Barragem em Mariana/MG Conjuntura dos Recursos Hídricos no Brasil**. Brasília: ANA, 2016.

AGÊNCIA NACIONAL DE ÁGUAS (BRASIL). **Atlas esgotos: Despoluição de bacias hidrográficas**. Brasília: ANA, 2023a. Disponível em: <<https://www.abntcatalogo.com.br/norma.aspx?Q=bTREVWUzMXNCZ0VBT1BUa1JoRXFnN3FJUzNiUURHL1U1ZElyc0F4SjJtaz0=>>>.

AGÊNCIA NACIONAL DE ÁGUAS E SANEAMENTO BÁSICO (ANA). **Plano integrado de recursos hídricos da bacia hidrográfica do rio Doce 2023-2042 e enquadramento dos corpos de água superficiais**. Brasília: ANA, 2023b.

AGUIAR, V. M. DE C.; NETO, J. A. B.; QUARESMA, V. DA S.; BASTOS, A. C.; ATHAYDE, J. P. M. DE. Bioavailability and ecological risks of trace metals in bottom sediments from Doce river continental shelf before and after the biggest ~ o dam environmental disaster in Brazil: The collapse of the Fund a. **Journal of Environmental Management**, v. 272, p. 1–14, 2020.

AGUILA, Y. A global pact for the environment: The logical outcome of 50 years of international environmental law. **Sustainability (Switzerland)**, v. 12, n. 14, p. 1–17, 2020.

ALAMEDDINE, I.; KARMAKAR, S.; QIAN, S. S.; PAERL, H. W.; RECKHOW, K. H.

Optimizing an estuarine water quality monitoring program through an entropy-based hierarchical spatiotemporal Bayesian framework. **Water Resources Research**, v. 49, n. 10, p. 6933–6945, 2013.

ALFONSO, L.; HE, L.; LOBBRECHT, A.; PRICE, R. Information theory applied to evaluate the discharge monitoring network of the Magdalena River. **Journal of Hydroinformatics**, v. 15, n. 1, p. 211–228, 2013.

ALFONSO, L.; LOBBRECHT, A.; PRICE, R. Information theory-based approach for location of monitoring water level gauges in polders. **Water Resources Research**, v. 46, n. 10, p. 1–14, 2010.

ALI, M. Z. M.; OTHMAN, F. Raingauge network optimization in a tropical urban area by coupling cross-validation with the geostatistical technique. **Hydrological Sciences Journal**, v. 63, n. 3, p. 474–491, 2018.

ALLMENDINGER, R.; EHRGOTT, M.; KLAMROTH, K.; LUQUE, M.; GANDIBLEUX, X.; GEIGER, M. J. Navigation in multiobjective optimization methods. **J Multi-Crit Decis Anal**, v. 24, p. 57–70, 2017.

ALMEIDA, R. G. B. DE; LAMPARELLI, M. C.; DODDS, W. K.; CUNHA, D. G. F. Spatial optimization of the water quality monitoring network in São Paulo State (Brazil) to improve sampling efficiency and reduce bias in a developing sub-tropical region. **Environmental Science and Pollution Research**, v. 29, n. 8, p. 11374–11392, 2022.

ALONSO, D. L.; PÉREZ, R.; OKIO, C. K. Y. A.; CASTILLO, E. Assessment of mining activity on arsenic contamination in surface water and sediments in southwestern area of Santurbán paramo, Colombia. **Journal of Environmental Management**, v. 264, n. February, p. 1-10, 2020.

ALTIERI, L.; COCCHI, D.; ROLI, G. A new approach to spatial entropy measures. **Environmental and Ecological Statistics**, v. 25, n. 1, p. 95–110, 2018.

ALVES, J. DO P. H.; FONSECA, L. C.; CHIELLE, R. DE S. A.; MACEDO, L. C. B. Monitoring water quality of the sergipe river basin: An evaluation using multivariate data analysis. **Revista Brasileira de Recursos Hidricos**, v. 23, p. 1–12, 2018.

AMIRI, V.; REZAEI, M.; SOHRABI, N. Groundwater quality assessment using entropy weighted water quality index (EWQI) in Lenjanat, Iran. **Environmental Earth Sciences**, v. 72, n. 9, p. 3479–3490, 2014.

ANDRADES, R.; MARTINS, R. F.; GUABIROBA, H. C.; RODRIGUES, V. L. A.; SZABLAK, F. T.; BASTOS, K. V.; BASTOS, P. G. P.; LIMA, L. R. S.; VILAR, C. C.; JOYEUX, J. C. Effects of seasonal contaminant remobilization on the community trophic dynamics in a Brazilian tropical estuary. **Science of the Total Environment**, v. 801, p. 1496–1470, 2021.

ANTANASIJEVIĆ, D.; POCAJT, V.; ANTANASIJEVIĆ, J.; PERIĆ-GRUJIĆ, A.; RISTIĆ, M. A novel SON2-based similarity index and its application for the rationalization of river water quality monitoring network. **River Res Applic.**, v. 34, p. 144–152, 2018.

ARMSTRONG, M.; PETTER, R.; PETTER, C. Why have so many tailings dams failed in recent years? **Resources Policy**, v. 63, n. November 2018, p. 1-11, 2019.

ASADI, A.; MOGHADDAM NIA, A.; BAKHTIARI ENAYAT, B.; ALILOU, H.; AHMADISHARAF, E.; KIMUTAI KANDA, E.; CHESSUM KIPKORIR, E. An integrated approach for prioritization of river water quality sampling points using modified Sanders, analytic network process, and hydrodynamic modeling. **Environmental Monitoring and Assessment**, v. 193, n. 8, p. 1–15, 2021.

AZEVEDO, A. M. **MultivariateAnalysis: Pacote Para Analise Multivariada**. [s.l.: s.n.]. Disponível em: <<https://cran.r-project.org/package=MultivariateAnalysis>>.

AZHAR, S. C.; ARIS, A. Z.; YUSOFF, M. K.; RAMLI, M. F.; JUAHIR, H. Classification of river water quality using multivariate analysis. **Procedia Environmental Sciences**, v. 30, p. 79–84, 2015.

BARBAROS, F. Entropy-assisted approach to determine priorities in water quality monitoring process. **Environmental Monitoring and Assessment**, v. 194, n. 12, p. 1-16, 2022.

BARCELLOS, D. DA S.; SOUZA, F. T. DE. Optimization of water quality monitoring

programs by data mining. **Water Research**, v. 221, n. June, p. 1–12, 2022.

BARCELOS, D. A.; PONTES, F. V. M.; FERNANDA, A. N. G.; CASTRO, D. C.; NATHALIA, O. A.; CASTILHOS, Z. C. Gold mining tailing: Environmental availability of metals and human health risk assessment. **Journal of Hazardous Materials**, v. 397, n. April, p. 1-8, 2020.

BARTRAM, J.; BALLANCE, R. **Water quality monitoring - A practical guide to the design and implementation of freshwater quality studies and monitoring programmes**. London and New York: UNEP/WHO, 1996.

BASTE, I. A.; WATSON, R. T. Tackling the climate, biodiversity and pollution emergencies by making peace with nature 50 years after the Stockholm Conference. **Global Environmental Change**, v. 73, n. February, p. 1-8, 2022.

BATTY, M. Spatial Entropy. **Geographical Analysis**, v. 6, n. 1, p. 1–31, 1974.

BAUDSON, E. *et al.* Trace metals in Rio Doce sediments before and after the collapse of the Fundão iron ore tailing dam, Southeastern Brazil. **Chemosphere**, v. 262, p. 1–8, 2021.

BEHMEL, S.; DAMOUR, M.; LUDWIG, R.; RODRIGUEZ, M. J. Water quality monitoring strategies — A review and future perspectives. **Science of the Total Environment**, v. 571, p. 1312–1329, 2016.

BERTOLDO, L. A.; RIBEIRO, A.; REIS, C. E. S.; FRACHINI, E.; KROETZ, B. L.; ABRÃO, T.; SANTOS, M. J. Environmental risk assessment of potentially toxic elements in Doce River watershed after mining sludge dam breakdown in Mariana, MG, Brazil. **Environmental Monitoring and Assessment**, v. 195, n. 5, p.1-16, 2023.

BORBA, R. P.; FIGUEIREDO, B. R.; RAWLINS, B.; MATSCHULLAT, J. Arsenic in water and sediment in the Iron quadrangle , State of Minas Gerais , Brazil. **Revista Brasileira de Geociências**, v. 30, n. 3, p. 558–561, 2000.

BORTOLETTO, E. C.; SILVA, H. A.; BONIFÁCIO, C. M.; TAVARES, C. R. G. Water quality monitoring of the Pirapó River watershed, Paraná, Brazil. **Brazilian Journal of**

Biology, v. 75, n. 4 suppl 2, p. 148–157, 2015.

BRASIL, R. M.; SILVA, M. A. DA. **Otimização de projetos de engenharia**. São Paulo: Editora Blucher, 2019.

BRASIL. Conselho Nacional do Meio Ambiente (CONAMA). **Resolução CONAMA 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: seção 1, Brasília, DF, n. 53, p. 58-63, 18 mar. 2005. Disponível em: <http://www2.mma.gov.br/port/conama/res/res05/res35705.pdf>. Acesso em: 17 mar. 2022.

BROWN, R.; DEMPSTER, M. A. H.; ADBY, P. R. **Introduction to optimization methods**. Londres: Chapman & Hall, 1974.

BURRITT, R. L.; CHRIST, K. L. Water risk in mining: Analysis of the Samarco dam failure. **Journal of Cleaner Production**, v. 178, p. 196–205, 2018.

CACCIUTTOLO, C.; CANO, D. Environmental impact assessment of mine tailings spill considering metallurgical processes of Gold and Copper mining : Case studies in the Andean countries of Chile. **Water**, v. 14, p. 1–30, 2022.

CAGNIN, R. C. *et al.* Macronutrients and dissolved iron in a land-ocean approach: Influences of contamination by ore tailings in Southeastern Brazil. **Frontiers in Marine Science**, v. 9, n. November, p. 1–19, 2022.

CALAZANS, G. M.; PINTO, C. C.; COSTA, E. P. DA; PERINI, A. F.; OLIVEIRA, S. C. Using multivariate techniques as a strategy to guide optimization projects for the surface water quality network monitoring in the Velhas river basin, Brazil. **Environmental Modeling and Assessment**, v. 190, n. 726, p. 1–15, 2018.

CAMARA, M.; JAMIL, N. R.; ABDULLAH, A. F. BIN; HASHIM, R. BINTI; ALIYU, A. G. Economic and efficiency based optimisation of water quality monitoring network for land use impact assessment. **Science of the Total Environment**, v. 737, p. 1-14, 2020.

CANADIAN COUNCIL OF MINISTERS OF THE ENVIRONMENT (CCME). **A Canada-wide framework for water quality monitoring**. Winnipeg: CCME, 2006.

CANADIAN COUNCIL OF MINISTERS OF THE ENVIRONMENT (CCME). **Guidance manual for optimizing water quality monitoring program design**. Winnipeg: CCME, 2015.

CARVALHO, D. W. DE. The ore tailings dam rupture disaster in mariana, Brazil 2015: What we have to learn from anthropogenic disasters. **Natural Resources Journal**, v. 59, n. 2, p. 281–300, 2019.

CARVALHO, G. *et al.* Genotoxic, biochemical and bioconcentration effects of manganese on *Oreochromis niloticus* (Cichlidae). **Ecotoxicology**, v. 27, p. 1150–1160, 2018.

CASTELLETTI, A.; PIANOSI, F.; RESTELLI, M. A multiobjective reinforcement learning approach to water resources systems operation: Pareto frontier approximation in a single run. **Water Resources Research**, v. 49, n. 6, p. 3476–3486, 2013.

ÇENGEL, Y. A. On entropy, information, and conservation of information. **Entropy**, v. 23, n. 6, p. 1-19, 2021.

CHANG, C. L.; LIN, Y. T. Using the VIKOR method to evaluate the design of a water quality monitoring network in a watershed. **International Journal of Environmental Science and Technology**, v. 11, n. 2, p. 303–310, 2014.

CHAPRA, S. C.; CAMACHO, L. A.; MCBRIDE, G. B. Impact of global warming on dissolved oxygen and BOD assimilative capacity of the world ' s rivers : Modeling analysis. **Water**, v. 13, p. 1-20, 2021.

CHASALOW, S. **Combinat: combinatorics utilities**, 2012. Disponível em: <<https://cran.r-project.org/package=combinat>>

CHEN, D.; MENG-CHUEN; BODIRSKY, B. L.; KRUEGER, T.; MISHRA, A.; POPP, A. The world ' s growing municipal solid waste : trends and impacts Recent citations The world ' s growing municipal solid waste: trends and impacts. **Environmental**

Research Letters, v. 15, p. 1-13, 2020.

COLLINS, A.; OHANDJA, D. G.; HOARE, D.; VOULVOULIS, N. Implementing the Water Framework Directive: A transition from established monitoring networks in England and Wales. **Environmental Science and Policy**, v. 17, p. 49–61, 2012.

COMITÊ DE BACIA HIDROGRÁFICA DA BACIA DO RIO DOCE (CBH-DOCE). **Sigaweb Doce: solos.** Disponível em: <<https://sigaaguas.org.br/sigaweb/apps/doce/>>. Acesso em: 8 set. 2024.

CORDEIRO, M. C. *et al.* Insights on the freshwater microbiomes metabolic changes associated with the world's largest mining disaster. **Science of the Total Environment**, v. 654, p. 1209–1217, 2019.

COSTA, A. T.; NALINI, H. A.; TARSO AMORIM CASTRO, P. DE; TATUMI, S. H. Stratigraphic analysis and distribution of arsenic in quaternary sedimentary deposits of the southeastern portion of the Quadrilátero Ferrífero, basin of the Ribeirão Carmo, MG. **Revista Escola de Minas**, v. 63, n. 4, p. 703–714, 2010.

COSTA, E. S.; CAGNIN, R. C.; SILVA, C. A. DA; LONGHINI, C. M.; SÁ, F.; LIMA, A. T.; GOMES, L. E. DE O.; BERNARDINO, A. F.; NETO, R. R. Iron ore tailings as a source of nutrients to the coastal zone. **Marine Pollution Bulletin**, v. 171, n. April, p. 112725, 2021.

CUPOLILLO, F.; ABREU, M. L. DE; VIANELLO, R. L. Climatologia da bacia do Rio Doce e sua relação com a topografia local. **Geografias**, v. 04, n. 1, p. 45–60, 2008.

DA CUNHA RICHARD, E.; AGUIAR DUARTE, H. DE; CALDERUCIO DUQUE ESTRADA, G.; BECHTOLD, J. P.; GUSSO MAIOLI, B.; ARAUJO DE FREITAS, A. H.; ELIZABETH WARNER, K.; MELGES FIGUEIREDO, L. H. Influence of Fundão tailings dam breach on water quality in the Doce River watershed. **Integrated Environmental Assessment and Management**, v. 16, n. 5, p. 583–595, 2020.

DASH, S.; KALAMDHAD, A. S. Science mapping approach to critical reviewing of published literature on water quality indexing. **Ecological Indicators**, v. 128, p. 1–18, 2021.

DAUS, B.; WENNRICH, R.; MORGENSTERN, P.; WEISS, H.; EUGE, H.; BIRGIT DAUS¹;?, RAINER WENNRICH¹, PETER MORGENSTERN¹, HOLGER WEISS¹, H. E. L. P.; NALINI, H. A.; LEONEL, L. V.; MONTEIRO, R. P. G.; MOREIRA, R. M. Arsenic speciation in plant samples from the Iron Quadrangle , Brazil. **Microchim Acta**, v. 180, p. 175–180, 2005.

DAVILA, R. B.; FONTES, M. P. F.; PACHECO, A. A.; FERREIRA, M. DA S. Heavy metals in iron ore tailings and floodplain soils affected by the Samarco dam collapse in Brazil. **Science of the Total Environment**, v. 709, p. 136–151, 2020.

DE CARVALHO, G. O.; PINHEIRO, A. DE A.; SOUSA, D. M. DE; PADILHA, J. DE A.; SOUZA, J. S.; GALVÃO, P. M.; PAIVA, T. DE C.; FREIRE, A. S.; SANTELLI, R. E.; MALM, O.; TORRES, J. P. M. Metals and arsenic in water supply for riverine communities affected by the largest environmental disaster in Brazil: The dam collapse on Doce river. **Orbital**, v. 10, n. 4, p. 299–307, 2018.

DENIS ABESSA, G.; JR., A. B.; CERVI, E. C.; SIMPSON, S. L.; STUBBLEFIELD, W.; RIBEIRO, C. C.; CRUZ, A. C. F.; KRUGER, G.; SMITH, R. Has the Rio Doce time bomb been defused Using a weight-of-evidence. **Integr Environ Assess Manag** 00, v. 0, n. 2023, p. 1-11, 2023.

DEONARINE, A.; BARTOV, G.; JOHNSON, T. M.; RUHL, L.; VENGOSH, A.; HSU-KIM, H. Environmental impacts of the Tennessee Valley Authority Kingston coal ash spill. 2. effect of coal ash on methylmercury in historically contaminated river sediments. **Environmental Science and Technology**, v. 47, n. 4, p. 2100–2108, 2013.

DESCHAMPS, E.; CIMINELLI, V. S. T.; LANGE, F. T.; MATSCHULLAT, J.; RAUE, B.; SCHMIDT, H. Soil and sediment geochemistry of the iron quadrangle, Brazil: The case of Arsenic. **Journal of Soils and Sediments**, v. 2, n. 4, p. 216–222, 2002.

DHARMAPRANI, D.; DYKES, L.; MCGAVIGAN, A. D.; KUKLIK, P.; POPE, K.; GANESAN, A. N. Information theory and atrial fibrillation (AF): A review. **Frontiers in Physiology**, v. 9, n. JUL, p. 1–19, 2018.

DING, X.; FANG, P. Accident trend prediction of heavy metal pollution in the

Heshangshan drinking water source area based on integrating a two-dimensional water quality model and GIS. **Sustainability (Switzerland)**, v. 11, n. 15, p. 1-20, 2019.

DIXON, W.; CHISWELL, B. Review of aquatic monitoring program design. **Water Research**, v. 30, n. 9, p. 1935–1948, 1996.

DO, H. T.; LO, S. L.; CHIUUEH, P. TE; THI, L. A. P.; SHANG, W. T. Optimal design of river nutrient monitoring points based on an export coefficient model. **Journal of Hydrology**, v. 406, p. 129–135, 2011.

DO, H. T.; LO, S. L.; PHAN THI, L. A. Calculating of river water quality sampling frequency by the analytic hierarchy process (AHP). **Environmental Monitoring and Assessment**, v. 185, n. 1, p. 909–916, 2013.

DOS SANTOS, G. R.; MAIA, L. C.; LOBO, F. A.; SANTIAGO, A. DA F.; SILVA, G. A. DA. A model based on a multivariate classification for assessing impacts on water quality in a DOCE river watershed after the Fundão tailings dam failure. **Environmental Pollution**, v. 334, n. July, p. 1-10, 2023.

DUARTE, E. B.; NEVES, M. A. Main chemical and mineralogical components of the Rio Doce sediments and the iron ore tailing from the Fundão Dam disaster , Southeastern Brazil. **Environmental Monitoring and Assessment**, p. 1–13, 2023.

ECOPLAN-LUME. **Plano integrado de recursos hídricos da bacia hidrográfica do Rio Doce – VOLUME I: Relatório final**. Belo Horizonte: IGAM, 2010.

ENGEORPS. **Revisão e atualização do PIRH Doce e propostas de enquadramento: Produto 03 – Consolidação do estado da arte sobre os recursos hídricos na bacia** Agência Nacional de Águas e Saneamento. Barueri: ANA, 2021.

ESPINDOLA, H. S. Vale do Rio Doce: Fronteira, industrialização e colapso socioambiental. **Fronteiras: Journal of Social, Technological and Environmental Science**, v. 4, n. 1, p. 160-206, 2015.

ESPINDOLA, H. S.; ELISA, I.; MIFARREG, G. Território da mineração: uma contribuição teórica. **Revista Brasileira de Geografia**, v. 62, n. 2, p. 67–93, 2017.

EUROPEAN UNION (EU). **Guidance on monitoring for the water framework directive: Final version Europe**. Copenhagen: WFD, 2003.

FAGBOTE, E. O.; OLANIPEKUN, E. O.; UYI, H. S. Water quality index of the ground water of bitumen deposit impacted farm settlements using entropy weighted method. **International Journal of Environmental Science and Technology**, v. 11, n. 1, p. 127–138, 2014.

FAGERHOLM, J.; HAATAJA, J.; JÄRVINEN, J.; LYLÄ, M.; RABACK, P.; SAVOLAINEN, V. **Applied Parallel Computing: advanced scientific computing**. Berlin/ Germany: Springer. v. 1716, 2002.

FELIPPE, M. F.; JUNIOR, A. P. M.; MENDES, L. C.; COTA, G. E. M.; CARNEIRO, P. S.; GONTIJO, B. M. Conexões geo-históricas e contemporâneas entre ocupação territorial , degradação ambiental e rarefação hídrica na Bacia do Rio Doce. **Geografias**, v. Edição esp, p. 203–222, 2016.

FERRÉ, L. Selection of components in principal component analysis: A comparison of methods. **Computational Statistics and Data Analysis**, v. 19, n. 6, p. 669–682, 1995.

FIGUEIREDO, B. R. DE; BORBA, R. P.; ANGÉLICA, R. S. Arsenic occurrence in Brazil and human exposure. **Environ Geochem Health (2007)**, v. 29, p. 109–118, 2007.

FIGUEIREDO, J. A. DE; NORIEGA, C. D.; OLIVEIRA, E. M. C. DE; RODRIGUES NETO, R.; BARROSO, G. F.; ARAÚJO FILHO, M. Avaliação biogeoquímica de águas fluviais com ênfase no comportamento dos compostos de nitrogênio e fósforo total para diagnoses provenientes do sistema aquático Bacia do Rio Doce, no Espírito Santo. **Geochimica Brasiliensis**, v. 28, n. 2, p. 215–226, 2014.

FIGUEIREDO, M. D.; LAMEIRAS, F. S.; ARDISSON, J. D.; ARAUJO, M. H.; TEIXEIRA, A. P. DE C. Tailings from Fundão tragedy: Physical–Chemical properties of the material that remains by Candonga dam. **Integrated Environmental Assessment and Management**, v. 16, n. 5, p. 636–642, 2020.

FISCHER, G. Acelerações em escala regional: A transformação do vale do Rio Doce, ca. 1880-1980. **Varia Historia**, v. 34, n. 65, p. 445–474, 2018.

FOROOZAND, H.; WEIJS, S. V. Objective functions for information-theoretical monitoring network design: What is “optimal”? **Hydrology and Earth System Sciences**, v. 25, n. 2, p. 831–850, 2021.

FRAGA, M. DE S.; SILVA, D. D. DA; REIS, G. B.; GUEDES, H. A. S.; ELESBON, A. A. A. Temporal and spatial trend analysis of surface water quality in the Doce River basin, Minas Gerais, Brazil. **Environment, Development and Sustainability**, v. 23, n. 8, p. 12124–12150, 2021.

FREEDMAN, D.; DIACONIS, P. On the histogram as a density estimator:L2 theory. **Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete**, v. 57, n. 4, p. 453–476, 1981.

FREITAS, C. M.; ANDRADE DA SILVA, M. Work accidents which become disasters: mine tailing dam failures in Brazil. **Revista Brasileira de Medicina do Trabalho**, v. 17, n. 1, p. 21–29, 2019.

GABRIEL, F. Â.; FERREIRA, A. D.; QUEIROZ, H. M.; VASCONCELOS, A. L. S.; FERREIRA, T. O.; BERNARDINO, A. F. Long-term contamination of the Rio Doce estuary as a result of Brazil’s largest environmental disaster. **Perspectives in ecology and conservation**, v. 19, p. 417–428, 2021.

GABRIEL, F. A.; SILVA, A. G.; QUEIROZ, H. M.; FERREIRA, T. O.; HAUSER-DAVIS, R. A.; BERNARDINO, A. F. Ecological risks of metal and metalloid contamination in the Rio Doce estuary. **Integrated Environmental Assessment and Management**, v. 16, n. 5, p. 655–660, 2020.

GHORBANI, M.; LILTVED, H.; REZA, M.; HINDAR, A. Assessing optimal water quality monitoring network in road construction using integrated information-theoretic techniques. **Journal of Hydrology**, v. 589, n. July, p. 1–12, 2020.

GOMES, G. R. S. **Uniformidade e entropia**. Lisboa: Universidade de Lisboa, 2013.

GOUDARZI, F.; HEDAYATIAGHMASHHADI, A.; KAZEMI, A.; FÜRST, C. Optimal location of water quality monitoring stations using an artificial neural network modeling in the Qarah-Chay River Basin, Iran. **Water (Switzerland)**, v. 14, n. 870, p. 1–16, 2022.

GRAHAM, A.; WILCOX, D. A. The impacts of Marcellus Shale gas drilling accidents on amphibians in a Pennsylvania fen. **Wetlands Ecology and Management**, v. 29, n. 1, p. 155–167, 2021.

GRAVELINE, N.; MATON, L.; RINAUDO, J. D.; LÜCKGE, H.; INTERWIES, E.; ROUILLARD, J.; STROSSER, P.; PALKANIETE, K.; TAVERNE, D. An operational perspective on potential uses and constraints of emerging tools for monitoring water quality. **TrAC - Trends in Analytical Chemistry**, v. 29, n. 5, p. 378–384, 2010.

GUEVARA, Y. Z. C.; SOUZA, J. J. L. L. DE; VELOSO, G. V.; VELOSO, R. W.; ROCHA, P. A.; ABRAHÃO, W. A. P.; FERNANDES FILHO, E. I. Reference values of soil quality for the Rio Doce Basin. **Revista Brasileira de Ciencia do Solo**, v. 42, p. 1–16, 2018.

GUO, G.; DUAN, R. Simulation and assessment of a water pollution accident caused by phenol leakage. **Water Policy**, v. 23, n. 3, p. 750–764, 2021.

HAIR JR, J. F.; BLACK, W. C.; BABIN, B. J.; ANDERSON, R. E.; TATHAM, R. L. **Multivariate data analysis**. 6^a ed. New Jersey: Prentice Hall, 2006.

HAJIGHOLIZADEH, M.; MELESSE, A. M. Assortment and spatiotemporal analysis of surface water quality using cluster and discriminant analyses. **Catena**, v. 151, p. 247–258, 2017.

HALABI, A. L. M.; SIACARA, A. T.; SAKANO, V. K.; PILEGGI, R. G.; FUTAI, M. M. Tailings Dam Failures: A historical analysis of the risk. **Journal of Failure Analysis and Prevention**, v. 22, n. 2, p. 464–477, 2022.

HAMED, M. A. R. Application of surface water quality classification models using principal components analysis and cluster analysis. **Water and Energy International**, v. 62r, n. 1, p. 54–62, 2019.

HÄRDLE, W. K.; SIMAR, L. **Applied multivariate statistical analysis, fourth edition**. 4^a ed. Berlin: Springer, 2015.

HATJE, V.; PEDREIRA, R. M. A.; REZENDE, C. E. DE; SCHETTINI, C. A. F.; SOUZA, G. C. DE; MARIN, D. C.; HACKSPACHER, P. C. The environmental impacts of one of

the largest tailing dam failures worldwide. **Scientific Reports**, v. 7, n. 1, p. 1–13, 2017.

HAUSSER, J.; STRIMMER, K. **Entropy: Estimation of entropy, mutual information and related quantities**, 2021. Disponível em: <<https://cran.r-project.org/package=entropy>>

HE, F.; MA, J.; LAI, Q.; SHUI, J.; LI, W. Environmental impact assessment of a wharf oil spill emergency on a river water source. **Water (Switzerland)**, v. 15, n. 2, p. 1-16, 2023.

HLAVÁČ, V. **Theory and practice of informatics : proceedings / SOFSEM 2000, 27th Conference on Current Trends in Theory and Practice of Informatics**. Verlag Berlin Heidelberg NewYork: Springer, 2000. v. 1963

HOU, D.; GE, X.; HUANG, P.; ZHANG, G.; LOÁICIGA, H. A real-time, dynamic early-warning model based on uncertainty analysis and risk assessment for sudden water pollution accidents. **Environmental Science and Pollution Research**, v. 21, n. 14, p. 8878–8892, 2014.

HOU, Y. Environmental accident and its treatment in a developing country: A case study on China. **Environmental Monitoring and Assessment**, v. 184, n. 8, p. 4855–4859, 2012.

ICAGA, Y. Genetic algorithm usage in water quality monitoring networks optimization in Gediz (Turkey) River Basin. **Environmental Monitoring and Assessment**, v. 108, n. 1–3, p. 261–277, 2005.

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA (IBGE). **Manual Técnico de Pedologia**. 3ª edição. ed. Rio de Janeiro: IBGE, 2015. v. 3

INSTITUTO BRASILEIRO DO MEIO AMBIENTE E DOS RECURSOS NATURAIS RENOVÁVEIS (IBAMA). **Laudo técnico preliminar: Impactos ambientais decorrentes do desastre envolvendo o rompimento da barragem de Fundão, em Mariana, Minas Gerais**. Brasília: IBAMA, 2015. Disponível em: <<https://jornalismosocioambiental.files.wordpress.com/2016/01/laudo-preliminar-do-ibama-sobre-mariana.pdf>>.

INSTITUTO MINEIRO DE GESTÃO DAS ÁGUAS (IGAM). **Encarte especial: qualidade das águas do rio Doce após 4 anos do rompimento da barragem de Fundão - 2015/2019.** Belo Horizonte: IGAM, 2020a. Disponível em: <<http://www3.ana.gov.br/portal/ANA/sala-de-situacao/rio-doce/documentos-relacionados/encarte-qualidade-da-gua-do-rio-doce-dois-anos-apos-rompimento-de-barragem-de-fundao-1.pdf>>.

INSTITUTO MINEIRO DE GESTÃO DAS ÁGUAS (IGAM). **Encarte especial sobre a Qualidade das Águas do rio Doce após 5 anos do rompimento da Barragem de Fundão: 2015/2020.** Belo Horizonte: IGAM, 2020b. Disponível em: <<http://www3.ana.gov.br/portal/ANA/sala-de-situacao/rio-doce/documentos-relacionados/encarte-qualidade-da-gua-do-rio-doce-dois-anos-apos-rompimento-de-barragem-de-fundao-1.pdf>>.

JACOMINE, P. K. T. A nova classificação brasileira de solos. **Anais da Academia Pernambucana de Ciência Agrônômica**, v. 5 e 6, p. 161–179, 2009.

JALILI, S.; HOSSEINZADEH, Y. Design optimization of truss structures with continuous and discrete variables by hybrid of biogeography-based optimization and differential evolution methods. **Structural Design of Tall and Special Buildings**, v. 27, n. 14, p. 1–26, 2018.

JARDIM, F. A.; SPERLING, E. VON; JARDIM, B. F. DE M.; ALMEIDA, K. C. DE B. Determinants of cyanobacteria's bloom in water at Doce River, Minas Gerais, Brazil. **Engenharia Sanitaria e Ambiental**, v. 19, n. 3, p. 207–218, 2014.

JAYNES, E. T. Information theory and statistical mechanics. **Physical Review**, v. 106, n. 4, p. 620–630, 1957a.

JAYNES, E. T. Information theory and statistical mechanics II. **Physical review**, v. 108, n. 2, p. 171–190, 1957b.

JESUS, E. T. DE; AMORIM, J. DA S.; JUNQUEIRA, R.; VIOLA, M. R.; MELLO, C. R. DE. Meteorological and hydrological drought from 1987 to 2017 in Doce river basin, Southeastern Brazil. **Revista Brasileira de Recursos Hidricos**, v. 25, p. 1–10, 2020.

JIANG, J.; TANG, S.; HAN, D.; FU, G.; SOLOMATINE, D.; ZHENG, Y. A comprehensive review on the design and optimization of surface water quality monitoring networks. **Environmental Modelling and Software**, v. 132, n. 104792, p. 1–17, 2020a.

JIANG, D. YI; WANG, Y. YAN; LIAO, Q.; LONG, Z.; ZHOU, S. YANG. Assessment of water quality and safety based on multi-statistical analyses of nutrients, biochemical indexes and heavy metals. **Journal of Central South University**, v. 27, n. 4, p. 1211–1223, 2020b.

JING, P.; YANG, Z.; ZHOU, W.; HUAI, W.; LU, X. Inversion of multiple parameters for river pollution accidents using emergency monitoring data. **Water Environment Research**, v. 91, n. 8, p. 731–738, 2019.

JORDÃO, C. P.; PEREIRA, J. C.; BRUNE, W.; PEREIRA, J. L.; BRAATHEN, P. C. Heavy metal dispersion from industrial wastes in the vale do aço, minas gerais, Brazil. **Environmental Technology (United Kingdom)**, v. 17, n. 5, p. 489–500, 1996.

KAFANTARIS, E.; PIPER, I.; LO, T. Y. M.; ESCUDERO, J. Augmentation of dispersion entropy for handling missing and outlier samples in physiological signal monitoring. **Entropy**, v. 22, n. 3, p. 1-23, 2020.

KÄNDLER, M.; BLECHINGER, K.; SEIDLER, C.; PAVLŮ, V.; ŠANDA, M.; DOSTÁL, T.; KRÁSA, J.; VITVAR, T.; ŠTICH, M. Impact of land use on water quality in the upper Nisa catchment in the Czech Republic and in Germany. **Science of the Total Environment**, v. 586, p. 1316–1325, 2017.

KAO, J. J.; LI, P. H.; HU, W. S. Optimization models for siting water quality monitoring stations in a catchment. **Environmental Monitoring and Assessment**, v. 184, p. 43–52, 2012.

KAO, J. J.; LI, P. H.; LIN, C. L.; HU, W. H. Siting analyses for water quality sampling in a catchment. **Environmental Monitoring and Assessment**, v. 139, p. 205–215, 2008.

KARAMOUZ, M.; KERACHIAN, R.; AKHBARI, M.; HAFEZ, B. Design of river water

quality monitoring networks: A case study. **Environmental Modeling and Assessment**, v. 14, n. 6, p. 705–714, 2009a.

KARAMOUZ, M.; NOKHANDAN, A. K.; KERACHIAN, R.; MAKSIMOVIC, Č. Design of on-line river water quality monitoring systems using the entropy theory: A case study. **Environmental Monitoring and Assessment**, v. 155, n. 1–4, p. 63–81, 2009b.

KAVEH, A.; ILCHI GHAZAAN, M. Enhanced colliding bodies optimization for design problems with continuous and discrete variables. **Advances in Engineering Software**, v. 77, p. 66–75, 2014.

KESSLER, W.; MÜLLER, G. Minor and trace-element data of iron oxides from iron-formations of the Iron Quadrangle, Minas Gerais, Brazil. **Mineralogy and Petrology**, v. 39, n. 3–4, p. 245–250, 1988.

KETTENRING, J. R. The practice of cluster analysis. **Journal of Classification**, v. 23, n. 1, p. 3–30, 2006.

KEUM, J.; COULIBALY, P. Sensitivity of Entropy Method to Time Series Length in Hydrometric Network Design. **Journal of Hydrologic Engineering**, v. 22, n. 7, p. 1–13, 2017a.

KEUM, J.; COULIBALY, P. Information theory-based decision support system for integrated design of multivariable hydrometric networks. **Journal of the American Water Resources Association**, v. 53, p. 6239–6259, 2017b.

KEUM, J.; KORNELSEN, K. C.; LEACH, J. M.; COULIBALY, P. Entropy applications to water monitoring network design: A review. **Entropy**, v. 19, n. 11, p. 1–21, 2017.

KHALIL, B.; OU, C.; PROULX-MCINNIS, S.; ST-HILAIRE, A.; ZANACIC, E. Statistical assessment of the surface water quality monitoring network in Saskatchewan. **Water, Air, and Soil Pollution**, v. 225, n. 10, p. 1–22, 2014.

KHALIL, B.; OUARDA, T. B. M. J.; ST-HILAIRE, A. A statistical approach for the assessment and redesign of the Nile Delta drainage system water-quality-monitoring locations. **Journal of Environmental Monitoring**, v. 13, p. 2190–2205, 2011.

KHALIL, B.; OUARDA, T. B. M. J.; ST-HILAIRE, A.; CHEBANA, F. A statistical approach for the rationalization of water quality indicators in surface water quality monitoring networks. **Journal of Hydrology**, v. 386, n. 1–4, p. 173–185, 2010.

KHISHE, M.; MOSAVI, M. R. Chimp optimization algorithm. **Expert Systems with Applications**, v. 149, p. 1–26, 2020.

KONG, Y. S.; KHAN, R. To examine environmental pollution by economic growth and their impact in an environmental Kuznets curve (EKC) among developed and developing countries. **PLoS ONE**, v. 14, n. 3, p. 1–23, 2019.

KOO, Y. H.; YANG, Y. S.; SONG, K. W. Radioactivity release from the Fukushima accident and its consequences: A review. **Progress in Nuclear Energy**, v. 74, p. 61–70, 2014.

KOSSOFF, D.; DUBBIN, W. E.; ALFREDSSON, M.; EDWARDS, S. J.; MACKLIN, M. G.; HUDSON-EDWARDS, K. A. Mine tailings dams: Characteristics, failure, environmental impacts, and remediation. **Applied Geochemistry**, v. 51, p. 229–245, 2014.

KÜTTER, V. T.; MARTINS, G. S.; BRANDINI, N.; CORDEIRO, R. C.; ALMEIDA, J. P. A.; MARQUES, E. D. Impacts of a tailings dam failure on water quality in the Doce river: The largest environmental disaster in Brazil. **Journal of Trace Elements and Minerals**, v. 5, n. August 2022, p. 1-13, 2023.

LAU, L. C.; LEE, K. T.; MOHAMED, A. R. Global warming mitigation and renewable energy policy development from the Kyoto Protocol to the Copenhagen Accord - A comment. **Renewable and Sustainable Energy Reviews**, v. 16, n. 7, p. 5280–5284, 2012.

LE, S.; JOSSE, J.; HUSSON, F. FactoMineR: A package for multivariate analysis. **Journal of Statistical Software**, v. 25, n. 1, p. 1–18, 2008.

LEE, C.; PAIK, K.; YOO, D. G.; KIM, J. H. Efficient method for optimal placing of water quality monitoring stations for an ungauged basin. **Journal of Environmental Management**, v. 132, p. 24–31, 2014.

LI, C.; SINGH, V. P.; MISHRA, A. K. Entropy theory-based criterion for hydrometric network evaluation and design: Maximum information minimum redundancy. **Water Resources Research**, v. 48, n. 5, p. 1–15, 2012.

LI, P. Y.; QIAN, H.; WU, J. H. Groundwater Quality Assessment Based on Improved Water Quality Index in Pengyang County, Ningxia, Northwest China. **E-Journal of Chemistry**, v. 7, n. S1, p. 1–9, 2010.

LI, P.; TIAN, R.; LIU, R. Solute geochemistry and multivariate analysis of water quality in the Guohua phosphorite mine, Guizhou Province, China. **Exposure and Health**, v. 11, n. 2, p. 81–94, 2019.

LI, M.; YAO, X. Quality evaluation of solution sets in multiobjective optimisation: A survey. **ACM Computing Surveys**, v. 52, n. 2, p. 1-38, 2019.

LI, R.; DUAN, X.; GUO, X.; HE, W.; LV, Y. Adaptive compressive sensing of images using spatial entropy. **Computational Intelligence and Neuroscience**, v. 2017, p. 1–9, 2017.

LIU, Y.; ZHENG, B. H.; WANG, M.; XU, Y. X.; QIN, Y. W. Optimization of sampling frequency for routine river water quality monitoring. **Science China Chemistry**, v. 57, n. 5, p. 772–778, 2014.

LIYANAGE, C. P.; MARASINGHE, A.; YAMADA, K. Comparison of optimized selection methods of sampling sites network for water quality monitoring in a river. **International Journal of Affective Engineering**, v. 15, n. 2, p. 195–204, 2016.

LOBO, E. A.; SCHUCH, M.; HEINRICH, C. G.; COSTA, A. BEN DA; DÜPONT, A.; WETZEL, C. E.; ECTOR, L. Development of the trophic water quality index (TWQI) for subtropical temperate Brazilian lotic systems. **Environmental Monitoring and Assessment**, v. 187, n. 6, p. 1-13, 2015.

LUSWETI, E.; KANDA, E. K.; OBANDO, J.; MAKOKHA, M. Effects of oil exploration on surface water quality – a review. **Water Practice and Technology**, v. 17, n. 10, p. 2171–2185, 2022.

LUZ, N. DA; TOBIASON, J. E.; KUMPEL, E. Water quality monitoring with purpose: Using a novel framework and leveraging long-term data. **Science of the Total Environment**, v. 818, p. 1-12, , 2022.

MAHJOURI, N.; KERACHIAN, R. Revising river water quality monitoring networks using discrete entropy theory: The Jajrood River experience. **Environmental Monitoring and Assessment**, v. 175, p. 291–302, 2011.

MAITY, R. **Statistical methods in hydrology and hydroclimatology**. Singapore: Springer, 2018. v. 555

MANLEY, S. F.; CONNER, R.; HOLTHOUSE PUTZ, A. R. Chicago's response to a Hexavalent Chromium Spill. **Journal - American Water Works Association**, v. 112, n. 3, p. 22–29, 2020.

MANNA, A.; BISWAS, D. Assessment of drinking water quality using water quality index: A review. **Water Conservation Science and Engineering**, v. 8, n. 1, p. 23–25, 2023.

MARINHO JÚNIOR, L. D.; ESPINDOLA, H. S.; NODARI, E. S. A charcoal frontier: the steel industry and forest in twentieth-century Minas Gerais (Brazil). **Ambiente & Sociedade**, v. 27, p. 1-24, 2024.

MARINHO, G.; CAMPOS, D. F.; TEIXEIRA, A.; TREMBLAY, M.; MENDIONDO, E. M.; GHIGLIENO, F. Advances in technological research for online and in situ water quality monitoring — A review. **sustainability**, v. 14, n. 5059, p. 1–28, 2022.

MARINHO, L. D.; ESPINDOLA, H. S. Iron Mines, forest and rivers: Environmental Impacts of iron factories in Brazilian Minas Gerais State in 19th Century. **Historia Ambiental Latinoamericana y Caribena**, v. 11, n. 3, p. 93–117, 2021.

MARTINS, B. D. A.; TAKAHASHI, J. A. Metal - Rich Mine - Tailing spills in Brazil and the consequences for the surrounding water bodies. **Water, Air, & Soil Pollution**, p. 1–14, 2022.

MATSUMOTO, M.; NISHIMURA, T. Mersenne Twister: A 623-dimensionally

equidistributed uniform pseudorandom number generator. **Discrete Mathematics**, 1998.

MAVUKKANDY, M. O.; KARMAKAR, S.; HARIKUMAR, P. S. Assessment and rationalization of water quality monitoring network: A multivariate statistical approach to the Kabbini River (India). **Environmental Science and Pollution Research**, v. 21, p. 10045–10066, 2014.

MAZZOTTI, B. F. **Co-projeto de Hardware/Software do filtro de partículas para localização em tempo real de robôs móveis**. 2009. Dissertação (Mestrado em Ciências de Computação e Matemática Computacional) - Universidade de São Paulo, São Paulo, 2009.

MEDEIROS, A. O.; MISSAGIA, B. S.; BRANDÃO, L. R.; CALLISTO, M.; BARBOSA, F. A. R.; ROSA, C. A. Water quality and diversity of yeasts from tropical lakes and rivers from the Rio Doce basin in Southeastern Brazil. **Brazilian Journal of Microbiology**, v. 43, n. 4, p. 1582–1594, 2012.

MELLO, C. L.; METELO, C. M. S.; SUGUIO, K.; KOHLER, H. C. Quaternary sedimentation, neotectonics and the evolution of the Doce river middle valley lake system (Southeastern Brazil). **Revista do Instituto Geológico**, v. 20, p. 29–36, 1999.

MENDES, L. B.; MELLO, F. A. DE; CHAGAS, K. R.; CAMPELO, R. P. M.; MEDEIROS, L. C. C.; SMITH, R. E. W.; FURLEY, T. H. Ecotoxicological assessment of the Doce River surface water after the Fundão dam collapse. **Integrated Environmental Assessment and Management**, v. 16, n. 5, p. 608–614, 2020.

MENDES, R. G.; VALLE JUNIOR, R. F. DO; MELO SILVA, M. M. A. P. DE; MORAIS FERNANDES, G. H. DE; FERNANDES, L. F. S.; FERNANDES, A. C. P.; PISSARRA, T. C. T.; MELO, M. C. DE; VALERA, C. A.; PACHECO, F. A. L. A partial least squares-path model of environmental degradation in the Paraopeba River, for rainy seasons after the rupture of B1 tailings dam, Brumadinho, Brazil. **Science of the Total Environment**, v. 851, p. 1–14, 2022.

MEYER, P. E. **infotheo: Information-theoretic measures**, 2022. Disponível em: <<https://cran.r-project.org/package=infotheo>>

MILLER, M. E.; GHISOLFI, R. D.; BARROSO, G. F. Remote sensing monitoring of mining tailings in the fluvial-estuarine-coastal ocean continuum of the Lower Doce River Valley (Brazil). **Environmental monitoring and assessment**, v. 195, n. 5, p. 1-20, 2023.

MINAS GERAIS. Deliberação Normativa Conjunta COPAM-CERH/MG nº 8, de 21 de novembro de 2022. Estabelece diretrizes e procedimentos para o licenciamento ambiental de atividades e empreendimentos no estado de Minas Gerais. **Diário Oficial do Estado de Minas Gerais**, Belo Horizonte, [data da publicação]. Disponível em: <<http://www.siam.mg.gov.br/sla/download.pdf?idNorma=56521>>. Acesso em: 17 out. 2023.

MINGOTI, S. A. **Análise de dados através de métodos de estatística multivariada: uma abordagem aplicada**. 1ª reimpre ed. Belo Horizonte: Editora UFMG, 2005.

MIRAUDA, D.; OSTOICH, M. MIMR criterion application: Entropy approach to select the optimal quality parameter set responsible for river pollution. **Sustainability (Switzerland)**, v. 12, n. 5, p. 1-22, 2020.

MLADEN A. VOUK. Cloud computing – Issues, research and implementations. **Acta Geologica Hispanica**, v. 32, n. 3–4, p. 171–184, 1997.

MOREIRA, E.; HERMÍNIO, S.; NALINI, A.; ADRIANA, J.; ABREU, T.; BRANDÃO, L. Mobilization of heavy metals in river sediments from the region impacted by the Fundão dam rupture , Brazil. **Environmental Earth Sciences**, v. 80, n. 24, p. 1–10, 2021.

MUÑOZ-COBO, J. L.; MENDIZÁBAL, R.; MIQUEL, A.; BERNA, C.; ESCRIVÁ, A. Use of the principles of maximum entropy and maximum relative entropy for the determination of uncertain parameter distributions in engineering applications. **Entropy**, v. 19, n. 9, p. 1-37, 2017.

NAIR, U. S. The standard error of Gini's Mean Difference. **Biometrika**, v. 28, n. 3, p. 428–436, 1936.

NGUYEN, T. H.; HELM, B.; HETTIARACHCHI, H.; CAUCCI, S.; KREBS, P. The selection of design methods for river water quality monitoring networks: a review.

Environmental Earth Sciences, v. 78, n. 3, p. 1–17, 2019.

NING, S. K.; CHANG, N. BIN. Screening the relocation strategies of water quality monitoring stations by compromise programming. **Journal of the American Water Resources Association**, v. 41, n. 5, p. 1039–1052, 2005.

NOGUEIRA, L. B.; SOUSA, S. M.; SANTOS, C. G. L.; ARAÚJO, G. S.; OLIVEIRA, L.; NOGUEIRA, K. O. P. C. Water quality from gualaxo do Norte and Carmo Rivers (Minas Gerais, Brazil) after the fundão dam failure. **Anuario do Instituto de Geociencias**, v. 44, p. 1–11, 2021.

NOOGHABI, M. J.; NOOGHABI, E. K. On entropy of a Pareto distribution in the presence of outliers. **Communications in Statistics - Theory and Methods**, v. 45, n. 17, p. 5234–5250, 2016.

OEHRIG, J.; KANANIZADEH, N.; WILD, M.; ROUHANI, S.; ODLE, W. Applying multivariate techniques to fingerprint water quality impact of the Fundão Dam breach within the Rio Doce basin. **Integrated Environmental Assessment and Management**, v. 00, n. 00, p. 1–15, 2023.

OJIMA, D. S.; MELILLO, J. M. Looking Beyond Stockholm +50: The ecological science and moving toward solutions. **The Bulletin of the Ecological Society of America**, v. 103, n. 1, p. 1–3, 2022.

OLIVEIRA SERRANO, L. DE; BORGES, A. C.; PRUSKI, F. F.; MELO, M. C. DE. A new approach to use load duration curves to evaluate water quality: A study in the Doce River basin, Brazil. **Water (Switzerland)**, v. 12, n. 3, p. 1–21, 2020.

OLIVEIRA, K. S. S.; QUARESMA, V. DA S. Temporal variability in the suspended sediment load and streamflow of the Doce River. **Journal of South American Earth Sciences**, v. 78, p. 101–115, 2017.

OLIVER, M. A.; WEBSTER, R. A tutorial guide to geostatistics: Computing and modelling variograms and kriging. **Catena**, v. 113, p. 56–69, 2014.

OLSEN, R. L.; CHAPPELL, R. W.; LOFTIS, J. C. Water quality sample collection, data

treatment and results presentation for principal components analysis - literature review and Illinois River watershed case study. **Water Research**, v. 46, n. 9, p. 3110–3122, 2012.

OUYANG, H.-T.; YU, H.; LU, C.-H.; LUO, Y.-H. Design optimization of river sampling network using Genetic Algorithms. **Journal of Water Resources Planning and Management**, v. 134, n. 1, p. 83–87, 2008.

P FERNANDES, A. C.; R FONSECA, A.; PACHECO, F. A. L.; SANCHES FERNANDES, L. F. Water quality predictions through linear regression - A brute force algorithm approach. **MethodsX**, v. 10, n. March, p. 1-22, 2023.

PACHECO, F. A. L.; VALLE JUNIOR, R. F. DO; MELO SILVA, M. M. A. P. DE; TARLÉ PISSARRA, T. C.; SOUZA ROLIM, G. DE; MELO, M. C. DE; VALERA, C. A.; MOURA, J. P.; SANCHES FERNANDES, L. F. Geochemistry and contamination of sediments and water in rivers affected by the rupture of tailings dams (Brumadinho, Brazil). **Applied Geochemistry**, v. 152, p. 1–17, 2023.

PÁDUA, L. H. R. DE; NASCIMENTO, N. DE O.; SILVA, F. E. O. E.; ALFONSO, L. Analysis of the fluviometric network of rio das velhas using entropy. **Revista Brasileira de Recursos Hidricos**, v. 24, p. 1–14, 2019.

PALMER, R. N.; ASCE, A. M.; MACKENZIE, M. C. Optimization of water quality monitoring networks. **Journal of Water Resources Planning and Management**, v. 111, n. 4, p. 478–493, 1985.

PARAGUASSÚ, L.; LEITE, M. G. P.; MOREIRA, F. W. A.; MENDONÇA, F. P. C.; ESKINAZI-SANT'ANNA, E. M. Impacts of mining in artificial lake of Iron Quadrangle-MG: past marks and changes of the present. **Environmental Earth Sciences**, v. 78, n. 5, p. 1–10, 2019.

PARK, S. Y.; CHOI, J. H.; WANG, S.; PARK, S. S. Design of a water quality monitoring network in a large river system using the genetic algorithm. **Ecological Modelling**, v. 199, p. 289–297, 2006.

PASSOS, J. B. DE M. C.; TEIXEIRA, D. B. DE S.; CAMPOS, J. A.; LIMA, R. P. C.;

FERNANDES-FILHO; INÁCIO, E.; SILVA; DA, D. D. Multivariate statistics for spatial and seasonal quality assessment of water in the Doce River basin, Southeastern Brazil. **Environ Monit Assess**, v. 193, n. 3, p. 1–16, 2021.

PEDROSA, P. Optical resilience of the Paraíba do Sul River (Brazil) during a toxic spill of a wood-pulping factory: The Cataguazes accident. **Environmental Monitoring and Assessment**, v. 129, n. 1–3, p. 137–150, 2007.

PEÑA-GUZMÁN, C. A.; SOTO, L.; DIAZ, A. A proposal for redesigning the water quality network of the Tunjuelo River in Bogotá, Colombia through a spatio-temporal analysis. **Resources**, v. 8, n. 2, p. 1-14, 2019.

PERDIGÃO, R. A. P.; EHRET, U.; KNUTH, K. H.; WANG, J. Debates: Does information theory provide a new paradigm for earth science? Emerging concepts and pathways of information physics. **Water Resources Research**, v. 56, n. 2, p. 1–13, 2020.

PEREIRA, S. F. P.; ROCHA, R. M.; PINHEIRO, L. S.; NOGUEIRA, D. P. Integration of statistical models and computer simulation in environmental accidents: A study on leakage of red mud in the Pará River, Amazon, Brazil. **Journal of the Brazilian Chemical Society**, v. 32, n. 10, p. 1997–2008, 2021.

PÉREZ, C. J.; VEGA-RODRÍGUEZ, M. A.; REDER, K.; FLÖRKE, M. A multi-objective artificial bee colony-based optimization approach to design water quality monitoring networks in river basins. **Journal of Cleaner Production**, v. 166, p. 579–589, 2017.

PETRUCIO, M. M.; MEDEIROS, A. O.; ROSA, C. A.; BARBOSA, F. A. R. Trophic state and microorganisms community of major sub-basins of the middle Rio Doce basin, southeast Brazil. **Brazilian Archives of Biology and Technology**, v. 48, n. 4, p. 623–633, 2005.

PINTO, C. C.; ALMEIDA, K. DE B.; OLIVEIRA, S. C. Spatial evaluation of the water quality from the from the Velhas river channel, in the state of Minas Gerais. **Periódico tchê química**, v. 15, n. 30, p. 75–86, 2018.

POURSHAHABI, S. ET AL. An Entropy-Based Approach to Fuzzy Multi-objective Optimization of Reservoir Water Quality Monitoring Networks Considering

Uncertainties. **Water Resources Management**, v. 32, p. 4425–4443, 2018.

PURI, D.; BOREL, K.; VANCE, C.; KARTHIKEYAN, R. Optimization of a water quality monitoring network using a spatially referenced water quality model and a genetic algorithm. **Water**, v. 9, n. 704, p. 1–11, 2017.

PURI, D.; KARTHIKEYAN, C. V. AND R. Optimization of a Water Quality Monitoring Network Using a Spatially Referenced Water Quality Model and a Genetic Algorithm. **Water**, v. 9, n. 704, p. 1–11, 2017.

QUEIROZ, H. M.; NÓBREGA, G. N.; FERREIRA, T. O.; ALMEIDA, L. S.; ROMERO, T. B.; SANTAELLA, S. T.; BERNARDINO, A. F.; OTERO, X. L. The Samarco mine tailing disaster: A possible time-bomb for heavy metals contamination? **Science of the Total Environment**, v. 637–638, p. 498–506, 2018.

QUEIROZ, H. M.; FERREIRA, T. O.; BARCELLOS, D.; NÓBREGA, G. N.; ANTELO, J.; OTERO, X. L.; BERNARDINO, A. F. From sinks to sources: The role of Fe oxyhydroxide transformations on phosphorus dynamics in estuarine soils. **Journal of Environmental Management**, v. 278, n. October 2020, p. 1-11, 2021a.

QUEIROZ, H. M.; YING, S. C.; BERNARDINO, A. F.; BARCELLOS, D.; NÓBREGA, G. N.; OTERO, X. L.; FERREIRA, T. O. Role of Fe dynamic in release of metals at Rio Doce estuary: Unfolding of a mining disaster. **Marine Pollution Bulletin**, v. 166, p. 1–7, 2021b.

R CORE TEAM. **R: A language and environment for statistical computing**. Vienna, Austria, 2021. Disponível em: <<https://www.r-project.org/>>

RADELYUK, I.; TUSSUPOVA, K.; KLEMEŠ, J. J.; PERSSON, K. M. Oil refinery and water pollution in the context of sustainable development: Developing and developed countries. **Journal of Cleaner Production**, v. 302, p. 1-13, 2021.

RAVEN, P. J.; GEORG, J. J. Recovery by riffle macroinvertebrates in a river after a major accidental spillage of Chiorpyrifos. **Environmental pollution**, v. 59, p. 55–70, 1989.

RAZMAN, N. A.; ISMAIL, W. Z. W.; KAMIL, N. A. I. M.; ZAINURIN, S. N.; ISMAIL, I.; JAMALUDIN, J.; SAHRIM, M.; ARIFFIN, K. N. Z.; BALAKRISHNAN, S. R. A review on water quality monitoring methods based on electronics and optical sensing. **Journal of advanced research in applied sciences and engineering technology**, v. 2, n. 2, p. 1–7, 2022.

REINA-GARCÍA, J.; TORO-VÉLEZ, A. F.; PEÑA-VARÓN, M. R.; OLAYA-OCHOA, J.; FIGUEROA-CASAS, A. Methodological design for the macro-location of a micropollutants monitoring network in tropical rivers: a case study in Cauca River. **Environmental Monitoring and Assessment**, v. 192, n. 4, p. 1-13, 2020.

REIS BRAGA, M. C. DOS; CASTRO FONSECA, A. DE F. Reactive improvement of environmental policies: lessons from the Mariana and Brumadinho disasters. **Sustentabilidade em Debate**, v. 12, n. 3, p. 182–196, 2021.

REIS, D. A. DOS; FONSECA SANTIAGO, A. DA; NASCIMENTO, L. P.; ROESER, H. M. P. Influence of environmental and anthropogenic factors at the bottom sediments in a Doce River tributary in Brazil. **Environmental Science and Pollution Research**, v. 24, n. 8, p. 7456–7467, 2017.

RENOVA. **Programa de monitoramento quali-quantitativo sistemático de água e sedimentos – PMQQS Relatório técnico dezembro 2017**. Belo Horizonte: RENOVA, 2017a.

RENOVA. **Programa de monitoramento quali-quantitativo sistemático de água e sedimentos – PMQQS Relatório técnico abril 2017**. Belo Horizonte: RENOVA, 2017b.

RENOVA. **PMQQS: Relatório parcial - Junho 2018**. Belo Horizonte: RENOVA, 2018.

RICO, M.; BENITO, G.; DÍEZ-HERRERO, A. Floods from tailings dam failures. **Journal of Hazardous Materials**, v. 154, n. 1–3, p. 79–87, 2008.

ROCHE, C.; THYGESEN, K.; BAKER, E. **Mine tailings storage: Safety is no accident**. UNEP: United Nations Environment Programme and GRID-Arendal, Nairobi and Arendal, 2017.

RODRIGUES, V.; ESTRANY, J.; RANZINI, M.; CICCIO, V. DE; MARTÍN-BENITO, J. M. T.; HEDO, J.; LUCAS-BORJA, M. E. Effects of land use and seasonality on stream water quality in a small tropical catchment: The headwater of Córrego Água Limpa, São Paulo (Brazil). **Science of the Total Environment**, v. 622–623, p. 1553–1561, 2018.

SÁ, F.; LONGHINI, C. M.; COSTA, E. S.; SILVA, C. A. DA; CAGNIN, R. C.; GOMES, L. E. DE O.; LIMA, A. T.; BERNARDINO, A. F.; NETO, R. R. Time-sequence development of metal(loid)s following the 2015 dam failure in the Doce river estuary, Brazil. **Science of the Total Environment**, v. 769, p. 1–12, 2021.

SABZIPOUR, B.; ASGHARI, O.; SARANG, A. Evaluation and optimal redesigning of river water-quality monitoring networks (RWQMN) using geostatistics approach (case study: Karun, Iran). **Sustainable Water Resources Management**, v. 5, n. 2, p. 439–455, 2019.

SALMAN, R.; NIKOO, M. R.; SHOJAEZADEH, S. A.; BEIGLOU, P. H. B.; SADEGH, M.; ADAMOWSKI, J. F.; ALAMDARI, N. A novel Bayesian maximum entropy-based approach for optimal design of water quality monitoring networks in rivers. **Journal of Hydrology**, v. 603, p. 1–15, 2021.

SÁNCHEZ, L. E.; ALGER, K.; ALONSO, L.; BARBOSA, F. A. R.; BRITO, M. C. W.; LAUREANO, F. V.; MAY, P.; ROESER, H.; KAKABADSE, Y. **Os impactos do rompimento da Barragem de Fundão: O caminho para uma mitigação sustentável e resiliente**. Gland/Suíça: IUCN, 2018.

SANTOLIN, C. V. A.; CIMINELLI, V. S. T.; NASCENTES, C. C.; WINDMOÖLLER, C. C. Distribution and environmental impact evaluation of metals in sediments from the Doce River Basin , Brazil. **Environ Earth Sci**, v. 74, p. 1235–1248, 2015.

SANTOS, G.; REIS, D.; SANTOS, I.; LIMA, L.; ROCHA, C.; CASTRO, P.; LOBO, F.; SANTIAGO, A.; SILVA, G. DA. Assessment of Gualaxo do Norte River water quality (Minas Gerais, Brazil) affected by the dam breach of Fundão utilizing exploratory multivariate techniques. **Environmental Monitoring and Assessment**, v. 195, n. 2, p. 1–13, 2023.

SAPATA GONZALEZ, R.; APARECIDA DA SILVEIRA ROSSI, R.; GUSTAVO MARTINS VIEIRA, L. Economic and financial consequences of process accidents in Brazil: Multiple case studies. **Engineering Failure Analysis**, v. 132, n. April 2021, p. 1-24, 2022.

SCHLIEMANN, S. A.; GREVSTAD, N.; BRAZEAU, R. H. Water quality and spatio-temporal hot spots in an effluent-dominated urban. **Hydrological Processes**, v. 35, p. 1–17, 2021.

SCOTT, D. W. On optimal and data-based histograms. **Biometika**, v. 66, n. 3, p. 605–610, 1979.

SEGURA, F. R.; NUNES, E. A.; PANIZ, F. P.; PAULELLI, A. C. C.; RODRIGUES, G. B.; BRAGA, G. Ú. L.; REIS PEDREIRA FILHO, W. DOS; BARBOSA, F.; CERCHIARO, G.; SILVA, F. F.; BATISTA, B. L. Potential risks of the residue from Samarco's mine dam burst (Bento Rodrigues, Brazil). **Environmental Pollution**, v. 218, p. 813–825, 2016.

SEMENOVA, G. Environmental disasters as a factor of environmental pollution. **E3S Web of Conferences**, v. 217, p. 1–8, 2020.

SEYFANG, G. Environmental mega-conferences - From Stockholm to Johannesburg and beyond. **Global Environmental Change**, v. 13, n. 3, p. 223–228, 2003.

SHANNON, C. E. A mathematical theory of communication. **Bell System Technical Journal**, v. 27, p. 379–423, 1948.

SHI, B.; JIANG, J.; SIVAKUMAR, B.; ZHENG, Y.; WANG, P. Quantitative design of emergency monitoring network for river chemical spills based on discrete entropy theory. **Water Research**, v. 134, p. 140–152, 2018.

SILVA, C. DE L. E; COSTA, A. T. DA; LANDA, G. G.; FONSECA, H. F. C.; SILVEIRA, A. Evaluation of sediment contamination by trace elements and the zooplankton community analysis in area affected by gold exploration in Southeast (SE) of the Iron Quadrangle, Alto Rio Doce, (MG) Brazil. **Acta Limnologica Brasiliensia**, v. 25, n. 2, p. 150–157, 2013.

SILVA, D. DE C.; BELLATO, C. R.; MARQUES NETO, J. DE O.; FONTES, M. P. F. Trace elements in river waters and sediments before and after a mining dam breach (Bento Rodrigues, Brazil). **Quimica Nova**, v. 41, n. 8, p. 857–866, 2018.

DA SILVA, D. L.; LIMA, A. R. B. DE; LIMA SOUZA, J. R. DE; ADAM, M. L. Environmental monitoring for genomic damage after an environmental accident in a river in the Brazilian Northeast. **Water Air Soil Pollut**, v. 233, n. 506, p. 1–10, 2022a.

DA SILVA, G. J.; BORGES, A. C.; MOREIRA, M. C.; ROSA, A. P. Statistical process control in assessing water quality in the Doce river basin after the collapse of the Fundão dam (Mariana, Brazil). **Journal of Environmental Management**, v. 317, n. May, p. 1–11, 2022b.

SILVA, K. R. DA. **Avaliação da qualidade da água de uma microbacia hidrográfica rural dos tabuleiros costeiros do Brasil**. 2020. Dissertação (Mestrado em Recursos Hídricos) - Universidade Federal de Sergipe, São Cristovão, 2020.

SIMEONOV, V.; SIMEONOVA, P.; TSAKOVSKI, S.; LOVCHINOV, V. Lake water monitoring data assessment by multivariate statistics. **Journal of Water Resource and Protection**, v. 02, n. 04, p. 353–361, 2010.

SINGH, K. R.; DUTTA, R.; KALAMDHAD, A. S.; KUMAR, B. An investigation on water quality variability and identification of ideal monitoring locations by using entropy based disorder indices. **Science of the Total Environment**, v. 647, p. 1444–1455, 2019a.

SINGH, K. R.; DUTTA, R.; KALAMDHAD, A. S.; KUMAR, B.. Information entropy as a tool in surface water quality assessment. **Environmental Earth Sciences**, v. 78, n. 1, p. 1–12, 2019b.

SINGH, V. P. **Entropy theory and its application in environmental and water engineering**. 1^a edition ed. Chichester: John Wiley & Sons, 2013.

SINGH, V. P. Entropy theory for derivation of infiltration equations. **Water Resources Research**, v. 46, n. 3, p. 1–20, 2010.

SOARES, P. F. **Projeto e avaliação de desempenho de redes de monitoramento**

de qualidade da água utilizando o conceito de entropia. 2001. Tese (Doutorado em Engenharia) - Universidade de São Paulo, São Paulo, 2001.

SOUZA FRAGA, M. DE; SILVA, D. D. DA; ALDEN ELESBON, A. A.; SOARES GUEDES, H. A. Methodological proposal for the allocation of water quality monitoring stations using strategic decision analysis. **Environmental Monitoring and Assessment**, v. 191, n. 12, p. 1-18, 2019.

SOUZA PEREIRA, M. A. DE; CAVALHERI, P. S.; OLIVEIRA, M. Â. C. DE; MAGALHÃES FILHO, F. J. C. A multivariate statistical approach to the integration of different land-uses, seasons, and water quality as water resources management tool. **Environmental Monitoring and Assessment**, v. 191, n. 9, p. 1-19, 2019.

SPERLING, M. VON. **Estudos e modelagem da qualidade da água de rios.** 1ª ed. Belo Horizonte: DESA/UFGM, 2007.

SPERLING, M. VON; VERBYLA, M. E.; OLIVEIRA, S. M. A. C. **Assessment of treatment plant performance and water quality data assessment of treatment plant.** London: IWA, 2020.

SPIER, C. A.; OLIVEIRA, S. M. B. DE; SIAL, A. N.; RIOS, F. J. Geochemistry and genesis of the banded iron formations of the Cauê Formation, Quadrilátero Ferrífero, Minas Gerais, Brazil. **Precambrian Research**, v. 152, n. 3-4, p. 170-206, 2007.

SREEPARVATHY, V.; SRINIVAS, V. V. A fuzzy entropy approach for design of hydrometric monitoring networks. **Journal of Hydrology**, v. 586, n. March, p. 1-18, 2020a.

STOSIC, T.; STOSIC, B.; SINGH, V. P. Optimizing streamflow monitoring networks using joint permutation entropy. **Journal of Hydrology**, v. 552, p. 306-312, 2017.

STROBL, R. O.; ROBILLARD, P. D. Network design for water quality monitoring of surface freshwaters: A review. **Journal of Environmental Management**, v. 87, n. 4, p. 639-648, 2008.

STROBL, R. O.; ROBILLARD, P. D.; SHANNON, R. D.; DAY, R. L.; MCDONNELL, A.

J. A water quality monitoring network design methodology for the selection of critical sampling points: Part I. **Environmental Monitoring and Assessment**, v. 112, n. 1–3, p. 137–158, 2006a.

STROBL, R. O.; ROBILLARD, P. D.; DAY, R. L.; SHANNON, R. D.; MCDONNELL, A. J. A water quality monitoring network design methodology for the selection of critical sampling points: Part II. **Environmental Monitoring and Assessment**, v. 122, n. 1–3, p. 319–334, 2006b.

STRONG, M. F. The promises and challenges of UNCED '92. **Ocean and Coastal Management**, v. 18, n. 1, p. 5–14, 1992.

STURGES, H. A. The choice of a class interval. **Journal of the American Statistical Association**, v. 21, n. 153, p. 65–66, 1926.

SUN, T.; WANG, X. Adoption of financial inclusion in a world of depleting natural resources: The importance of information and communication technology in emerging economies. **Resources Policy**, v. 85, n. PB, p. 1-20, 2023.

SUTADIAN, A. D.; MUTTIL, N.; YILMAZ, A. G.; PERERA, B. J. C. Development of river water quality indices: a review. **Environmental Monitoring and Assessment**, v. 188, n. 1, p. 1–29, 2016.

SZEKELY, G. J.; RIZZO, M. L. Hierarchical clustering via joint between-within distances: Extending Ward's minimum variance method. **Journal of Classification**, v. 22, n. 2, p. 151–183, 2005.

TAN, J. Water quality monitoring and standardization in China : A review. **IOP Conf. Series: Earth and Environmental Science** 546, v. 546, p. 1–5, 2020.

TANG, X.; ZHAI, A.; DING, X.; ZHU, Q. Safety guarantee system of drinking water source in Three Gorges reservoir area and its application in Huangjuedu drinking water source area. **Sustainability (Switzerland)**, v. 11, n. 24, p. 1-13, 2019.

TANOS, P.; KOVÁCS, J.; KOVÁCS, S.; ANDA, A.; HATVANI, I. G. Optimization of the monitoring network on the River Tisza (Central Europe, Hungary) using combined

cluster and discriminant analysis, taking seasonality into account. **Environmental Monitoring and Assessment**, v. 187, n. 9, p. 1-14, 2015.

TEIXEIRA, M. C.; SANTOS, A. C.; FERNANDES, C. S.; NG, J. C. Arsenic contamination assessment in Brazil – Past, present and future concerns: A historical and critical review. **Science of the Total Environment**, v. 730, p. 1-21, 2020.

TELICI, I. T.; NAM, K.; GUAN, J.; ARAL, M. M. Optimal water quality monitoring network design for river systems. **Journal of Environmental Management**, v. 90, n. 10, p. 2987–2998, 2009.

TERRELL, G. R.; SCOTT, D. W. Oversmoothed nonparametric density estimates. **Journal of the American Statistical Association**, v. 80, n. 389, p. 209–214, 1985.

TIMMERMAN, J. G.; BEINAT, E.; TERMEER, K.; COFINO, W. Analyzing the data-rich-but-information-poor syndrome in dutch water management in historical perspective. **Environmental Management**, v. 45, n. 5, p. 1231–1242, 2010.

UDDIN, M. G.; NASH, S.; OLBERT, A. I. A review of water quality index models and their use for assessing surface water quality. **Ecological Indicators**, v. 122, p. 1–21, 2021.

UNITED NATIONS ENVIRONMENT PROGRAMME (UNEP). **The Songua river spill China, December 2005**. China: UNEP, 2005.

UNITED NATIONS ENVIRONMENT PROGRAMME (UNEP). **The UN - Status report on the application of integrated approaches to water resources management** United Nations Environment Programme. Nairobi/Kenia: UNEP, 2012.

Disponível em: https://www.un.org/waterforlifedecade/pdf/un_water_status_report_2012.pdf.

UNITED STATES GEOLOGICAL SURVEY (USGS). **The strategy for improving water-quality monitoring in the United States: final report of the intergovernmental task force on monitoring water quality**. Reston/Virginia United States Geological Survey, 1995.

URSULAK, J.; COULIBALY, P. Integration of hydrological models with entropy and multi-objective optimization based methods for designing specific needs streamflow monitoring networks. **Journal of Hydrology**, v. 593, n. May 2020, p. 1-16, 2021.

VEGA-RODRÍGUEZ, M. A.; PÉREZ, C. J.; REDER, K.; FLÖRKE, M. A stage-based approach to allocating water quality monitoring stations based on the WorldQual model: The Jubba River as a case study. **Science of the Total Environment**, v. 762, p. 1-15, 2021.

VÉLEZ-UPEGUI, J. J.; LONDOÑO-CARVAJAL, A.; SOLÓRZANO-GÓMEZ, M. V.; IGNACIO, J.; BOTERO-FERNÁNDEZ, V.; VÉLEZ, J. J.; LONDOÑO, A. Redesign of a water quantity , quality , and sediment-monitoring network in a tropical. **Revista Facultad de Ingeniería, Universidad de Antioquia**, n. 96, p. 64–77, 2020.

VICQ, R. DE; MATSCHULLAT, J.; LEITE, M. G. P.; NALINI JUNIOR, H. A.; MENDONÇA, F. P. C. Iron Quadrangle stream sediments, Brazil: geochemical maps and reference values. **Environmental Earth Sciences**, v. 74, n. 5, p. 4407–4417, 2015.

VILLAS-BOAS, M. D.; OLIVERA, F.; AZEVEDO, J. P. S. DE. Assessment of the water quality monitoring network of the Piabanha River experimental watersheds in Rio de Janeiro, Brazil, using autoassociative neural networks. **Environmental Monitoring and Assessment**, p. 189–439, 2017.

VILMIN, L.; FLIPO, N.; ESCOFFIER, N.; GROLEAU, A. Estimation of the water quality of a large urbanized river as defined by the European WFD: what is the optimal sampling frequency? **Environmental Science and Pollution Research**, v. 25, p. 23485–23501, 2018.

WANG, Z.; RANGAIAH, G. P. Application and analysis of methods for selecting an optimal solution from the Pareto-Optimal front obtained by multiobjective optimization. **Industrial & Engineering Chemistry Research**, v. 56, p. 560–574, 2017.

WARD, R. C.; LOFTIS, J. C.; MCBRIDE, G. B. The “data-rich but information-poor” syndrome in water quality monitoring. **Environmental Management**, v. 10, n. 3, p. 291–297, 1986.

WEI, T.; SONG, S. Confidence interval estimation for precipitation quantiles based on principle of maximum entropy. **Entropy**, v. 21, n. 3, p. 1-18, 2019.

WILD, M.; ROUHANI, S.; OEHRIG, J.; ALVES, P. H. G.; ODLE, W.; GASPAR, D. F. A. Using spatiotemporal ratio analyses to quantitatively estimate water quality recovery of the Rio Doce. **Integrated Environmental Assessment and Management**, v. 00, n. 00, p. 1–13, 2023.

YAMAKA, W.; SRIBOONCHITTA, S. Forecasting using information and entropy based on belief functions. **Complexity**, v. 2020, p. 1-17, 2020.

YAMAMOTO, F. Y.; PAULY, G. F. E.; NASCIMENTO, L. S.; FERNANDES, G. M.; SANTOS, M. P.; FIGUEIRA, R. C. L.; CAVALCANTE, R. M.; GRASSI, M. T.; ABESSA, D. M. S. Explaining the persistence of hazardous chemicals in the Doce River (Brazil) by multiple sources of contamination and a major environmental disaster. **Journal of Hazardous Materials Advances**, v. 9, n. November 2022, p. 1-15, 2023.

YAMAMOTO, F. Y.; SOUZA, A. T. C.; PAULA, V. DE C. S. DE; BEVERARI, I.; GARCIA, J. R. E.; PADIAL, A. A.; SOUZA ABESSA, D. M. DE. From molecular endpoints to modeling longer-term effects in fish embryos exposed to the elutriate from Doce River. **Science of the Total Environment**, v. 846, n. March, p. 1-14, 2022.

YANG, W.; ZHAO, Y.; WANG, D.; WU, H.; LIN, A.; HE, L. Using principal components analysis and idw interpolation to determine spatial and temporal changes of Surfacewater quality of Xin'Anjiang river in huangshan, china. **International Journal of Environmental Research and Public Health**, v. 17, n. 8, p. 1–14, 2020.

YAZDI, J. Water quality monitoring network design for urban drainage systems, an entropy method. **Urban Water Journal**, v. 15, n. 3, p. 227–233, 2018.

YENILMEZ, F.; DÜZGÜN, S.; AKSOY, A. An evaluation of potential sampling locations in a reservoir with emphasis on conserved spatial correlation structure. **Environmental Monitoring and Assessment**, v. 187, n. 1, p. 1-21, 2015.

YU, X.; ZHANG, X. Multiswarm comprehensive learning particle swarm optimization for solving multiobjective optimization problems. **Plos one**, v. 12, n. 2, p. 1–21, 2017.

YUAN, Y.; XU, H.; WANG, B.; ZHANG, B.; YAO, X. Balancing convergence and diversity in decomposition-based many-objective optimizers. **IEEE Transactions on Evolutionary Computation**, v. 20, n. 2, p. 180–198, 2016.

ZHANG, F.; WANG, J.; WANG, X. Recognizing the relationship between spatial patterns in water quality and land-use/cover types: A case study of the Jinghe Oasis in Xinjiang, China. **Water (Switzerland)**, v. 10, n. 5, p. 1–17, 2018.

ZHU, X.; YUE, Y.; WONG, P. W. H.; ZHANG, Y.; DING, H. Designing an optimized water quality monitoring network with reserved monitoring locations. **Water (Switzerland)**, v. 11, n. 4, p. 1–25, 2019.

ZHU, X.; YUE, Y.; WONG, P. W. H.; ZHANG, Y.; TAN, J. Optimum water quality monitoring network design for bidirectional river systems. **International Journal of Environmental Research and Public Health**, v. 15, n. 2, p. 1–19, 2018.

APÊNDICES

Tabela A1 - Estações de monitoramento selecionadas

Estação	Código	Latitude	Longitude
Acaiaca - Carmo 01	RCA-01	-20.347100	-43.112700
Aimorés - Manhauçu 01	RMH-01	-19.483700	-41.094100
Baixo Guandu - Doce 11	RDO-11	-19.506100	-41.013900
Baixo Guandu - Guandu 01	RGU-01	-19.624700	-41.017900
Barra Longa - Carmo 02	RCA-02	-20.283000	-43.032500
Barra Longa - Carmo 03	RCA-03	-20.281700	-43.025500
Barra Longa - Carmo 04	RCA-04	-20.271800	-42.998700
Barra Longa - Carmo 05	RCA-05	-20.258800	-42.990200
Barra Longa - Carmo 06	RCA-06	-20.268000	-42.971500
Barra Longa - Gualaxo Norte 08	RGN-08	-20.286100	-43.065800
Belo Oriente - Doce 05	RDO-05	-19.320700	-42.364600
Colatina - Doce 12	RDO-12	-19.499300	-40.758700
Colatina - Doce 13	RDO-13	-19.535300	-40.635600
Colatina - Doce 14	RDO-14	-19.510600	-40.554900
Conselheiro Pena - Caratinga 01	RCR-01	-19.063600	-41.531100
Governador Valadares - Doce 07	RDO-07	-18.970800	-42.088300
Governador Valadares - Doce 08	RDO-08	-18.883300	-41.952600
Governador Valadares - Suaçuí 01	RSG-01	-18.854100	-41.786400
Linhares - Doce 15	RDO-15	-19.407800	-40.064600
Linhares - Doce 16	RDO-16	-19.646400	-39.823200
Mariana - Gualaxo do Norte 01	RGN-01	-20.276500	-43.431700
Mariana - Gualaxo do Norte 02	RGN-02	-20.266900	-43.307100
Mariana - Gualaxo do Norte 03	RGN-03	-20.269100	-43.300800
Mariana - Gualaxo do Norte 04	RGN-04	-20.283400	-43.290400
Mariana - Gualaxo do Norte 05	RGN-05	-20.292700	-43.279200
Mariana - Gualaxo do Norte 06	RGN-06	-20.303700	-43.249500
Mariana - Gualaxo do Norte 07	RGN-07	-20.280900	-43.079200
Mariana - Piracicaba 01	RPC-01	-20.178300	-43.501500
Mariana - Piracicaba 02	RPC-02	-20.159300	-43.419200
Naque - Santo Antônio 01	RSA-01	-19.232400	-42.327100
Periquito - Doce 06	RDO-06	-19.095600	-42.155000
Ponte Nova - Piranga 01	RPG-01	-20.383600	-42.902600
Raul Soares - Matipó 01	RMA-01	-20.099800	-42.449100
Bom Jesus do Galho - Doce 04	RDO-04	-19.554200	-42.521400
Resplendor - Doce 10	RDO-10	-19.328800	-41.252800
Rio Doce - Doce 01	RDO-01	-20.248093	-42.885247
São Domingos do Prata - Doce 03	RDO-03	-20.014400	-42.744600
Timoteo - Piracicaba 03	RPC-03	-19.525000	-42.644200
Tumiritinga - Doce 09	RDO-09	-18.971500	-41.641600

Tabela A2 – Conjunto original de variáveis de qualidade da água – rede PMQQS

Variável de qualidade da água	Unidade
Alcalinidade total	mgCaCO ₃ /L
Alumínio dissolvido	mg/L
Alumínio total	mg/L
Antimônio dissolvido	mg/L
Antimônio total	mg/L
Arsênio dissolvido	mg/L
Arsênio total	mg/L
Bário dissolvido	mg/L
Bário total	mg/L
Berílio dissolvido	mg/L
Berílio total	mg/L
Boro dissolvido	mg/L
Boro total	mg/L
Cádmio dissolvido	mg/L
Cádmio total	mg/L
Cálcio dissolvido	mg/L
Cálcio total	mg/L
Carbono orgânico dissolvido	mg/L
Carbono orgânico total	mg/L
Chumbo dissolvido	mg/L
Chumbo total	mg/L
Cianeto livre	mg/L
Cloreto total	mg/L
Clorofila a	µg/L
Cobalto dissolvido	mg/L
Cobalto total	mg/L
Cobre dissolvido	mg/L
Cobre total	mg/L
Condutividade <i>in situ</i>	µs/cm
Condutividade lab	µs/cm
Cor verdadeira	mgPt/L
Cromo dissolvido	mg/L
Cromo total	mg/L
DBO	mgO ₂ /L
Dureza total	mgCaCO ₃ /L
Escherichia coli	NMP/100mL
Feoftina	µg/L
Ferro II	mg/L
Ferro III	mg/L
Ferro dissolvido	mg/L
Ferro total	mg/L
Fluoreto	mg/L
Fosfato	mg/L
Fósforo dissolvido	mg/L
Fósforo total	mg/L

Variável de qualidade da água	Unidade
Magnésio dissolvido	mg/L
Magnésio total	mg/L
Manganês dissolvido	mg/L
Manganês total	mg/L
Mercúrio dissolvido	mg/L
Mercúrio total	mg/L
Metilmercúrio	mg/L
Molibdênio dissolvido	mg/L
Molibdênio total	mg/L
Níquel dissolvido	mg/L
Níquel total	mg/L
Nitrato	mg/L
Nitrito	mg/L
Nitrogênio amoniacal	mg/L
Nitrogênio kjeldahl total	mg/L
Nitrogênio orgânico	mg/L
Oxigênio dissolvido <i>in situ</i>	mgO ₂ /L.
Oxigênio dissolvido saturado <i>in situ</i>	%
Polifosfato	mg/l
Potássio dissolvido	mg/l
Potencial redox <i>in situ</i>	Mv
Prata dissolvida	mg/L
Prata total	mg/L
Selênio dissolvido	mg/L
Selênio total	mg/L
Sódio dissolvido	mg/L
Sódio total	mg/L
Sólidos dissolvidos totais	mg/L
Sólidos sedimentáveis	mg/L
Sólidos suspensos totais	mg/L
Sólidos totais	mg/L
Sulfato	mg/L
Sulfetos como H ₂ S indissociável	mg/L
Sulfetos totais	mg/L
Temperatura ambiente <i>in situ</i>	°C
Temperatura da amostra <i>in situ</i>	°C
Transparência da água <i>in situ</i>	m
Turbidez <i>in situ</i>	NTU
Turbidez lab	NTU
Vanádio dissolvido	mg/L
Vanádio total	mg/L
Zinco dissolvido	mg/L
Zinco total	mg/L
pH <i>in situ</i>	-
pH lab	-

Tabela A3 – Variáveis de qualidade da água selecionada para pesquisa

Variáveis de qualidade da água
Alcalinidade total
Alumínio dissolvido
Alumínio total
Antimônio dissolvido
Antimônio total
Arsênio dissolvido
Arsênio total
Bário dissolvido
Bário total
Berílio dissolvido
Berílio total
Boro dissolvido
Boro total
Cádmio dissolvido
Cádmio total
Cálcio total
Carbono orgânico dissolvido
Carbono orgânico total
Chumbo dissolvido
Chumbo total
Cianeto livre
Cloreto total
Clorofila a
Cobalto dissolvido
Cobalto total
Cobre dissolvido
Cobre total
Condutividade <i>in situ</i>
Cor verdadeira
Cromo dissolvido
Cromo total
DBO
Dureza total
Escherichia coli
Feoftina
Ferro II
Ferro III
Ferro dissolvido
Ferro total
Fósforo dissolvido
Fósforo total
Magnésio total
Manganês dissolvido
Manganês total
Mercúrio dissolvido

Variáveis de qualidade da água

Mercúrio total
 Molibdênio dissolvido
 Molibdênio total
 Níquel dissolvido
 Níquel total
 Nitrato
 Nitrito
 Nitrogênio amoniacal
 Nitrogênio kjeldahl total
 Nitrogênio orgânico
 Oxigênio dissolvido *in situ*
 Oxigênio dissolvido saturado *in situ*
 Polifosfato
 Potencial redox *in situ*
 Prata dissolvida
 Prata total
 Selênio dissolvido
 Selênio total
 Sódio total
 Sólidos dissolvidos totais
 Sólidos sedimentáveis
 Sólidos suspensos totais
 Sólidos totais
 Sulfetos como H₂S indissociável
 Sulfetos totais
 Temperatura da amostra *in situ*
 Turbidez *in situ*
 Vanádio dissolvido
 Vanádio total
 Zinco dissolvido
 Zinco total

Tabela A4 – Variáveis de qualidade da água eliminadas

Variável de qualidade da água	%	Não-conformidade
Condutividade lab	35,8	Excesso de dados inválidos
Cálcio dissolvido	31,8	Excesso de dados faltantes
Fluoreto	31,8	Excesso de dados faltantes
Fosfato	31,8	Excesso de dados faltantes
Magnésio dissolvido	31,8	Excesso de dados faltantes
Metilmercúrio	75,5	Excesso de dados faltantes
Potássio dissolvido	31,8	Excesso de dados faltantes
Sódio dissolvido	31,8	Excesso de dados faltantes
Sulfato	31,8	Excesso de dados faltantes
Transparência da água <i>in situ</i>	97,9	Excesso de dados faltantes
Turbidez lab	31,8	Excesso de dados faltantes

Table A5 – Estações de monitoramento da rede PMQQS agrupadas em subregiões de qualidade da água

Estação de monitoramento	Código	Sub-região
Acaiaca - Carmo 01	RCA-01	E
Barra Longa - Carmo 02	RCA-02	E
Barra Longa - Carmo 03	RCA-03	E
Barra Longa - Carmo 04	RCA-04	E
Barra Longa - Carmo 05	RCA-05	E
Barra Longa - Carmo 06	RCA-06	E
Conselheiro Pena - Caratinga 01	RCR-01	L
Rio Doce - Doce 01	RDO-01	G
São Domingos do Prata - Doce 03	RDO-03	G
Bom Jesus do Galho - Doce 04	RDO-04	G
Belo Oriente - Doce 05	RDO-05	I
Periquito - Doce 06	RDO-06	G
Governador Valadares - Doce 07	RDO-07	J
Governador Valadares - Doce 08	RDO-08	J
Tumiritinga - Doce 09	RDO-09	J
Resplendor - Doce 10	RDO-10	J
Baixo Guandu - Doce 11	RDO-11	M
Colatina - Doce 12	RDO-12	M
Colatina - Doce 13	RDO-13	M
Colatina - Doce 14	RDO-14	M
Linhares - Doce 15	RDO-15	M
Linhares - Doce 16	RDO-16	N
Mariana - Gualaxo Norte 01	RGN-01	A
Mariana - Gualaxo Norte 02	RGN-02	B
Mariana - Gualaxo Norte 03	RGN-03	B
Mariana - Gualaxo Norte 04	RGN-04	B
Mariana - Gualaxo Norte 05	RGN-05	C
Mariana - Gualaxo Norte 06	RGN-06	C
Mariana - Gualaxo Norte 07	RGN-07	D
Barra Longa - Gualaxo Norte 08	RGN-08	D
Baixo Guandu - Guandu 01	RGU-01	K
Raul Soares - Matipó 01	RMA-01	G
Aimorés - Manhuaçu 01	RMH-01	M
Mariana - Piracicaba 01	RPC-01	F
Mariana - Piracicaba 02	RPC-02	F
Timoteo - Piracicaba 03	RPC-03	H
Ponte Nova - Piranga 01	RPG-01	G
Naque - Santo Antônio 01	RSA-01	F
Governador Valadares - Suaçuí 01	RSG-01	J

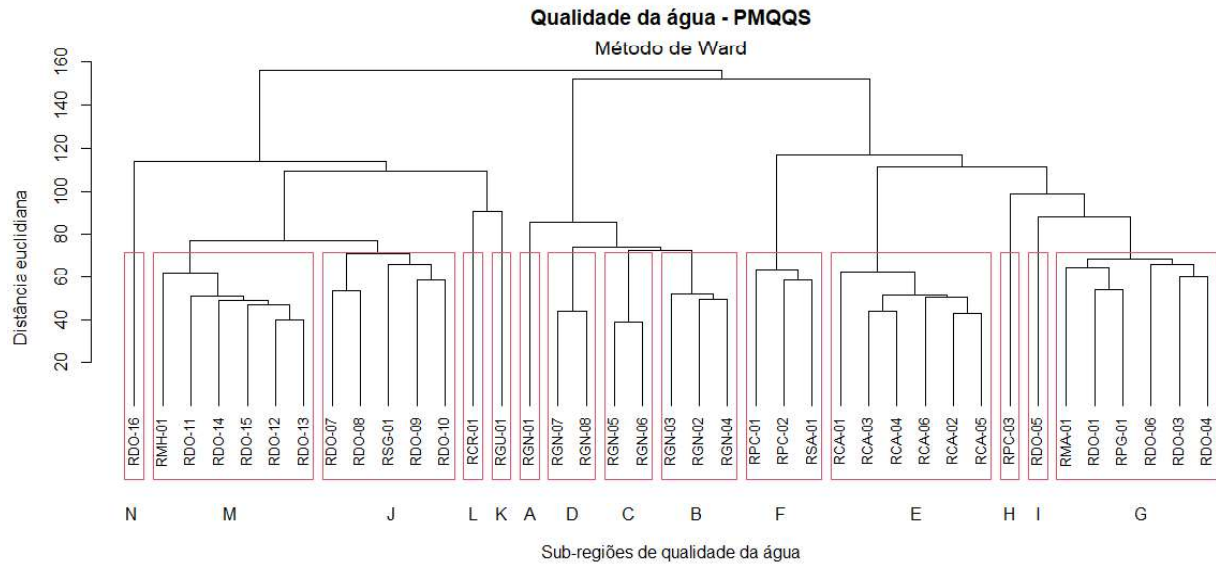
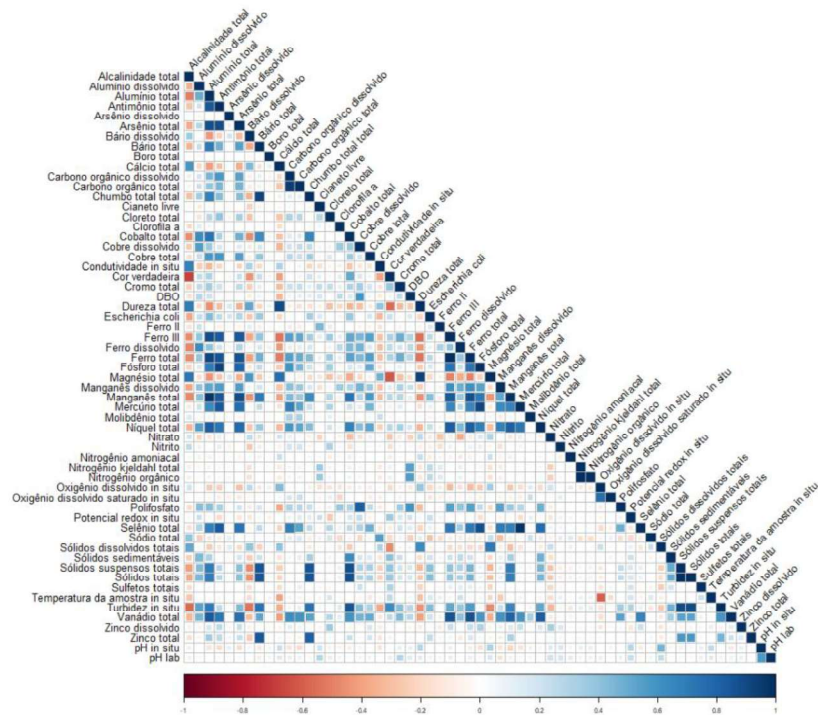
Figura A1 - Dendrograma com delimitação das subregiões de qualidade da água de A a N**Figura A2 – Matriz de correlação de variáveis de qualidade da água na subregião B**

Figura A3 – Matriz de correlação de variáveis de qualidade da água na subregião C

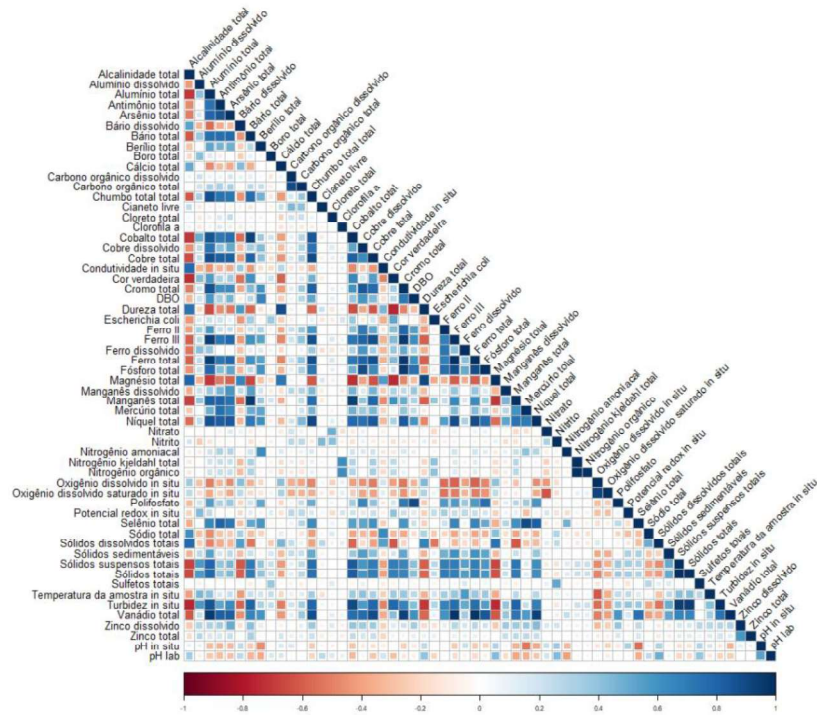


Figura A4 – Matriz de correlação de variáveis de qualidade da água na subregião D

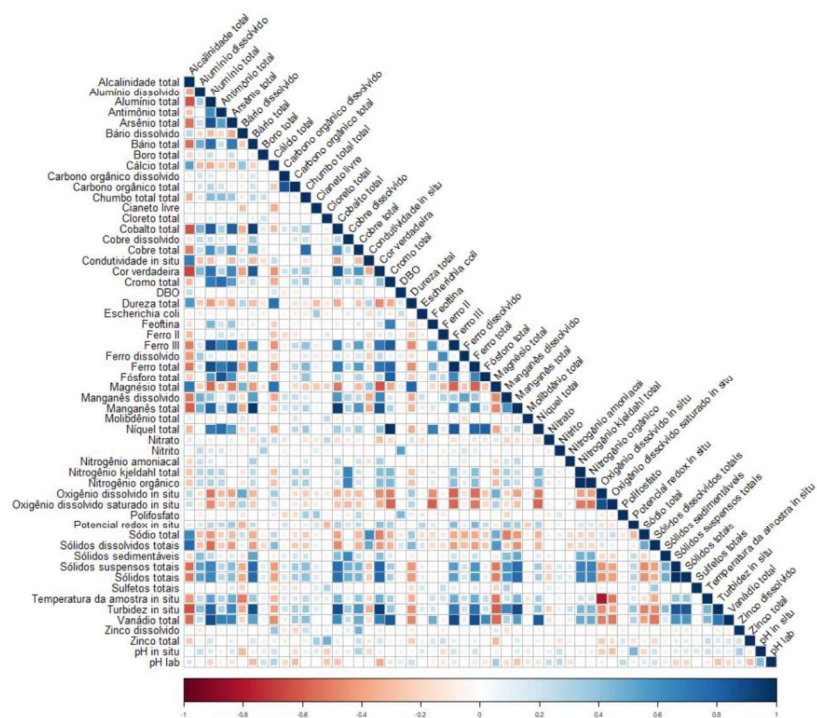


Figura A5 – Matriz de correlação de variáveis de qualidade da água na subregião E

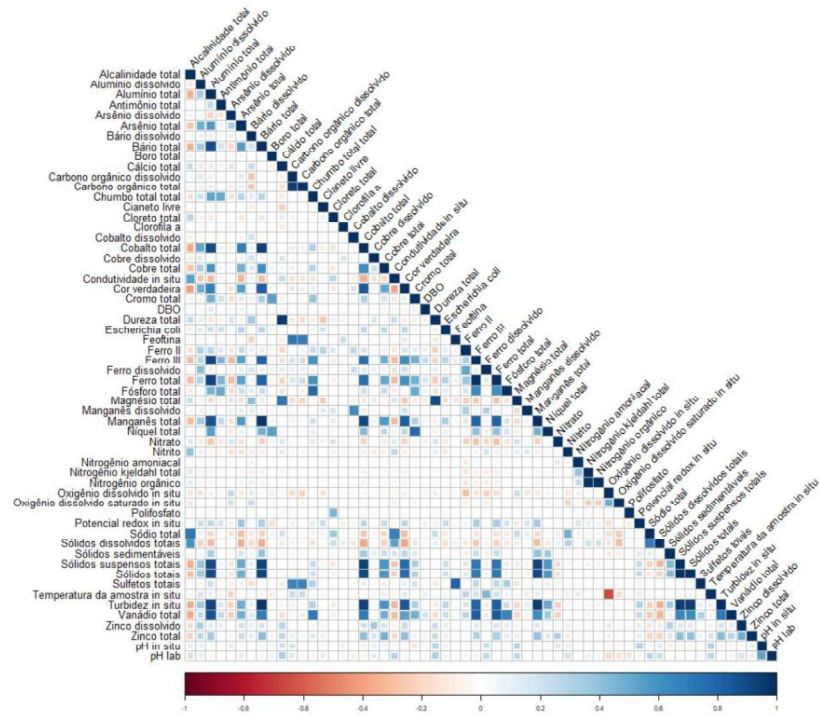


Figura A6 – Matriz de correlação de variáveis de qualidade da água na subregião F

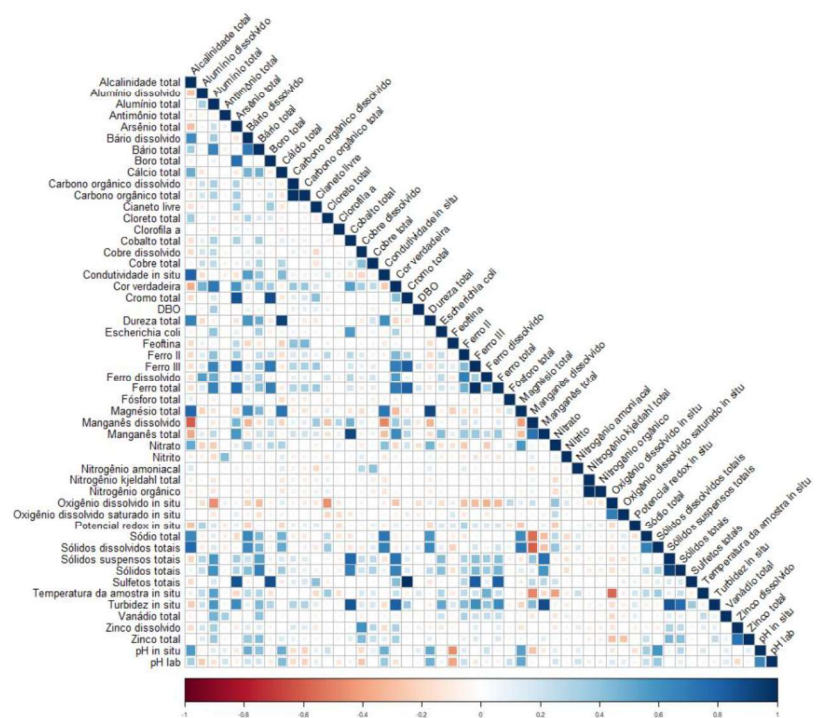


Figura A7 – Matriz de correlação de variáveis de qualidade da água na subregião G

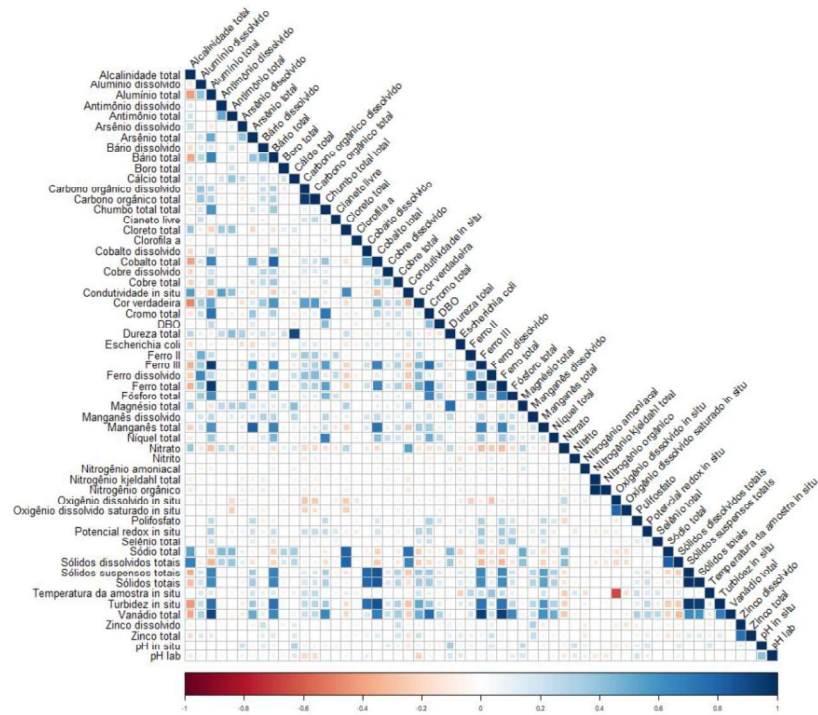


Figura A8 – Matriz de correlação de variáveis de qualidade da água na subregião H

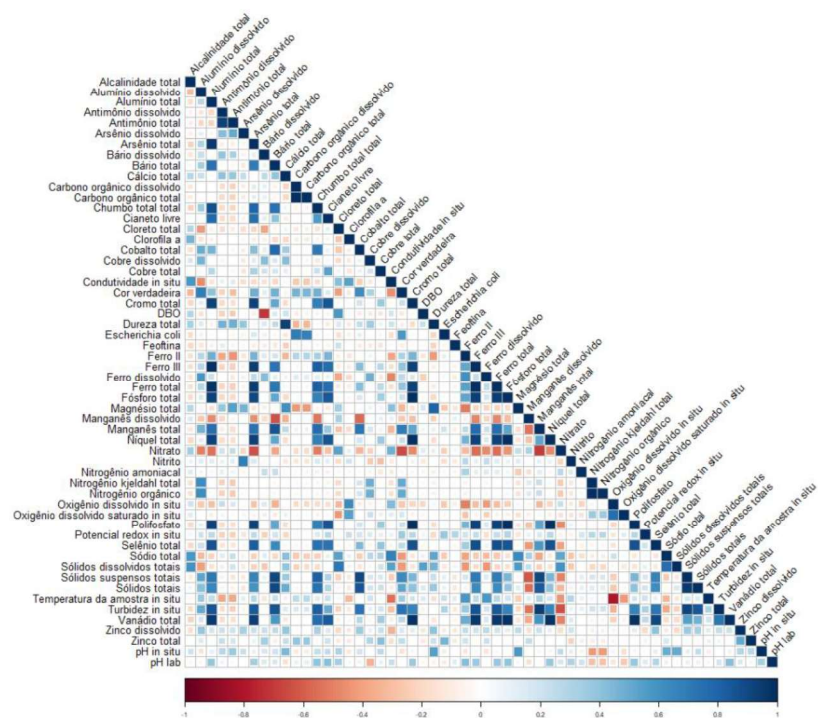


Figura A9 – Matriz de correlação de variáveis de qualidade da água na subregião I

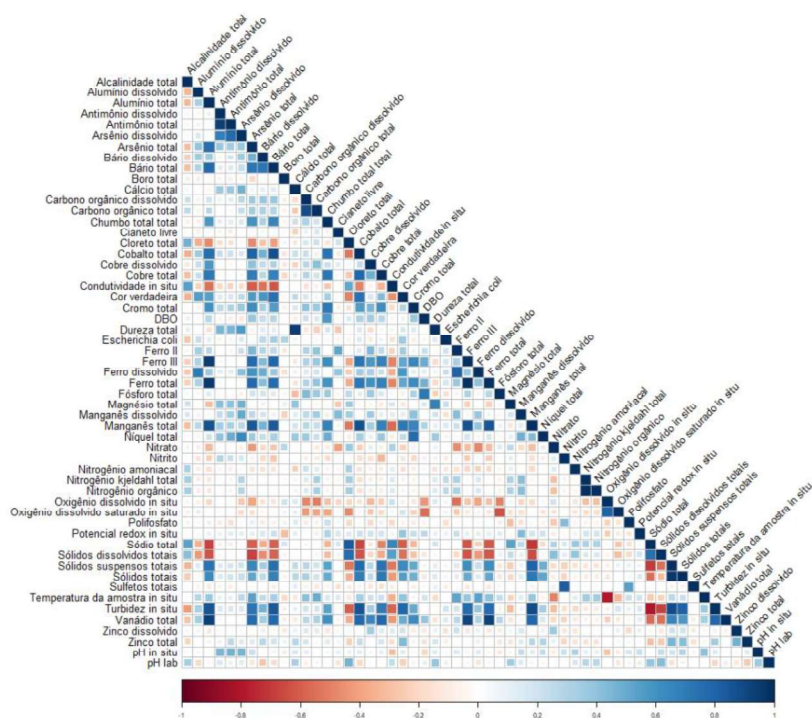


Figura A10 – Matriz de correlação de variáveis de qualidade da água na subregião J

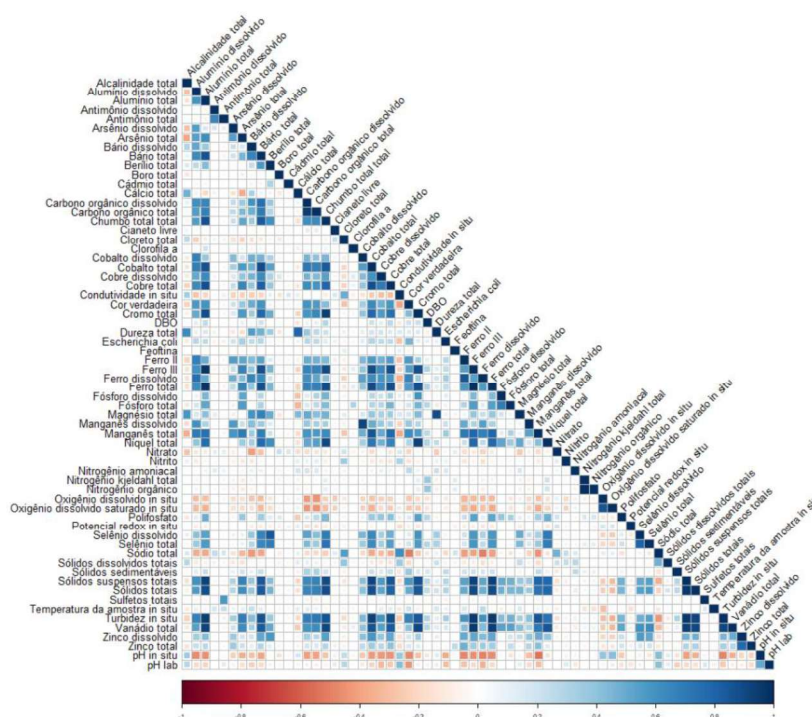


Figura A11 – Matriz de correlação de variáveis de qualidade da água na subregião K

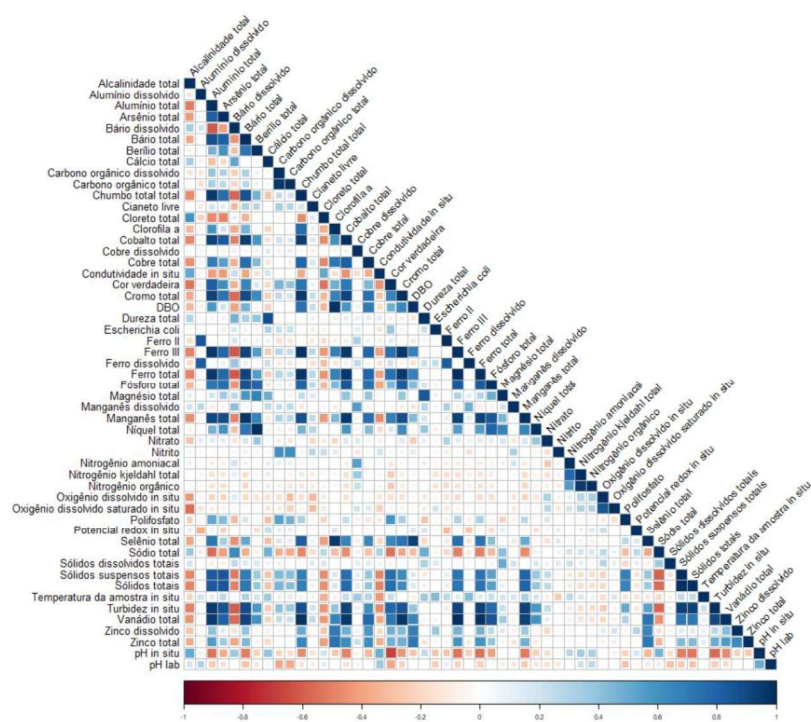


Figura A12 – Matriz de correlação de variáveis de qualidade da água na subregião L

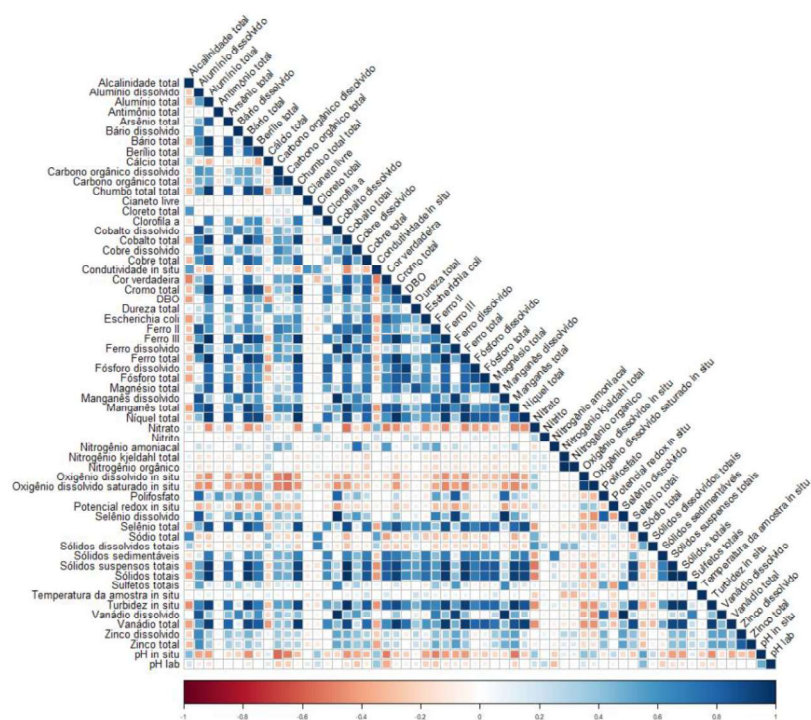


Figura A13 – Matriz de correlação de variáveis de qualidade da água na subregião M

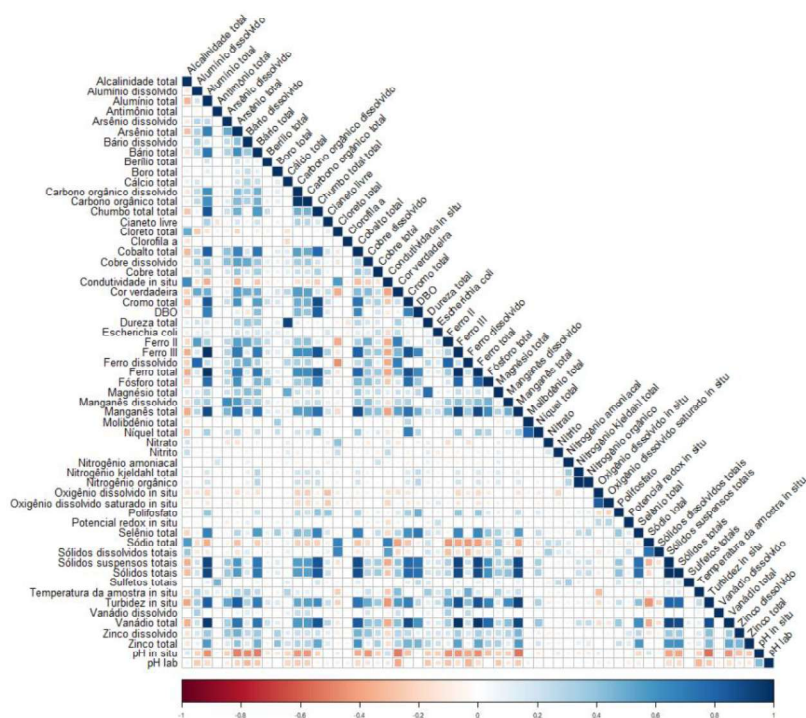


Figura A14 – Matriz de correlação de variáveis de qualidade da água na subregião N

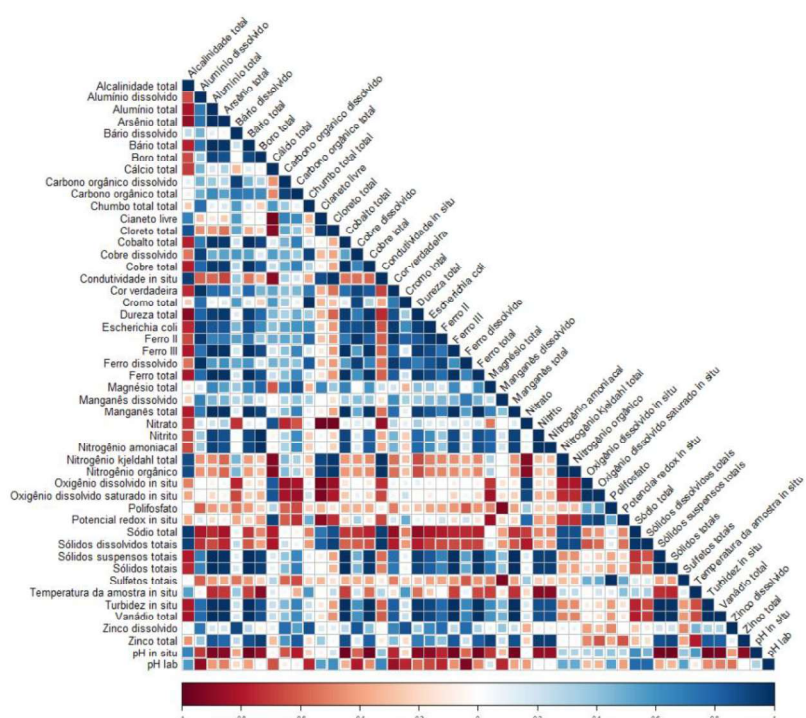


Tabela A6 - Variáveis de qualidade da água excluídas da ACP devido à multicolinearidade

A	B	C	D	E	F	G
Arsênio total	Antimônio total	Arsênio total	Arsênio total	Bário total	Carbono orgânico total	Carbono orgânico total
Carbono orgânico total	Arsênio total	Carbono orgânico total	Carbono orgânico total	Carbono orgânico total	Cromo total	Dureza total
Cobalto total	Carbono orgânico total	Chumbo total	Cobalto total	Cobalto total	Dureza total	Ferro III
Dureza total	Chumbo total	Cobalto total	Cor verdadeira	Dureza total	Ferro total	Ferro total
Ferro II	Dureza total	Cobre total	Ferro III	Ferro III	Manganês total	Manganês total
Ferro III	Ferro III	Cromo total	Ferro total	Ferro total	Nitrogênio orgânico	Nitrogênio orgânico
Ferro total	Ferro total	Dureza total	Fósforo total	Manganês total	Sólidos totais	Sódio total
Magnésio total	Manganês total	Ferro III	Manganês total	Nitrogênio orgânico	Sulfetos totais	Sólidos suspensos totais
Manganês total	Mercurio total	Ferro total	Níquel total	Sólidos suspensos totais	Turbidez <i>in situ</i>	Sólidos totais
Níquel total	Níquel total	Fósforo total	Nitrogênio orgânico	Sólidos totais	-	Turbidez <i>in situ</i>
Nitrogênio orgânico	Nitrogênio orgânico	Manganês total	Oxigênio dissolvido saturado <i>in situ</i>	Sulfetos totais	-	Vanádio total
Sólidos suspensos totais	Polifosfato	Níquel total	Sólidos totais	Turbidez <i>in situ</i>	-	-
Turbidez <i>in situ</i>	Selênio total	Nitrogênio orgânico	Temperatura da amostra <i>in situ</i>	Vanádio total	-	-
-	Sólidos suspensos totais	Oxigênio dissolvido saturado <i>in situ</i>	Turbidez <i>in situ</i>	-	-	-
-	Sólidos totais	Polifosfato	Vanádio total	-	-	-
-	Turbidez <i>in situ</i>	Selênio total	-	-	-	-
-	Vanádio total	Sólidos totais	-	-	-	-
-	Zinco total	Turbidez <i>in situ</i>	-	-	-	-
	-	Vanádio total	-	-	-	-
H	I	J	K	L	M	N
Antimônio total	Antimônio total	Bário total	Arsênio total	Arsênio total	Carbono orgânico total	Alumínio total
Arsênio total	Carbono orgânico total	Carbono orgânico total	Bário total	Bário total	Chumbo total	Arsênio total

A	B	C	D	E	F	G
Carbono orgânico total	Cobalto total	Chumbo total	Carbono orgânico total	Berílio total	Cobalto total	Boro total
Chumbo total	Dureza total	Cobalto total	Chumbo total	Carbono orgânico total	Cromo total	Carbono orgânico dissolvido
Cromo total	Ferro III	Cobre total	Cobalto total	Chumbo total	Dureza total	Cianeto livre
Dureza total	Ferro total	Cromo total	Cor verdadeira	Cobalto dissolvido	Ferro III	Cloreto total
Ferro III	Manganês total	Ferro III	Cromo total	Cobalto total	Ferro total	Cobalto total
Ferro total	Nitrogênio orgânico	Ferro dissolvido	DBO	Cobre total	Manganês total	Cobre dissolvido
Fósforo total	Oxigênio dissolvido saturado <i>in situ</i>	Ferro total	Dureza total	Cromo total	Nitrogênio orgânico	Cobre total
Manganês total	Sódio total	Magnésio total	Ferro II	Escherichia coli	Oxigênio dissolvido saturado <i>in situ</i>	Condutividade <i>in situ</i>
Níquel total	Sólidos totais	Manganês dissolvido	Ferro III	Ferro II	Sólidos suspensos totais	Cor verdadeira
Nitrogênio orgânico	Sulfetos totais	Nitrogênio orgânico	Ferro dissolvido	Ferro III	Sólidos totais	Cromo total
Polifosfato	Turbidez <i>in situ</i>	Oxigênio dissolvido saturado <i>in situ</i>	Ferro total	Ferro dissolvido	Turbidez <i>in situ</i>	Dureza total
Selênio total	Vanádio total	Selênio dissolvido	Manganês total	Ferro total	Vanádio total	Escherichia coli
Sólidos suspensos totais	-	Sólidos suspensos totais	Níquel total	Fósforo dissolvido	-	Ferro II
Sólidos totais	-	Sólidos totais	Nitrogênio orgânico	Fósforo total	-	Ferro III
Temperatura da amostra <i>in situ</i>	-	Turbidez <i>in situ</i>	Oxigênio dissolvido saturado <i>in situ</i>	Magnésio total	-	Ferro dissolvido
Turbidez <i>in situ</i>	-	Vanádio total	Selênio total	Manganês dissolvido	-	Ferro total
Vanádio total	-	-	Sólidos suspensos totais	Manganês total	-	Magnésio total
-	-	-	Sólidos totais	Níquel total	-	Manganês total
-	-	-	Turbidez <i>in situ</i>	Nitrogênio orgânico	-	Nitrato
-	-	-	Vanádio total	Oxigênio dissolvido saturado <i>in situ</i>	-	Nitrito
-	-	-	-	Selênio dissolvido	-	Nitrogênio amoniacal
-	-	-	-	Selênio total	-	Nitrogênio kjeldahl total

A	B	C	D	E	F	G
-	-	-	-	Sólidos suspensos totais	-	Nitrogênio orgânico
-	-	-	-	Sólidos totais	-	Oxigênio dissolvido <i>in situ</i>
-	-	-	-	Turbidez <i>in situ</i>	-	Oxigênio dissolvido saturado <i>in situ</i>
-	-	-	-	Vanádio dissolvido	-	Polifosfato
-	-	-	-	Vanádio total	-	Potencial redox <i>in situ</i>
-	-	-	-	Zinco total	-	Sódio total
-	-	-	-	-	-	Sólidos dissolvidos totais
-	-	-	-	-	-	Sólidos suspensos totais
-	-	-	-	-	-	Sólidos totais
-	-	-	-	-	-	Sulfetos totais
-	-	-	-	-	-	Turbidez <i>in situ</i>
-	-	-	-	-	-	Vanádio total
-	-	-	-	-	-	Zinco dissolvido
-	-	-	-	-	-	Zinco total
-	-	-	-	-	-	pH <i>in situ</i>
-	-	-	-	-	-	pH lab

Tabela A7 - Variáveis de qualidade da água selecionadas para ACP

A	B	C	D	E	F	G
Alcalinidade total	Alcalinidade total	Alcalinidade total	Alcalinidade total	Alcalinidade total	Alcalinidade total	Alcalinidade total
Alumínio dissolvido	Alumínio dissolvido	Alumínio dissolvido	Alumínio dissolvido	Alumínio dissolvido	Alumínio dissolvido	Alumínio dissolvido
Alumínio total	Alumínio total	Alumínio total	Alumínio total	Alumínio total	Alumínio total	Alumínio total
Antimônio total	Arsênio dissolvido	Antimônio total	Antimônio total	Antimônio total	Antimônio total	Antimônio dissolvido
Bário dissolvido	Bário dissolvido	Bário dissolvido	Bário dissolvido	Arsênio dissolvido	Arsênio total	Antimônio total
Bário total	Bário total	Bário total	Bário total	Arsênio total	Bário dissolvido	Arsênio dissolvido
Boro total	Boro total	Berílio total	Boro total	Bário dissolvido	Bário total	Arsênio total
Cálcio total	Cálcio total	Boro total	Cálcio total	Boro total	Boro total	Bário dissolvido
Carbono orgânico dissolvido	Carbono orgânico dissolvido	Cálcio total	Carbono orgânico dissolvido	Cálcio total	Cálcio total	Bário total
Cianeto livre	Cianeto livre	Carbono orgânico dissolvido	Chumbo total	Carbono orgânico dissolvido	Carbono orgânico dissolvido	Boro total

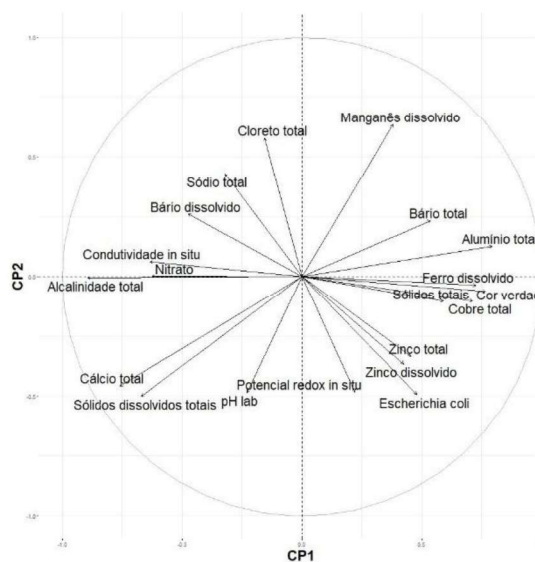
A	B	C	D	E	F	G
Cloreto total	Cloreto total	Cianeto livre	Cianeto livre	Chumbo total	Cianeto livre	Cálcio total
Clorofila a	Clorofila a	Cloreto total	Cloreto total	Cianeto livre	Cloreto total	Carbono orgânico dissolvido
Cobre dissolvido	Cobalto total	Clorofila a	Cobre dissolvido	Cloreto total	Clorofila a	Chumbo total
Cobre total	Cobre dissolvido	Cobre dissolvido	Cobre total	Clorofila a	Cobalto total	Cianeto livre
Condutividade <i>in situ</i>	Cobre total	Condutividade <i>in situ</i>	Condutividade <i>in situ</i>	Cobalto dissolvido	Cobre dissolvido	Cloreto total
Cor verdadeira	Condutividade <i>in situ</i>	Cor verdadeira	Cromo total	Cobre dissolvido	Cobre total	Clorofila a
Cromo total	Cor verdadeira	DBO	DBO	Cobre total	Condutividade <i>in situ</i>	Cobalto dissolvido
Escherichia coli	Cromo total	Escherichia coli	Dureza total	Condutividade <i>in situ</i>	Cor verdadeira	Cobalto total
Ferro dissolvido	DBO	Ferro II	Escherichia coli	Cor verdadeira	DBO	Cobre dissolvido
Manganês dissolvido	Escherichia coli	Ferro dissolvido	Feoftina	Cromo total	Escherichia coli	Cobre total
Molibdênio total	Ferro II	Magnésio total	Ferro II	DBO	Feoftina	Condutividade <i>in situ</i>
Nitrato	Ferro dissolvido	Manganês dissolvido	Ferro dissolvido	Escherichia coli	Ferro II	Cor verdadeira
Nitrito	Fósforo total	Mercúrio total	Magnésio total	Feoftina	Ferro III	Cromo total
Nitrogênio amoniacal	Magnésio total	Nitrato	Manganês dissolvido	Ferro II	Ferro dissolvido	DBO
Nitrogênio kjeldahl total	Manganês dissolvido	Nitrito	Molibdênio total	Ferro dissolvido	Fósforo total	Escherichia coli
Oxigênio dissolvido <i>in situ</i>	Molibdênio total	Nitrogênio amoniacal	Nitrato	Fósforo total	Magnésio total	Ferro II
Oxigênio dissolvido saturado <i>in situ</i>	Nitrato	Nitrogênio kjeldahl total	Nitrito	Magnésio total	Manganês dissolvido	Ferro dissolvido
Potencial redox <i>in situ</i>	Nitrito	Oxigênio dissolvido <i>in situ</i>	Nitrogênio amoniacal	Manganês dissolvido	Nitrato	Fósforo total
Sódio total	Nitrogênio amoniacal	Potencial redox <i>in situ</i>	Nitrogênio kjeldahl total	Níquel total	Nitrito	Magnésio total
Sólidos dissolvidos totais	Nitrogênio kjeldahl total	Sódio total	Oxigênio dissolvido <i>in situ</i>	Nitrato	Nitrogênio amoniacal	Manganês dissolvido
Sólidos totais	Oxigênio dissolvido <i>in situ</i>	Sólidos dissolvidos totais	Polifosfato	Nitrito	Nitrogênio kjeldahl total	Níquel total
Temperatura da amostra <i>in situ</i>	Oxigênio dissolvido saturado <i>in situ</i>	Sólidos sedimentáveis	Potencial redox <i>in situ</i>	Nitrogênio amoniacal	Oxigênio dissolvido <i>in situ</i>	Nitrato
Vanádio total	Potencial redox <i>in situ</i>	Sólidos suspensos totais	Sódio total	Nitrogênio kjeldahl total	Oxigênio dissolvido saturado <i>in situ</i>	Nitrito
Zinco dissolvido	Sódio total	Sulfetos totais	Sólidos dissolvidos totais	Oxigênio dissolvido <i>in situ</i>	Potencial redox <i>in situ</i>	Nitrogênio amoniacal
Zinco total	Sólidos dissolvidos totais	Temperatura da amostra <i>in situ</i>	Sólidos sedimentáveis	Oxigênio dissolvido saturado <i>in situ</i>	Sódio total	Nitrogênio kjeldahl total

A	B	C	D	E	F	G
pH <i>in situ</i>	Sólidos sedimentáveis	Zinco dissolvido	Sólidos suspensos totais	Polifosfato	Sólidos dissolvidos totais	Oxigênio dissolvido <i>in situ</i>
pH lab	Sulfetos totais	Zinco total	Sulfetos totais	Potencial redox <i>in situ</i>	Sólidos suspensos totais	Oxigênio dissolvido saturado <i>in situ</i>
-	Temperatura da amostra <i>in situ</i>	pH <i>in situ</i>	Zinco dissolvido	Sódio total	Temperatura da amostra <i>in situ</i>	Polifosfato
-	Zinco dissolvido	pH lab	Zinco total	Sólidos dissolvidos totais	Vanádio total	Potencial redox <i>in situ</i>
-	pH <i>in situ</i>	-	pH <i>in situ</i>	Sólidos sedimentáveis	Zinco dissolvido	Selênio total
-	pH lab	-	pH lab	Temperatura da amostra <i>in situ</i>	Zinco total	Sólidos dissolvidos totais
-	-	-	-	Zinco dissolvido	pH <i>in situ</i>	Temperatura da amostra <i>in situ</i>
-	-	-	-	Zinco total	pH lab	Zinco dissolvido
-	-	-	-	pH <i>in situ</i>	-	Zinco total
-	-	-	-	pH lab	-	pH <i>in situ</i>
-	-	-	-	-	-	pH lab
H	I	J	K	L	M	N
Alcalinidade total	Alcalinidade total	Alcalinidade total	Alcalinidade total	Alcalinidade total	Alcalinidade total	Alcalinidade total
Alumínio dissolvido	Alumínio dissolvido	Alumínio dissolvido	Alumínio dissolvido	Alumínio dissolvido	Alumínio dissolvido	Alumínio dissolvido
Alumínio total	Alumínio total	Alumínio total	Alumínio total	Alumínio total	Alumínio total	Bário dissolvido
Antimônio dissolvido	Antimônio dissolvido	Antimônio dissolvido	Bário dissolvido	Antimônio total	Antimônio total	Bário total
Arsênio dissolvido	Arsênio dissolvido	Antimônio total	Berílio total	Bário dissolvido	Arsênio dissolvido	Cálcio total
Bário dissolvido	Arsênio total	Arsênio dissolvido	Cálcio total	Cálcio total	Arsênio total	Carbono orgânico total
Bário total	Bário dissolvido	Arsênio total	Carbono orgânico dissolvido	Carbono orgânico dissolvido	Bário dissolvido	Chumbo total
Cálcio total	Bário total	Bário dissolvido	Cianeto livre	Cianeto livre	Bário total	Manganês dissolvido
Carbono orgânico dissolvido	Boro total	Berílio total	Cloreto total	Cloreto total	Berílio total	Temperatura da amostra <i>in situ</i>
Cianeto livre	Cálcio total	Boro total	Clorofila a	Clorofila a	Boro total	-
Cloreto total	Carbono orgânico dissolvido	Cádmio total	Cobre dissolvido	Cobre dissolvido	Cálcio total	-
Clorofila a	Chumbo total	Cálcio total	Cobre total	Condutividade <i>in situ</i>	Carbono orgânico dissolvido	-
Cobalto total	Cianeto livre	Carbono orgânico dissolvido	Condutividade <i>in situ</i>	Cor verdadeira	Cianeto livre	-
Cobre dissolvido	Cloreto total	Cianeto livre	Escherichia coli	DBO	Cloreto total	-
Cobre total	Cobre dissolvido	Cloreto total	Fósforo total	Dureza total	Clorofila a	-
Condutividade <i>in situ</i>	Cobre total	Clorofila a	Magnésio total	Nitrato	Cobre dissolvido	-

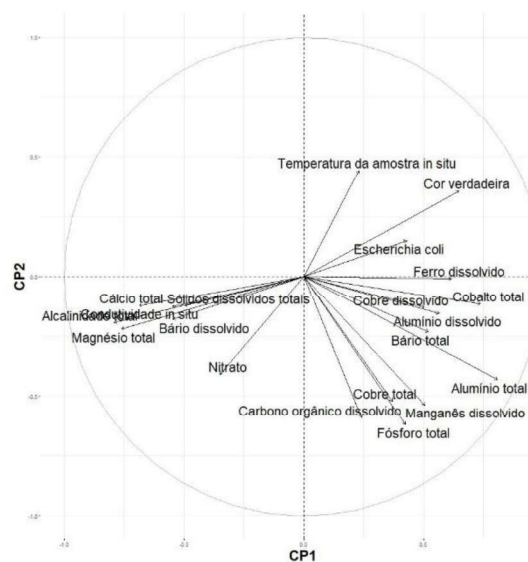
A	B	C	D	E	F	G
Cor verdadeira	Condutividade <i>in situ</i>	Cobalto dissolvido	Manganês dissolvido	Nitrito	Cobre total	-
DBO	Cor verdadeira	Cobre dissolvido	Nitrato	Nitrogênio amoniacal	Condutividade <i>in situ</i>	-
Escherichia coli	Cromo total	Condutividade <i>in situ</i>	Nitrito	Nitrogênio kjeldahl total	Cor verdadeira	-
Feoftina	DBO	Cor verdadeira	Nitrogênio amoniacal	Oxigênio dissolvido <i>in situ</i>	DBO	-
Ferro II	Escherichia coli	DBO	Nitrogênio kjeldahl total	Polifosfato	Escherichia coli	-
Ferro dissolvido	Ferro II	Dureza total	Oxigênio dissolvido <i>in situ</i>	Potencial redox <i>in situ</i>	Ferro II	-
Magnésio total	Ferro dissolvido	Escherichia coli	Polifosfato	Sódio total	Ferro dissolvido	-
Manganês dissolvido	Fósforo total	Feoftina	Potencial redox <i>in situ</i>	Sólidos dissolvidos totais	Fósforo total	-
Nitrato	Magnésio total	Ferro II	Sódio total	Sólidos sedimentáveis	Magnésio total	-
Nitrito	Manganês dissolvido	Fósforo dissolvido	Sólidos dissolvidos totais	Sulfetos totais	Manganês dissolvido	-
Nitrogênio amoniacal	Níquel total	Fósforo total	Temperatura da amostra <i>in situ</i>	Temperatura da amostra <i>in situ</i>	Molibdênio total	-
Nitrogênio kjeldahl total	Nitrato	Manganês total	Zinco dissolvido	Zinco dissolvido	Níquel total	-
Oxigênio dissolvido <i>in situ</i>	Nitrito	Níquel total	Zinco total	pH <i>in situ</i>	Nitrato	-
Oxigênio dissolvido saturado <i>in situ</i>	Nitrogênio amoniacal	Nitrato	pH <i>in situ</i>	pH lab	Nitrito	-
Potencial redox <i>in situ</i>	Nitrogênio kjeldahl total	Nitrito	pH lab	-	Nitrogênio amoniacal	-
Sódio total	Oxigênio dissolvido <i>in situ</i>	Nitrogênio amoniacal	-	-	Nitrogênio kjeldahl total	-
Sólidos dissolvidos totais	Polifosfato	Nitrogênio kjeldahl total	-	-	Oxigênio dissolvido <i>in situ</i>	-
Zinco dissolvido	Potencial redox <i>in situ</i>	Oxigênio dissolvido <i>in situ</i>	-	-	Polifosfato	-
Zinco total	Sólidos dissolvidos totais	Polifosfato	-	-	Potencial redox <i>in situ</i>	-
pH <i>in situ</i>	Sólidos suspensos totais	Potencial redox <i>in situ</i>	-	-	Selênio total	-
pH lab	Temperatura da amostra <i>in situ</i>	Selênio total	-	-	Sódio total	-
-	Zinco dissolvido	Sódio total	-	-	Sólidos dissolvidos totais	-
-	Zinco total	Sólidos dissolvidos totais	-	-	Sulfetos totais	-
-	pH <i>in situ</i>	Sólidos sedimentáveis	-	-	Temperatura da amostra <i>in situ</i>	-

A	B	C	D	E	F	G
-	pH lab	Sulfetos totais	-	-	Vanádio dissolvido	-
-	-	Temperatura da amostra <i>in situ</i>	-	-	Zinco dissolvido	-
-	-	Zinco dissolvido	-	-	Zinco total	-
-	-	Zinco total	-	-	pH <i>in situ</i>	-
-	-	pH <i>in situ</i>	-	-	pH lab	-
-	-	pH lab	-	-	-	-

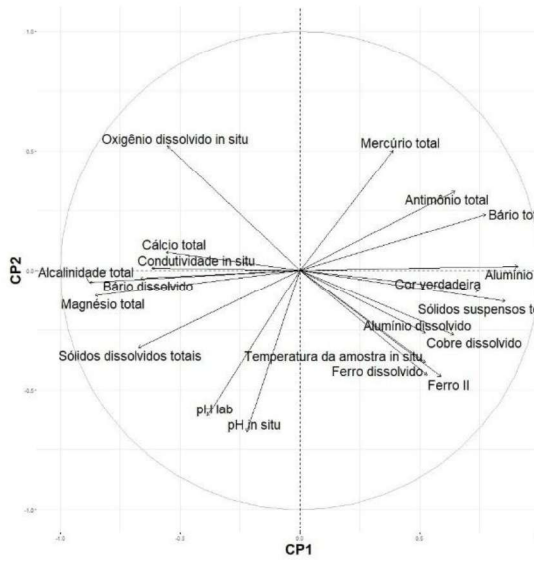
Figura A15 – Contribuição de variáveis de qualidade da água “não-críticas” as componentes CP1 e CP2 nas subregiões A (a), B (b), C (c), E (d), F (e), G (f), H (g), I (h), K (i), L (j), M (k) e N (l)



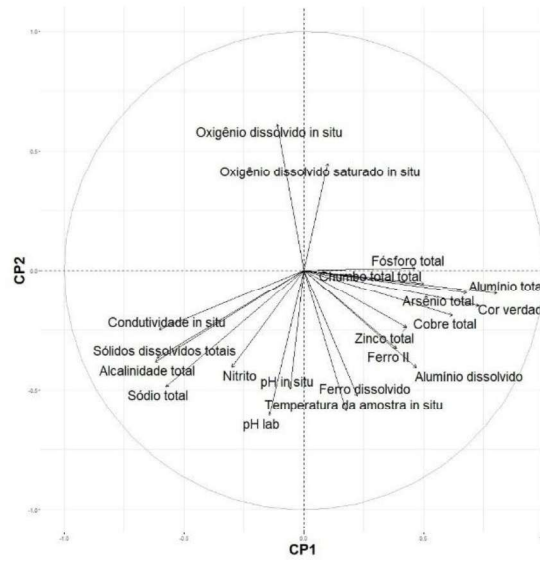
a



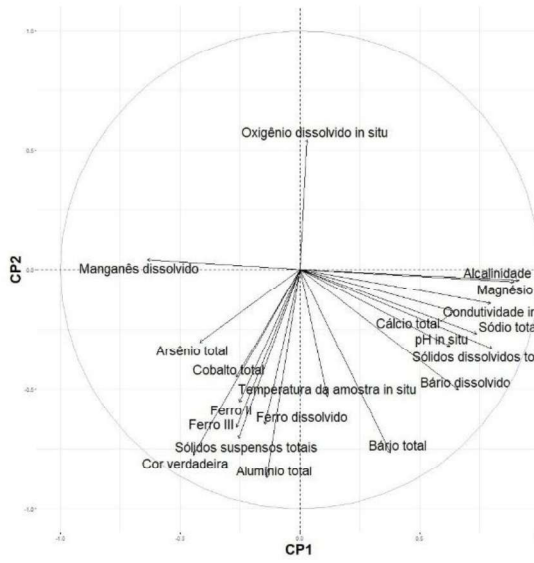
b



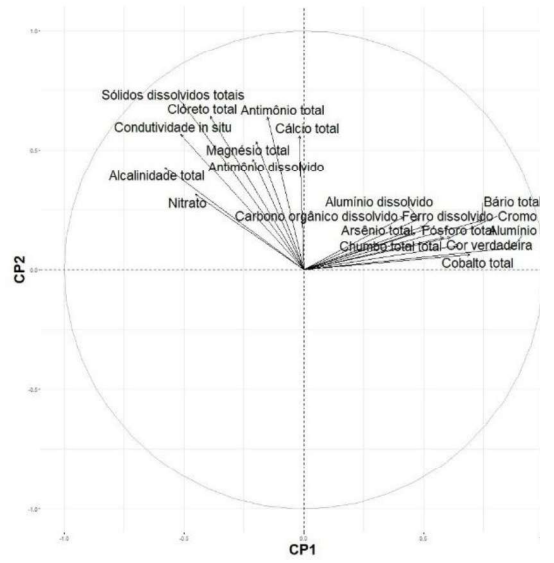
c



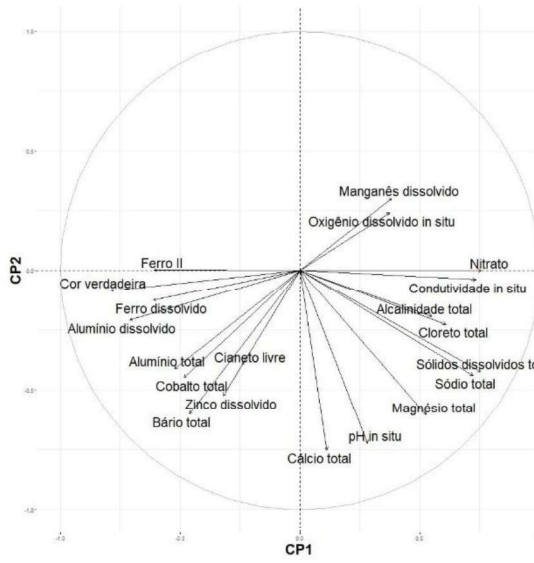
d



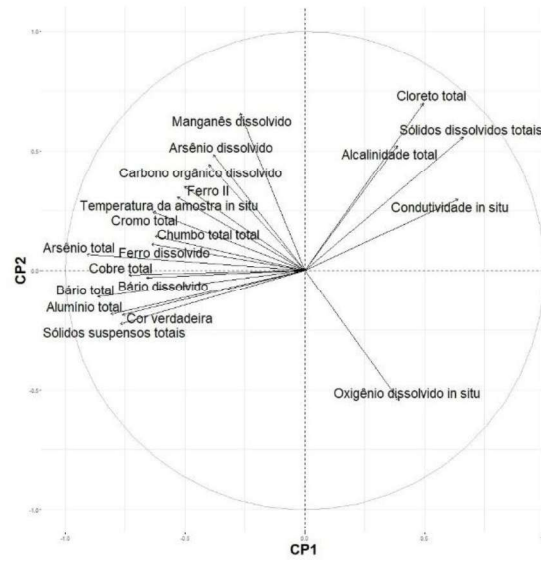
e



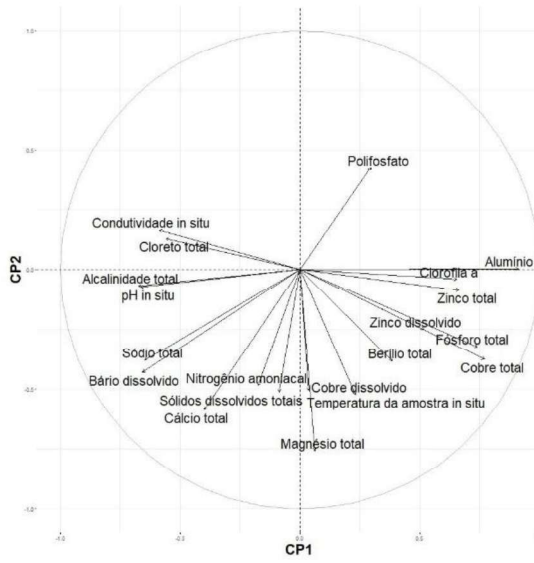
f



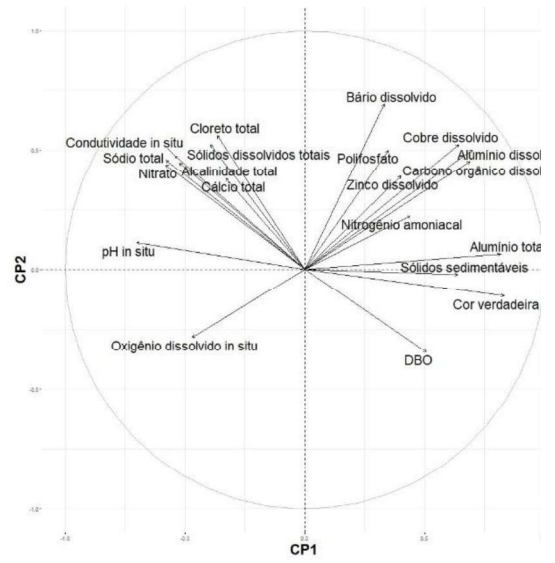
g



h



i



j

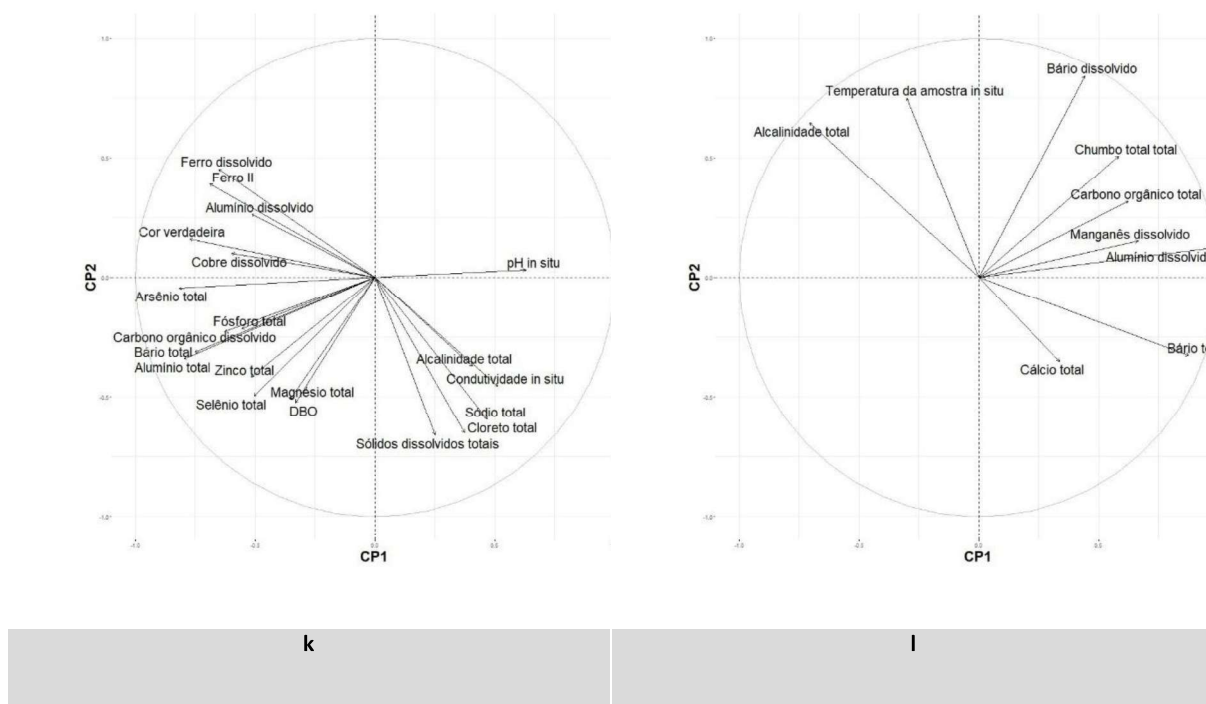
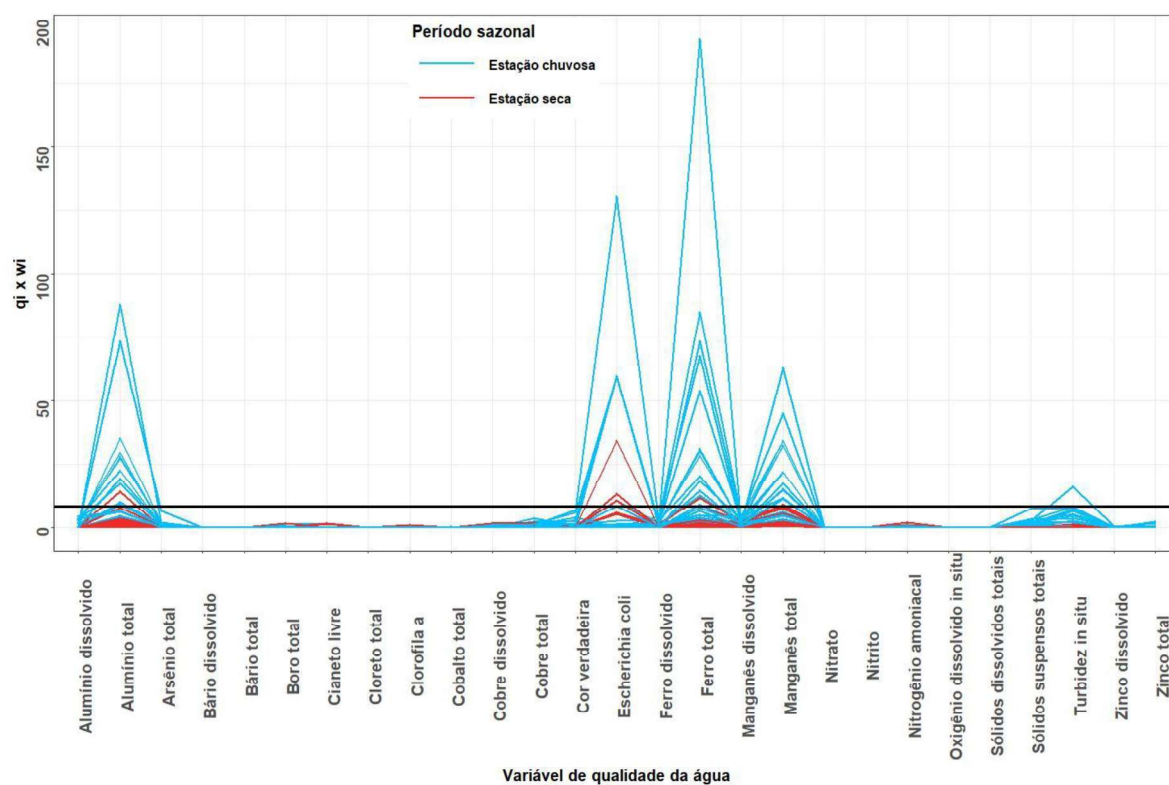
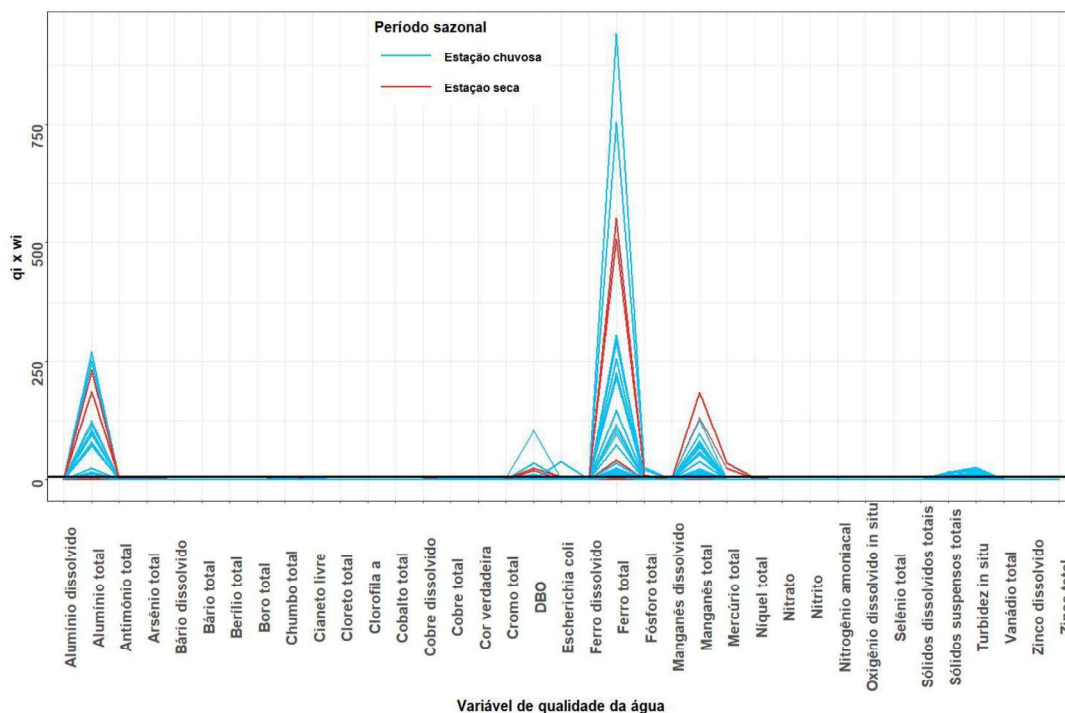


Figura A16 – Contribuição de variáveis de qualidade da água para EWQI na subregião A*



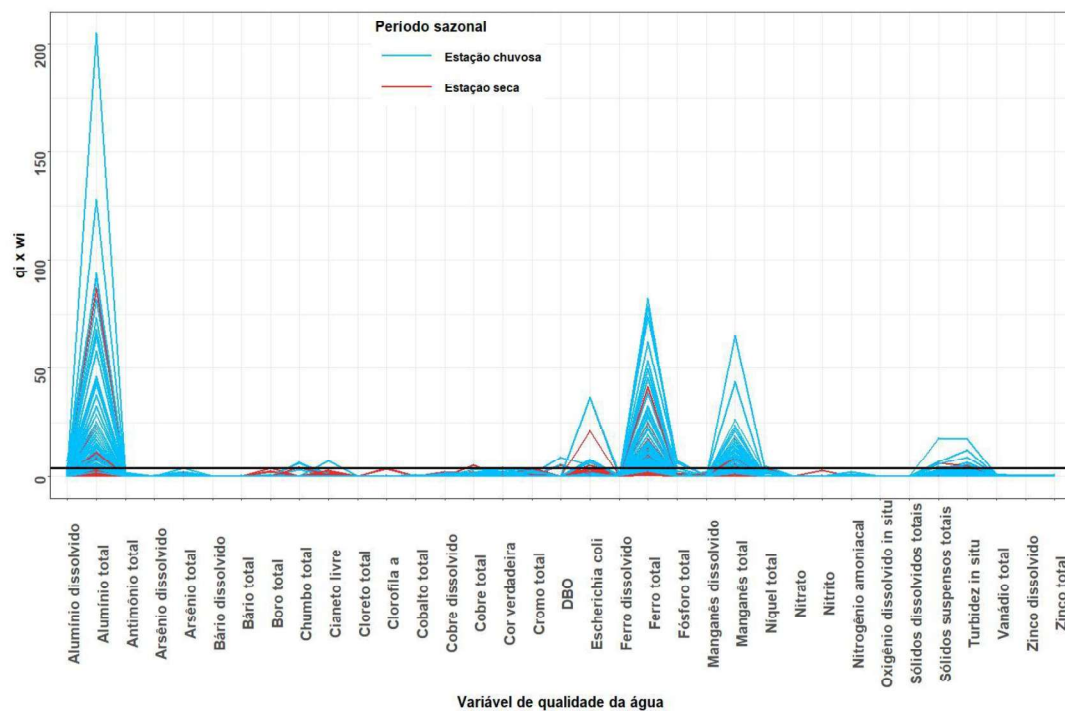
***A linha preta representa o percentil 95% de qi x wi de todas as campanhas da subregião analisada.**

Figura A17 – Contribuição de variáveis de qualidade da água para EWQI na subregião C*



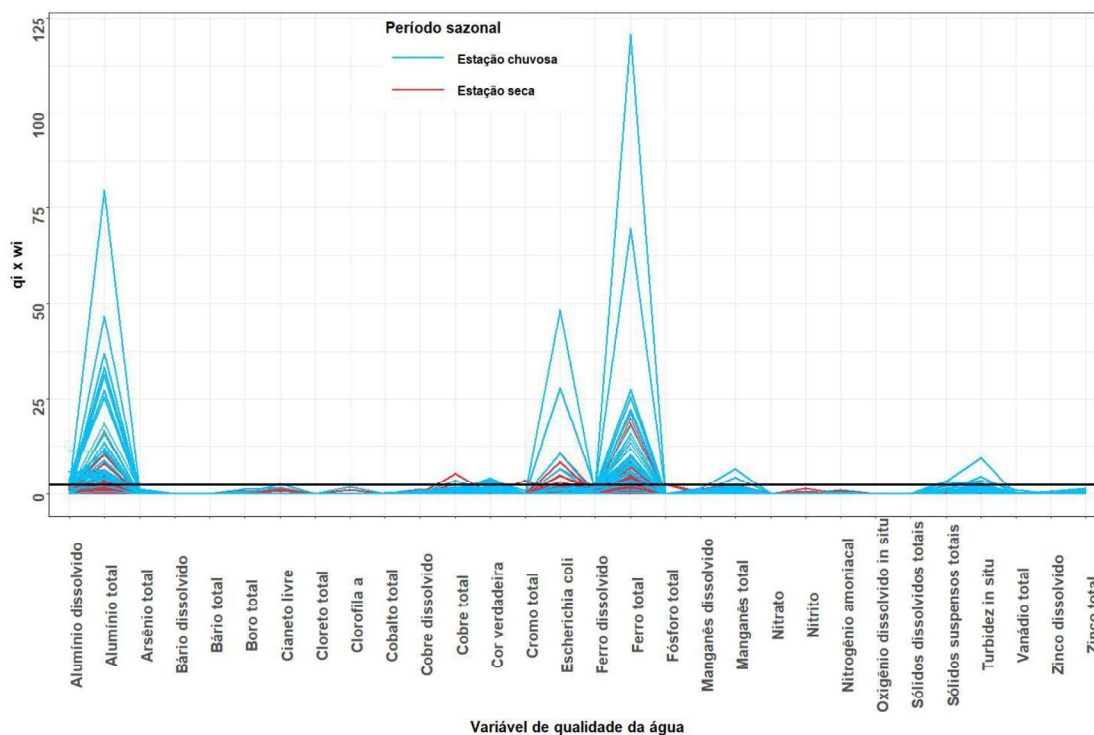
***A linha preta representa o percentil 95% de $q_i \times w_i$ de todas as campanhas da subregião analisada.**

Figura A18 – Contribuição de variáveis de qualidade da água para EWQI na subregião E*



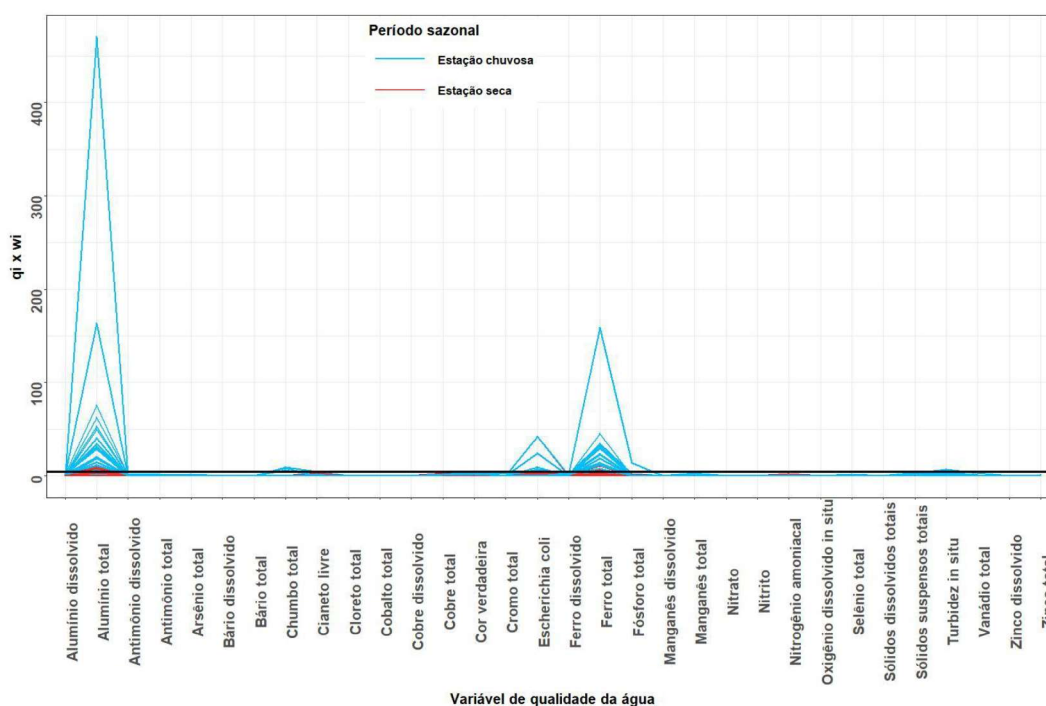
***A linha preta representa o percentil 95% de $q_i \times w_i$ de todas as campanhas da subregião analisada.**

Figura A19 – Contribuição de variáveis de qualidade da água para EWQI na subregião F*



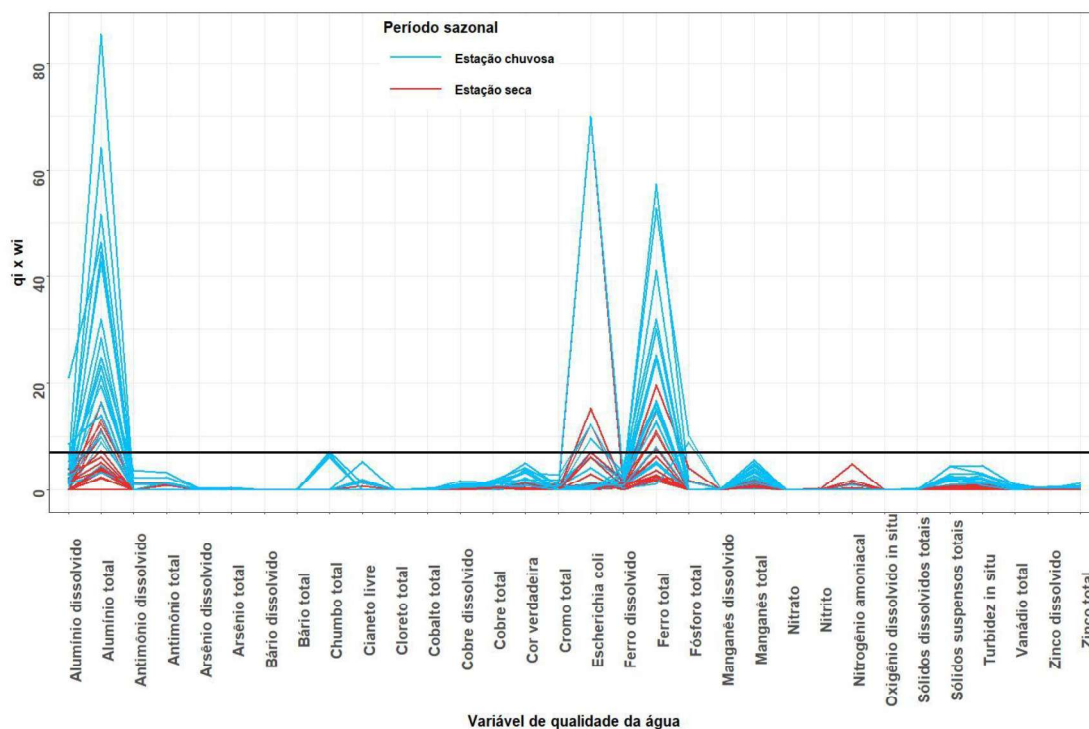
***A linha preta representa o percentil 95% de $q_i \times w_i$ de todas as campanhas da subregião analisada.**

Figura A20 – Contribuição de variáveis de qualidade da água para EWQI na subregião H*



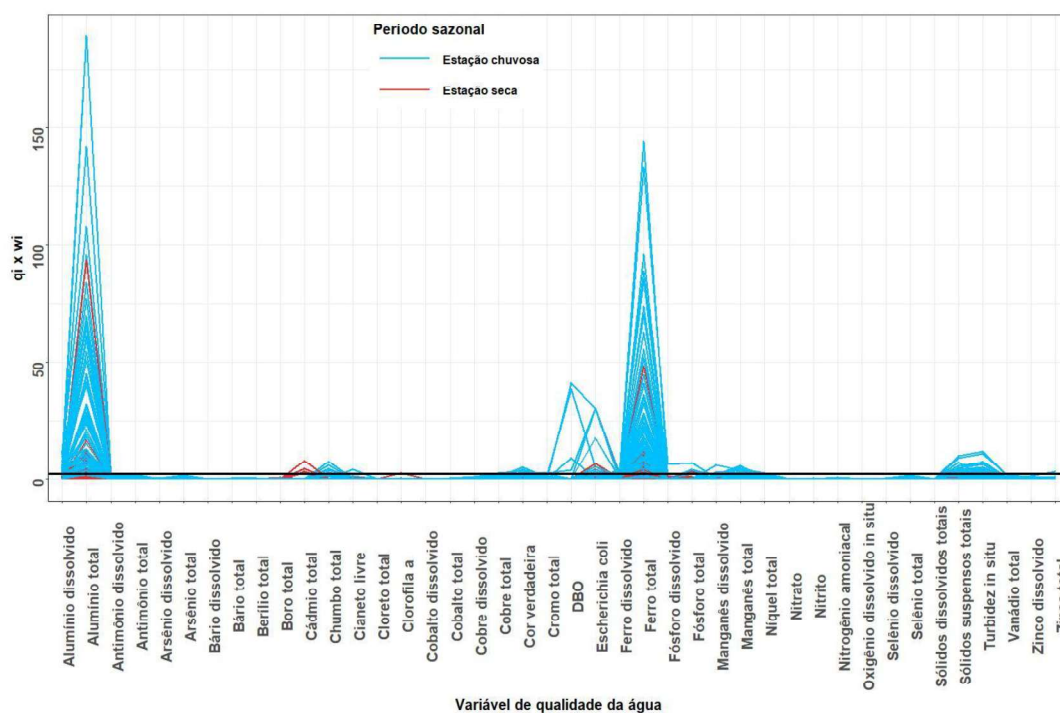
***A linha preta representa o percentil 95% de $q_i \times w_i$ de todas as campanhas da subregião analisada.**

Figura A21 – Contribuição de variáveis de qualidade da água para EWQI na subregião I*



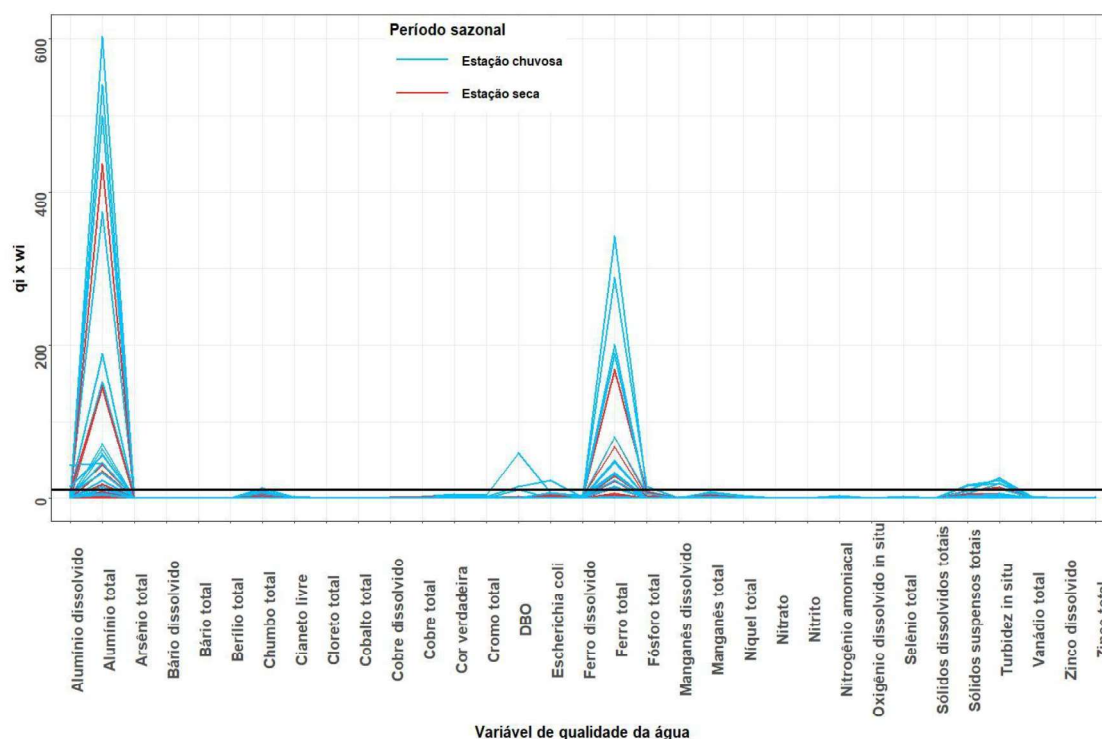
*A linha preta representa o percentil 95% de $q_i \times w_i$ de todas as campanhas da subregião analisada.

Figura A22 – Contribuição de variáveis de qualidade da água para EWQI na subregião J*



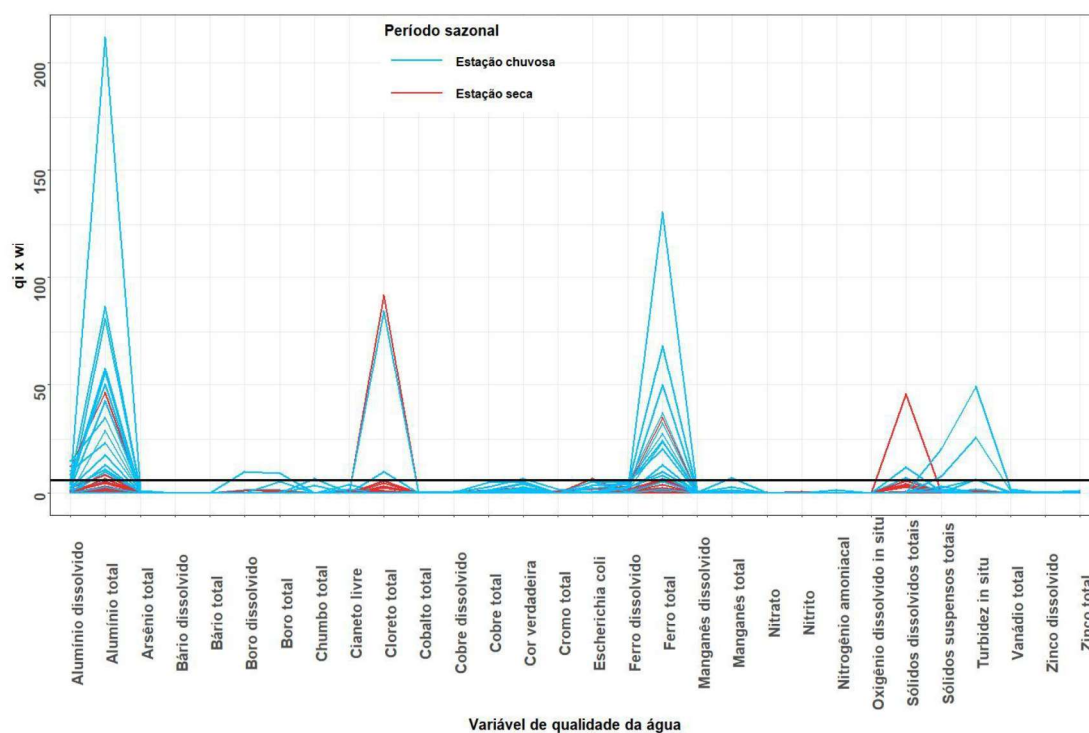
*A linha preta representa o percentil 95% de $q_i \times w_i$ de todas as campanhas da subregião analisada.

Figura A23 - Contribuição de variáveis de qualidade da água para EWQI na subregião K



***A linha preta representa o percentil 95% de $q_i \times w_i$ de todas as campanhas da subregião analisada.**

Figura A24 – Contribuição de variáveis de qualidade da água para EWQI na subregião N*



***A linha preta representa o percentil 95% de $q_i \times w_i$ de todas as campanhas da subregião analisada.**

Figura A25 – Análise de sensibilidade para determinação do tamanho ideal do *bin* para a subregião B

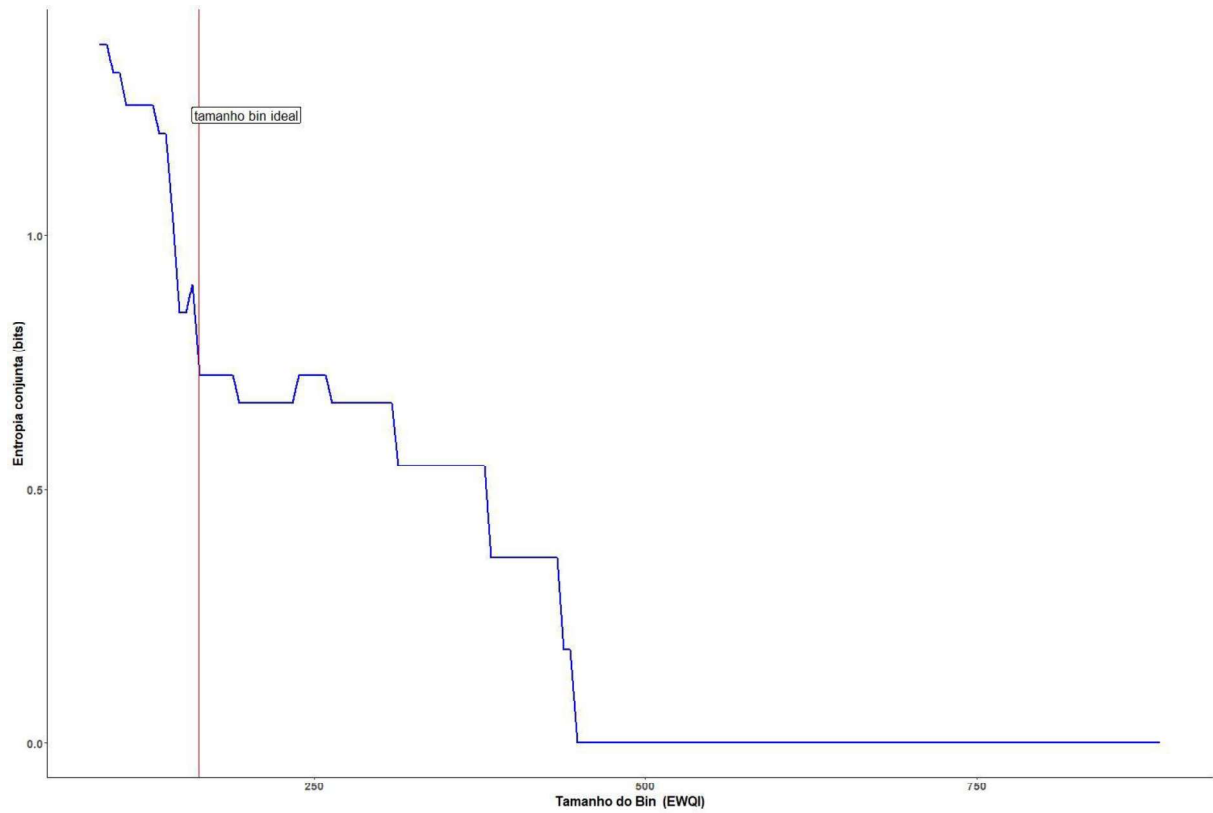


Figura A26 – Análise de sensibilidade para determinação do tamanho ideal do *bin* para a subregião C

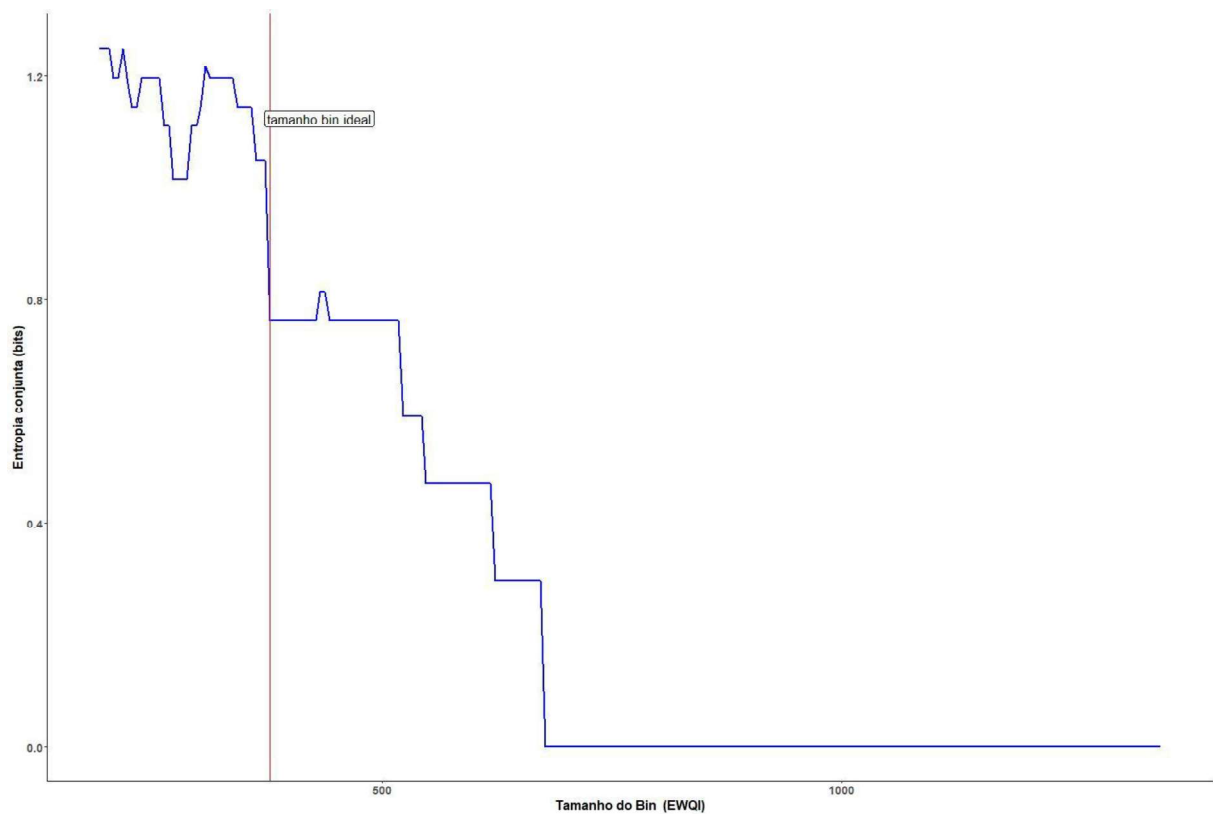


Figura A27 – Análise de sensibilidade para determinação do tamanho ideal do *bin* para a subregião D

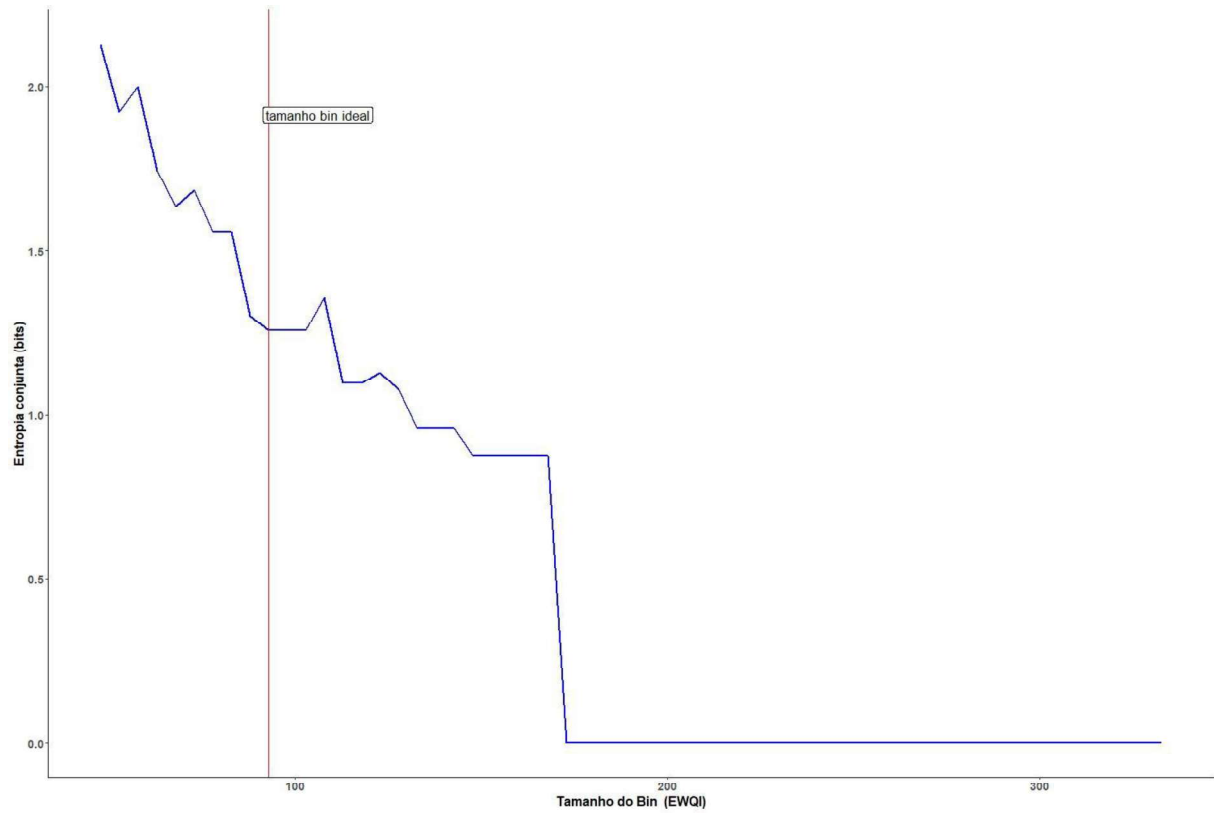


Figura A29 – Análise de sensibilidade para determinação do tamanho ideal do *bin* para a subregião F

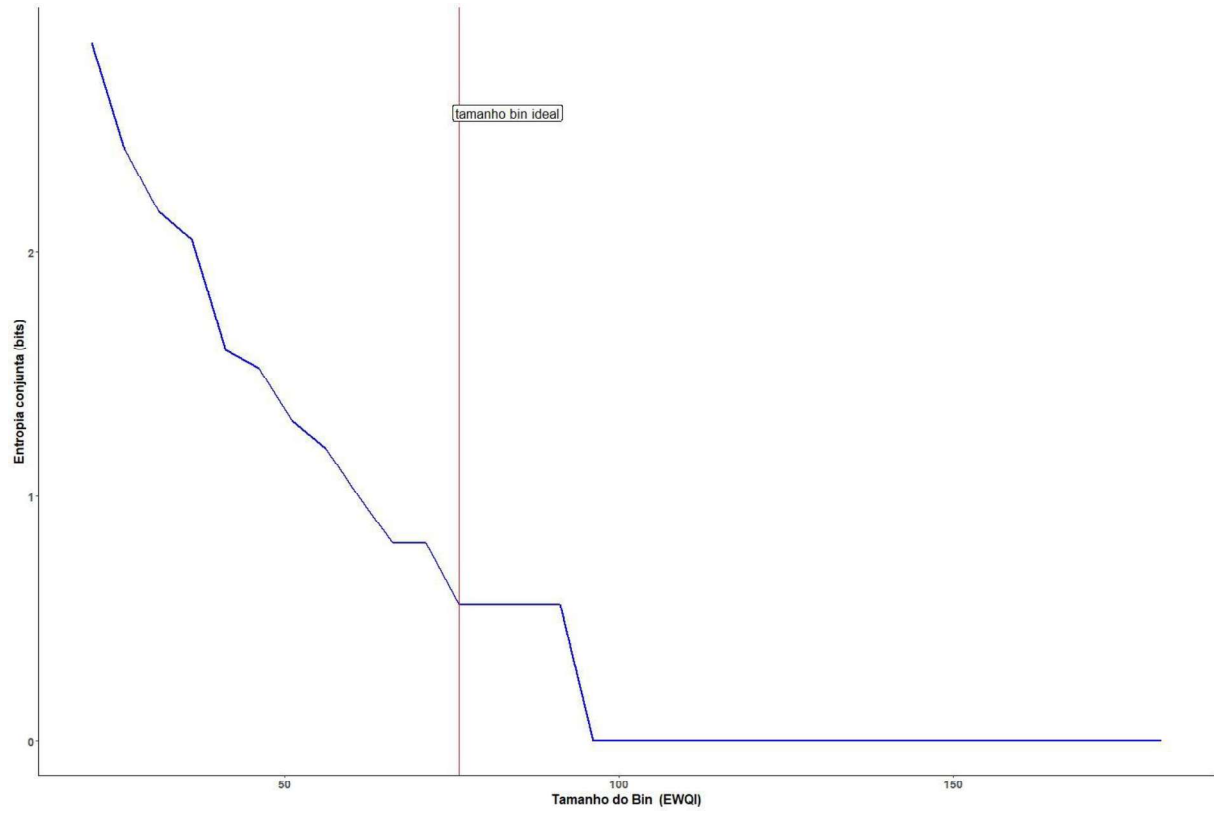


Figura A30 – Análise de sensibilidade para determinação do tamanho ideal do *bin* para a subregião G

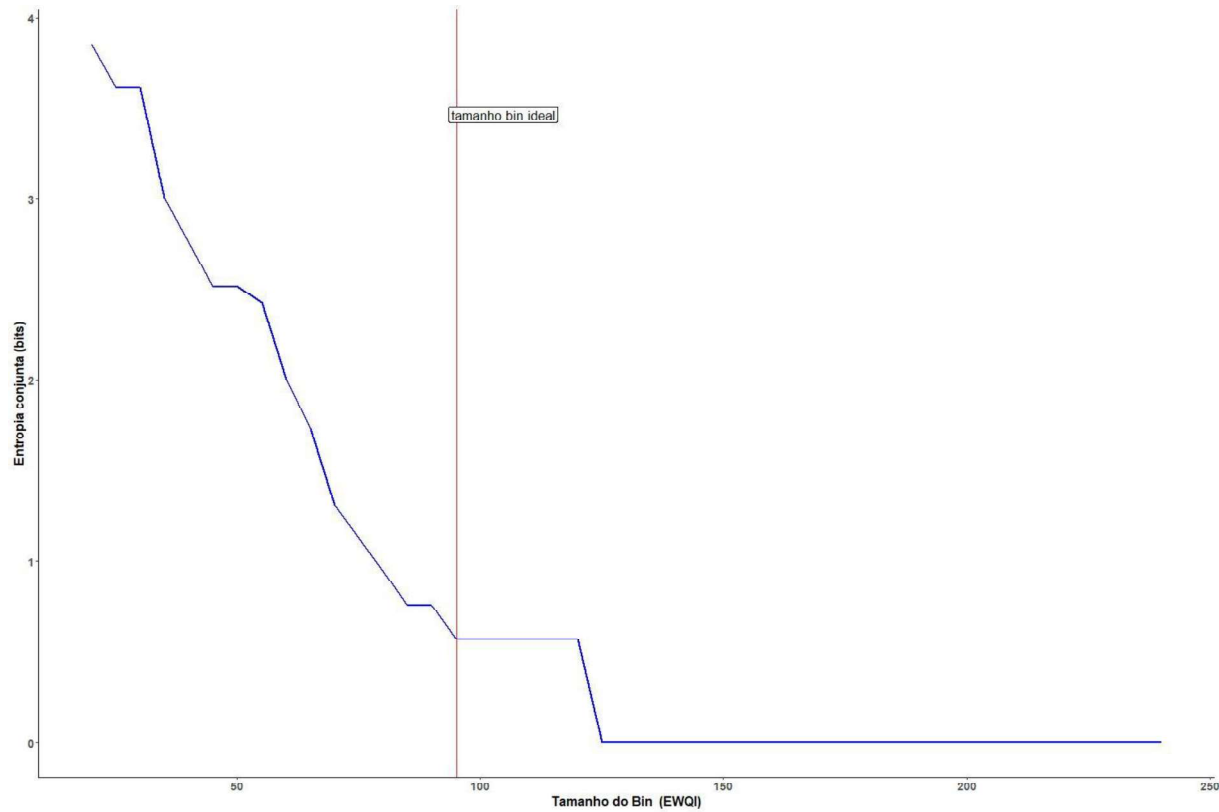


Figura A31 – Análise de sensibilidade para determinação do tamanho ideal do *bin* para a subregião J

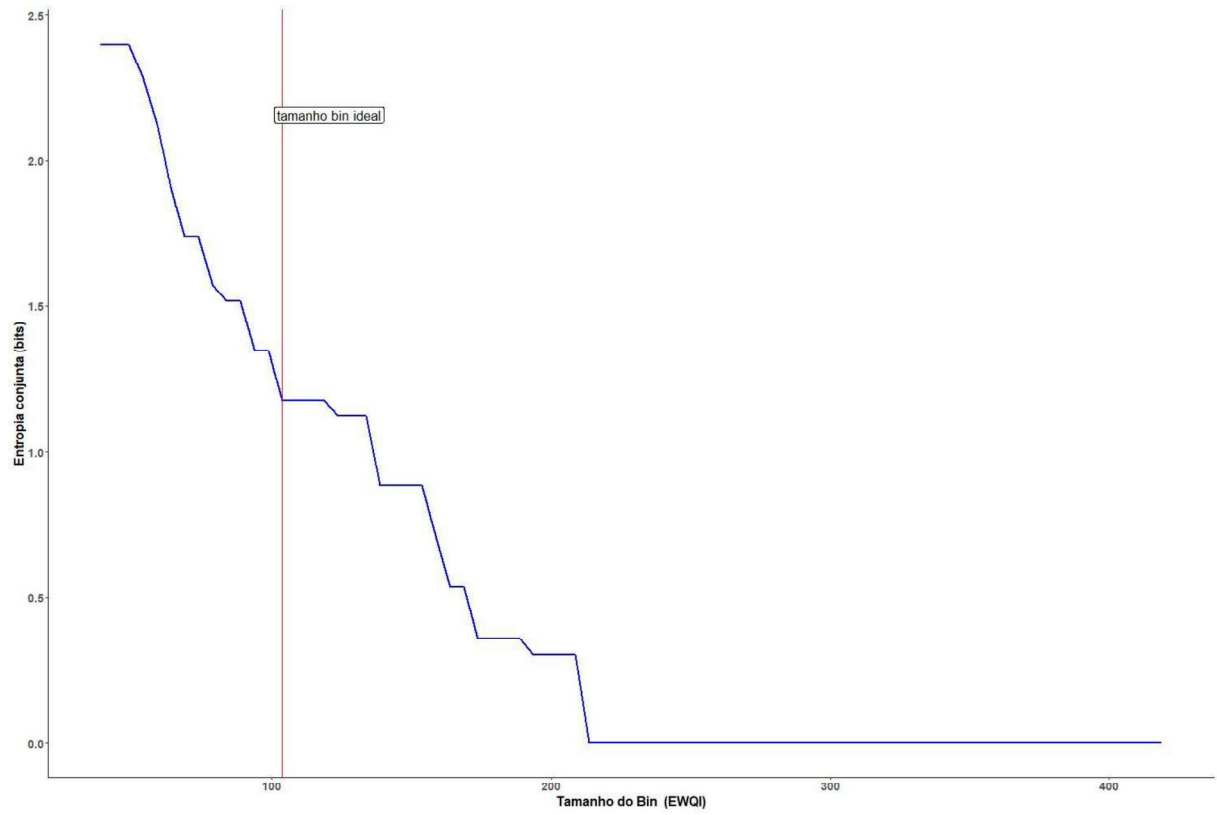


Figura A32 – Análise de sensibilidade para determinação do tamanho ideal do *bin* para a subregião M

