

Parâmetros significativos para monitoramento e avaliação da qualidade da água, Bacia Hidrográfica do Rio Cachoeira (Joinville, Santa Catarina, Brasil)

Significant parameters for monitoring and evaluation of water quality in the Cachoeira River Hydrographic Basin (Joinville, Santa Catarina, Brazil)

Glauber Rover Cadorin¹ , Anelise Destefani² , Carlos Eduardo Nogueira Martins² ,
André Luis Fachini de Souza² 

¹Companhia Águas de Joinville, Joinville, SC, Brasil. E-mail: roverkid@gmail.com

²Instituto Federal Catarinense – IFC, Araquari, SC, Brasil. E-mails: anelise.destefani@ifc.edu.br, carlos.martins@ifc.edu.br, andre.fachini@ifc.edu.br

Como citar: Cadorin, G. R., Destefani, A., Martins, C. E. N., & Souza, A. L. F. (2023). Parâmetros significativos para monitoramento e avaliação da qualidade da água, Bacia Hidrográfica do Rio Cachoeira (Joinville, Santa Catarina, Brasil). *Revista de Gestão de Água da América Latina*, 20, e14. <https://doi.org/10.21168/rega.v20e14>

RESUMO: Índices de qualidade de água são ferramentas utilizadas para sintetizar, por meio de um único valor, contribuições de variáveis físicas, químicas e biológicas de corpos hídricos. Dentre as várias ferramentas disponíveis para avaliar a qualidade da água, o Índice de Qualidade das Águas (IQA) é um dos instrumentos mais utilizados. Este trabalho tem como objetivo identificar entre os parâmetros constituintes do IQA, variáveis significativas para otimização do monitoramento da qualidade de corpos hídricos, buscando reduções de parâmetros, tempo e custo para realização de análises. Os dados utilizados contemplam campanhas em 19 rios que compõem a Bacia Hidrográfica do Rio Cachoeira na região nordeste de Santa Catarina (Brasil), no período de 2014 a 2021. Ferramentas estatísticas como análise de componentes principais, análise de similaridade e regressão linear múltipla foram utilizadas para análise dos dados experimentais. Os resultados sugerem oxigênio dissolvido, demanda bioquímica de oxigênio, coliformes, fósforo e turbidez como variáveis significativas para a realização dos monitoramentos. Modelos matemáticos gerados apresentaram assertividade média de 81% quando comparados ao IQA. A proposta de redução das variáveis sugeriu viabilidade de aplicação prática, com redução de tempo e custos.

Palavras-chave: Água Superficial; Índice de Qualidade; Recursos Hídricos.

ABSTRACT: Water quality indices are tools commonly used to synthesize, by a single value or category, the contributions of physical, chemical and biological variables of water bodies. Among the several tools available to assess water quality, the Water Quality Index (WQI) is one of the most commonly used instruments. In this sense, this work aims to identify, among the constituent parameters of the WQI, groups of significant variables for the optimization of the monitoring of the quality of water bodies, while seeking reductions in parameters, time and cost for carrying out analyses. The data used in this work included campaigns from 2014 to 2021 to collect surface liquid samples from 19 rivers that make up a watershed in the North Eastern region of Santa Catarina, Brazil, with tidal intrusion. To study the variables, statistical tools such as principal component analysis, similarity analysis and multiple linear regression were used in computer programs. The results suggest the parameters dissolved oxygen, biochemical oxygen demand, coliforms, phosphorus and turbidity as significant variables for carrying out the monitoring. The generated mathematical models showed an average assertiveness of 81% when their results were compared to the WQI. The proposal to reduce the variables indicates feasibility of practical application, mainly because they represent a reduction in time and cost to carry out the monitoring of the rivers studied.

Keywords: Surface Water; Quality Index; Water Resources.

INTRODUÇÃO

A água é um recurso natural de valor econômico, ambiental e social, fundamental para a sobrevivência dos seres humanos e dos ecossistemas. Em termos ambientais, é globalmente considerada um bem comum da humanidade, insubstituível e essencial para toda vida que existe no planeta Terra (Capellari & Capellari, 2018).

Quando assegurado por reconhecimentos internacionais, o acesso a água potável passa a ter um papel social importante por estar associado ao princípio da dignidade humana (Souza et al., 2020).

Cabe a União, Estados e Municípios, avaliar de forma regionalizada o acesso a água na busca por um gerenciamento mais eficiente desse recurso. Essa avaliação contribui para evitar o aparecimento de situações que apontem fragilidade social como a crise hídrica ocorrida no estado de São Paulo e recorrente em várias regiões do país (Zancul, 2015). Atualmente, dois terços da população mundial passam pelo menos uma vez ao ano por restrição de água (Prado et al., 2018).

O risco da escassez hídrica está relacionado ao uso e ocupação das áreas urbanas. A ocupação urbana brasileira ocorreu de forma acelerada, desordenada e sem considerar os impactos ao meio ambiente e ao próprio indivíduo, resultando em problemas como destruição de fauna e flora, enchentes e assoreamento dos leitos de rios (Reis et al., 2020). Consequentemente, esses rios tornaram-se fontes de transmissão de doenças por mosquitos e parasitas, representando riscos à saúde pública. Febre-amarela, dengue e cólera são alguns exemplos de doenças relacionadas a falta de saneamento básico e ainda endêmicas no Brasil (Camargo et al., 2017).

Nesse sentido, muitas bacias hidrográficas foram ocupadas pelo homem por possuírem água com qualidade e abundância para consumo, cultivo e até mesmo atividades industriais. Essa ocupação acarreta despejos inadequados de resíduos como lixo e esgoto doméstico. Como consequência, os rios acabam adquirindo consideráveis cargas poluidoras (Reis et al., 2020).

Com o compromisso de universalizar o acesso ao saneamento básico até 2033, foi aprovado pelo governo federal o novo marco legal do saneamento básico. O tema trata da demanda urgente em aumentar o acesso a água potável e a coleta e tratamento de esgoto em locais com pouca ou nenhuma infraestrutura (Araújo & Graças e Silva, 2021).

Para que políticas públicas possam ser executadas de maneira assertiva, é necessária a utilização de indicadores que permitam uma forma simples de classificar a qualidade dos rios, principalmente para um público não especialista. Estas informações, quando de fácil compreensão, irão auxiliar para que o gerenciamento dos ambientes aquáticos possa ser realizado de maneira eficaz (Pessoa et al., 2020).

As primeiras tentativas de definição de indicadores para qualidade da água foram introduzidas há mais de 150 anos. Na Alemanha, em meados de 1848, a presença ou ausência de determinados organismos servia como base para analisar a qualidade de rios. Apesar destas tentativas terem contribuído para a evolução dos indicadores de qualidade, somente em 1965 foi introduzido o que seria definido como o primeiro índice de qualidade da era moderna, o índice Horton (Abbasi & Abbasi, 2012).

Entre os anos de 2000 e 2019 muitas propostas para aplicação de parâmetros de controles de qualidade foram realizadas com variação dos ensaios empregados e das frequências de monitoramento (Rezaie-Balf et al., 2020). Os estudos foram realizados, principalmente, na China, Espanha, Índia, Malásia e Vietnã. A análise de pH foi sugerida como significativa em 95,83% dos estudos verificados, seguido de OD com 91,67%, DBO com 75% e coliformes fecais com 50%. As análises de turbidez (33,33%), temperatura (29,17%), fósforo (25%), sólidos totais (16,67%) e nitrogênio (12,50%) também foram citadas, mas em menos da metade dos estudos analisados (Rezaie-Balf et al., 2020).

A classificação baseada em indicadores de nascentes de águas para fins de abastecimento e irrigação foi utilizada na região montanhosa de Jammu Himalaia, na Índia. Também foi utilizada na comparação de diferentes indicadores para a classificação de reservatórios na região do rio Aliakmon, na Grécia (Taloor et al., 2020).

Adicionalmente, utilizaram-se indicadores para avaliar as mudanças sazonais e espaciais da qualidade da água em um dos projetos de desvio de água na bacia hidrográfica *Middle-Route* (MR), na China, considerada uma das mais longas do mundo (Nong et al., 2020).

A avaliação de possível impacto ambiental causado por atividades domésticas, industriais e agrícolas, por meio de indicadores de qualidade focados nas propriedades físico-químicas, foram realizados para avaliar a Bacia Hidrográfica de Coruh, na Turquia (Bilgin, 2015). Também, riscos de contaminação de águas subterrâneas, devido as atividades industriais na região de Camaçari (Bahia, Brasil), motivaram estudos para verificar a qualidade dos recursos hídricos presentes na região (Oliveira, 2018).

Estas categorias de estudos promoveram discussões quanto a variabilidade dos dados em função de sazonalidade, onde períodos secos e chuvosos foram estudados como variáveis significativas para

a avaliação. Em busca por alternativas para o abastecimento público na região metropolitana de Chittagong (Bangladesh), pesquisadores exploraram o uso de indicador para entender o nível de potabilidade de água em períodos chuvoso e de inverno (Molla et al., 2022). De maneira similar, a avaliação da variação dos parâmetros físico-químicos, relacionados à qualidade da água, em função da sazonalidade foi estudado no lago Hebbal (Bangalore, Índia) (Sudarshan et al., 2019).

Adicionalmente, tem-se buscado selecionar as características físicas e químicas que possam explicar a variabilidade da qualidade da água com a utilização de ferramentas estatísticas (Cecílio et al., 2013). Estudos propõem a redução de variáveis por meio de testes estatísticos em busca de um monitoramento mais vantajoso e menos custoso (Rabelo et al., 2020).

No Brasil, o Índice de Qualidade da Água (IQA) da Companhia Ambiental do Estado de São Paulo (CETESB) é um dos índices mais utilizados atualmente (Brasil, 2004). O IQA foi desenvolvido na década de 1970 pela *National Sanitation Foundation* (NSF), dos Estados Unidos. Foram definidos nove parâmetros para a composição do indicador, cada qual com seu respectivo peso em função de sua importância para a conformação global de qualidade da água: coliformes fecais (Termo tolerantes) (peso 0,15), pH (peso 0,12), demanda bioquímica de oxigênio (DBO) (peso 0,10), nitrogênio total (peso 0,10), fosfato total (peso 0,10), temperatura (peso 0,10), turbidez (peso 0,08), sólidos (resíduos) totais (peso 0,08) e saturação de oxigênio dissolvido (OD) (peso 0,17) (Castro & Venâncio, 2019).

Para o IQA, as medidas dos parâmetros analisados previamente são convertidas em subíndices. Estes valores, em seguida, são combinados de maneira a produzir uma pontuação única, o que apresenta vantagem de comunicação, pois a informação torna-se clara e resumida para o público leigo (Rodrigues et al., 2020).

Apesar de ser uma ferramenta amplamente utilizada no Brasil, o IQA apresenta limitações. Isto porque foi desenvolvido inicialmente para avaliação da qualidade de rios para fins de abastecimento, considerando aspectos relativos ao tratamento desta água (Oliveira et al., 2020).

Na cidade de Joinville-SC, o sistema IQA é utilizado para monitoramento da qualidade de bacias hidrográficas como a Bacia do Rio Cachoeira. Dentro das doze principais regiões hidrográficas do Brasil, a região de Joinville, localizada na parte nordeste do estado de Santa Catarina (SC), pertence à região hidrográfica do Atlântico Sul. Os rios compreendidos nesta bacia são classificados como de pequeno porte e que correm diretamente no sentido do Oceano Atlântico. Adicionalmente, manguezais e restingas fazem parte dos ecossistemas associados a essa região hidrográfica (Oliveira et al., 2017).

O município de Joinville é o mais populoso do estado e aparece na 85ª posição na classificação nacional de atendimento com rede coletora de esgoto dentre as 100 maiores cidades brasileiras (Instituto Trata Brasil, 2021). Com o aumento da cobertura do sistema de coleta e tratamento de esgoto doméstico na região de Joinville-SC, o IQA auxilia na avaliação da evolução da qualidade dos rios que deixaram de receber cargas de esgoto (Comitê de Gerenciamento das Bacias Hidrográficas da Bacia da Babitonga e Bacias Contíguas, 2016).

A aplicação de métodos estatísticos pode facilitar na manipulação de dados e ajudar na identificação de fontes de poluição no monitoramento de corpos hídricos. As análises de correlação e de variância servem de exemplos de métodos estatísticos que permitem identificar os possíveis motivos ou fontes responsáveis pelas alterações na qualidade da água do corpo hídrico (Santos et al., 2020). Para que a avaliação, por meio do uso dos indicadores, possa ser realizada de maneira adequada, estudos sugerem como alternativa a utilização de ferramentas estatísticas como a análise multivariada para a avaliação de múltiplas medidas sobre o objeto de investigação (Cecconello et al., 2018).

Nesse sentido, este estudo teve como objetivo identificar entre os parâmetros constituintes do IQA, grupos de variáveis significativas para otimização do monitoramento da qualidade de corpos hídricos da cidade de Joinville-SC.

Buscou-se, de maneira mais específica, definir variáveis mais significativas por meio de testes estatísticos, avaliar a possibilidade de redução de variáveis utilizadas para o monitoramento de rios e condensar as variáveis constituintes do IQA, por meio de modelos matemáticos mais adequados à realidade da região estudada, em um conjunto menor de componentes e com uma perda mínima de informação.

MATERIAIS E MÉTODOS

Área de estudo

Os estudos foram realizados em rios que compreendem a bacia hidrográfica do Rio Cachoeira, localizados na cidade de Joinville, Santa Catarina, Brasil (Figura 1). Esta bacia está geograficamente localizada em planícies costeiras de baixa altitude, sofrendo significativa influência de maré (Oliveira et al., 2017), além da proximidade com a Baía da Babitonga.



Figura 1- Local de realização dos estudos

Fonte: Google Maps, 2021.

As amostragens ocorreram em 19 pontos (Quadro 1). A numeração de cada ponto está relacionada à cronologia de início do monitoramento, onde o alvo de amostragem 1 foi o primeiro a ser avaliado, enquanto o ponto 19 o último a passar por amostragens e análise. A escolha dos alvos de amostragem é estratégica e definida por setores internos da Companhia Águas de Joinville (Companhia Águas de Joinville, 2021).

Alguns pontos já possuem rede coletora de esgoto, enquanto outros ainda recebem efluentes sanitários. Os dados das análises e as informações relacionadas à presença de rede coletora de esgoto foram obtidos a partir do banco de dados da companhia de saneamento do município.

Quadro 1 - Pontos de amostragem e respectivos níveis de cobertura de rede coletora de esgoto.

Identificação	Localização (Latitude / Longitude)	Rede Coletora de Esgoto (%)
Ponto 01	26°18'32.9"S 48°50'22.9"W	89,0 (desde 2014)
Ponto 02	26°18'19.6"S 48°50'22.8"W	100,0 (desde 2014)
Ponto 03	26°17'29.3"S 48°50'43.9"W	96,0 (desde 2014)
Ponto 04	26°17'06.2"S 48°50'24.6"W	100,0 (desde 2014)
Ponto 05	26°16'49.4"S 48°50'12.5"W	94,0 (desde 2014)
Ponto 06	26°16'23.2"S 48°50'00.9"W	99,0 (desde 2014)
Ponto 07	26°16'28.4"S 48°50'58.4"W	97,0 (desde 2014)
Ponto 08	26°16'20.2"S 48°50'57.3"W	95,0 (desde 2014)
Ponto 09	26°16'14.2"S 48°51'19.5"W	89,0 (desde 2014)
Ponto 10	26°16'04.6"S 48°51'38.0"W	92,0 (desde 2014)
Ponto 11	26°16'07.2"S 48°52'37.4"W	93,0 (desde 2014)
Ponto 12	26°19'55.3"S 48°49'21.7"W	2,0 (desde 2014)
Ponto 13	26°19'58.0"S 48°49'09.3"W	0,0
Ponto 14	26°22'21.3"S 48°49'31.8"W	0,0
Ponto 15	26°20'45.5"S 48°48'40.2"W	0,0
Ponto 16	26°19'20.7"S 48°50'19.3"W	46,0 (desde 2014)
Ponto 17	26°18'13.5"S 48°51'12.6"W	99,0 (desde 2014)
Ponto 18	26°16'38.8"S 48°50'54.8"W	100,0 (desde 2014)
Ponto 19	26°18'50.2"S 48°49'41.8"W	78,0 (desde 2018)

Fonte: Companhia de Águas de Joinville (2021).

Amostragem e parâmetros analíticos

As amostragens e os ensaios foram efetuados pelo Laboratório de Controle de Qualidade da Companhia Águas de Joinville (LCQ), reconhecido junto ao Instituto do Meio Ambiente de Santa Catarina, Brasil (IMA) e ocorreram em momento de maré vazante. Essa condição se aplica aos pontos 01, 06 e, principalmente 19, que sofrem intrusão da maré.

Os procedimentos de amostragem e preservação de amostras foram executados conforme os métodos padrões para análise de água e efluentes (American Public Health Association, 2017). Foram

coletados volumes aproximados de dois litros de amostras de água superficial para realização dos ensaios que compõem o IQA (Quadro 2).

Devido à instabilidade de alguns parâmetros e a presença de interferentes como algas, as leituras de oxigênio dissolvido (OD), pH e temperatura foram realizadas diretamente, em campo, nos primeiros 15 minutos após a amostragem, conforme preconizado pelo *Standard Methods For The Examination Of Water And Wastewater* (American Public Health Association, 2017).

Todos os equipamentos utilizados foram previamente ajustados e inspecionados por meio da aplicação de materiais de referência certificados (padrões) e verificados com o auxílio de brancos, além de verificações intermediárias com padrões de concentração.

Quadro 2 - Ensaios analíticos de composição do IQA.

Parâmetro	Limite de Quantificação	Referência
Coliformes	1,8 NMP/100ml	SMEWW 9223
Oxigênio dissolvido (OD)	0,4 mg/l O ₂	SMEWW - 4500-O G
pH	Não Aplicável	SMEWW - 4500-H+ B
Temperatura	0,1 °C	SMEWW - 2550
Demanda Bioquímica e Oxigênio (DBO)	4,0 mg/l O ₂	SMEWW - 5210D
Sólidos totais	15,0 mg/l	SMEWW - 2540B e 2540E
Turbidez	1,0 NTU	SMEWW - 2130B
Fósforo total	0,03 mg/l P-PO ₄	SMEWW - 4500-P C
Nitrogênio total	1,8 mg/l N	SMEWW - 4500N-C

Fonte: Autor

Tratamento de dados

Os dados das análises foram provenientes de 30 campanhas com início no ano de 2014 (mínimo quatro amostragens no ano, distribuídas ao longo das quatro estações) e término em 2020 (mensal desde 2019). Valores considerados discrepantes (*outliers*) foram excluídos.

Por se tratar de variáveis numéricas, resultado de medições por instrumentos, e possuírem casas decimais, pôde-se classificar os dados como quantitativos contínuos. Estes valores, provenientes das medições, foram agrupados em planilhas eletrônicas para composição do banco de dados.

Os resultados dos parâmetros medidos nos 19 pontos monitorados foram submetidos à análise estatística descritiva, estudo de correlação, análise de componentes principais, teste de agrupamento e regressão linear múltipla (Figura 2).

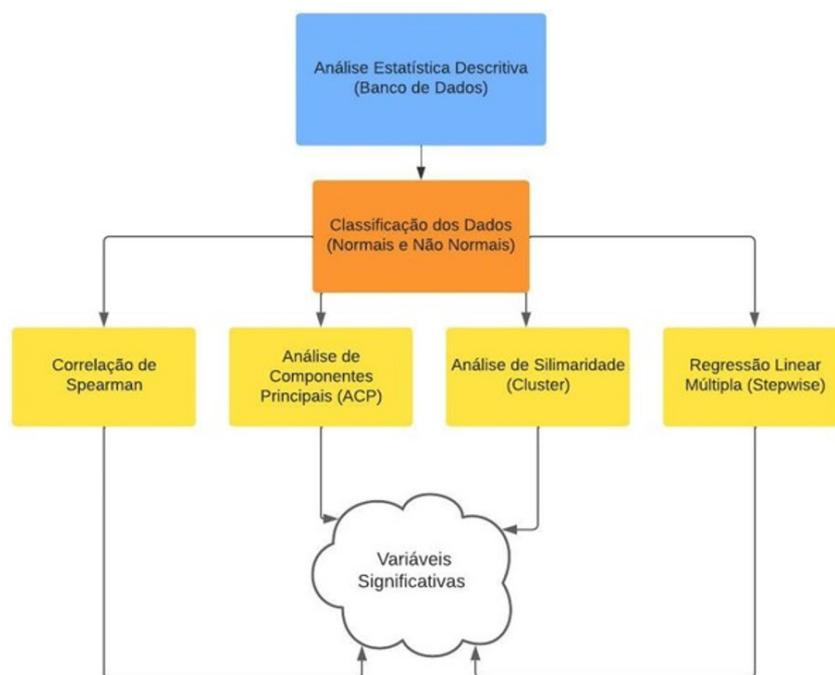


Figura 2 - Roteiro estatístico para definição de variáveis significativas.

Fonte: Autor

A análise estatística descritiva, correlações e regressão linear com as variáveis definidas foram realizadas com o programa Jamovi (Jamovi, 2021).

Para a realização da análise de componentes principais, classificação por similaridade (*cluster*) e regressão linear múltipla (*stepwise*) foi utilizado o programa R (RStudio Team, 2020).

Análise estatística

A análise estatística descritiva, incluindo aderência à normalidade, foi executada por meio do programa Jamovi, utilizando os seguintes parâmetros: Módulo de análises > exploração > estatística descritiva. Além da inclusão dos dados e seleção dos ensaios constituintes do IQA, foram selecionadas as seguintes opções: dimensão da amostra (n e omissão), tendência central (média e mediana), dispersão (desvio padrão, mínimo e máximo), normalidade (Shapiro-Wilk) e dispersão média (intervalo de confiança para a média de 95%).

Os testes de correlação foram executados também por meio do programa Jamovi, utilizando os seguintes parâmetros: módulo de análises > regressão > matriz de correlações. Após a inclusão dos dados, foram selecionadas as seguintes opções: coeficientes de correlação (Spearman, em função da normalidade obtida), hipóteses (correlacionado) e opções adicionais (relatar significância, sinalizar correlações significativas e intervalo de confiança em 95%).

A análise de componentes principais (ACP) foi executada no programa Rstudio por meio da função *Factoshiny* e *Factomine R*. Em meio às opções, a ACP foi selecionada onde 41 gráficos e descrição dos eixos foram obtidos.

Os testes de agrupamento também foram obtidos com o uso do programa Rstudio e ainda por meio da função *Factoshiny* e *Factomine R*. Em opções, a análise de *clustering* foi selecionada, onde gráficos e variáveis significativas foram obtidos.

A regressão linear múltipla foi realizada por meio da função *Stepwise R* no programa *Rstudio*. Importante ressaltar a necessidade de tratamento dos resultados em busca do modelo matemático que apresentar o menor *índice de Akaike*.

RESULTADOS E DISCUSSÃO

Análise estatística descritiva

Resumir uma série de valores de mesma natureza por meio de um conjunto de ferramentas e técnicas é considerada etapa fundamental para o início de qualquer análise estatística. Verificar mínimos e máximos utilizando ferramentas como o *Boxplot* (diagrama de caixas) auxiliam para evitar que erros sejam utilizados de maneira equivocada, invalidando todo o trabalho. Inclusive, uma análise estatística bem executada permite a retirada de informações atípicas, também conhecidas como *outliers* (Mancuso et al., 2018).

A avaliação da normalidade dos dados é um pré-requisito para muitos testes estatísticos porque os dados normais são uma suposição subjacente em testes paramétricos. Estes testes têm a vantagem de fazer um julgamento objetivo da normalidade, porém nem sempre são suficientemente sensíveis em bancos de dados com número muito pequeno de amostras e/ou excessivamente sensíveis a tamanhos de amostra grandes. Este tratamento prévio dos dados evita que estudos com diferentes normalidades sejam comparados. Dentre as diversas maneiras para o cálculo, as principais são os testes de Shapiro Wilk (amostras com até 2000 resultados) e Kolmogorov-Smirnov (amostras com mais de 2000 resultados) (Orcan, 2020).

Assim, o banco de dados foi introduzido no programa R para cálculo de estatística descritiva e realização do teste de Shapiro-Wilk para determinação da distribuição. Os valores medidos dos nove parâmetros analisados apresentaram tanto distribuição normal como não normal ao longo dos 19 pontos monitorados. O ensaio de OD apresentou distribuição normal em 100% das observações realizadas, seguido de temperatura (94,7%), pH (78,9%) e nitrogênio total (57,9%). Por outro lado, os parâmetros fósforo (42,1%), sólidos totais (31,6%), DBO (21,1%), turbidez (15,8%) e coliformes (5,3%) apresentaram distribuição não normal na maioria das observações.

Não foi identificado comportamento que justificasse mudança na distribuição em função da presença ou ausência de rede coletora de esgoto na região. O ensaio de fósforo teve alteração na distribuição onde se apresentava como normal em rios com cobertura de rede coletora inferior a 50%. Ainda assim, em outros três rios com rede coletora, houve classificação dos dados como não normal.

Correlação entre os parâmetros

Os coeficientes de correlação são ferramentas amplamente utilizadas para analisar diferentes tipos de associação entre variáveis em indivíduos pertencentes a populações em estudo. Eles são encontrados em praticamente todas as áreas do conhecimento empírico, dentre as mais utilizadas, destacam-se a de Pearson, definida para dados numéricos, e a de Spearman, para dados ordinais. Enquanto o método de Pearson avalia as relações lineares para a covariância de variáveis normalizadas (distribuição normal), a correlação de Spearman avalia as relações monotônicas de medidas não paramétricas (Makowski et al., 2020).

A análise baseada no cálculo de um coeficiente de correlação é apenas a primeira etapa de um estudo correlacional. Esta análise é complementada posteriormente com modelos de regressão, onde uma das variáveis é considerada explicativa e a outra como explicada. A ausência de argumentos que sustentem essas exigências limita a interpretação dos resultados, principalmente se agregar estudos de regressão (Pinilla & Rico, 2021).

Para verificação da significância das variáveis, matrizes de correlação foram geradas por meio do programa Jamovi para verificar se houve associação entre estas e, quando existentes, qual a intensidade destas correlações. Os valores obtidos passaram pelo teste de esfericidade (Bartlett), apresentando conformidade para todos os pontos ao apresentar valor de $p < 0,05$ (Cecconello et al., 2018). Neste caso, o teste indica que há relação entre os pontos estudados (Drose et al., 2020).

Dos 162 apontamentos de correlação, 80 foram listadas como fracas, 52 como moderadas e 30 correlações como fortes. Em 17 dos 19 rios monitorados, o maior nível de incidência ocorreu entre a relação “fósforo x sólidos”. Relações entre “coliformes x fósforo”, “fósforo x nitrogênio”, “DBO x nitrogênio” e “DBO x fósforo” apresentaram elevado índice de ocorrência em mais da metade dos alvos monitorados. As análises de DBO, nitrogênio total e fósforo foram mencionadas em oito correlações cada (incluindo entre si).

Foram classificadas como fortes e moderadas 82 observações. As relações “fósforo x sólidos” e “coliformes x fósforo” mantiveram-se como as mais representativas por terem sido indicadas em mais da metade dos pontos monitorados. Outras relações mencionadas acima como “DBO x nitrogênio” e “DBO x Fósforo” foram mantidas, mas todas em menor número de incidência.

Com a retirada das correlações fracas, somente o ponto 19 não apresentou nenhuma correlação entre as variáveis analisadas. De todos os pontos monitorados, este é o ponto sujeito a maior influência de intrusão de maré, o que pode explicar tal comportamento. As análises de sólidos, DBO e nitrogênio foram mencionadas em seis correlações (incluindo entre si), seguido do parâmetro coliformes com cinco menções. Fósforo, OD e pH foram correlacionadas em quatro casos e, temperatura e turbidez com três e duas menções, respectivamente.

Considerando somente correlações fortes, houve redução para 30 ocorrências. Destas, “fósforo x sólidos”, “coliformes x fósforo”, “DBO x nitrogênio” e “DBO x fósforo” foram mantidas, porém, em número menor de incidência. Com a retirada de correlações moderadas, somente 12 pontos apresentaram correlações entre as variáveis analisadas. A análise de sólidos foi mencionada em cinco correlações. Os ensaios de fósforo e DBO foram mencionados em quatro casos (incluindo em correlação entre si), seguido de nitrogênio, coliformes e pH com três menções. Oxigênio dissolvido apresentou correlação somente em uma ocasião (com pH).

Quando todas as correlações foram mantidas, pouca variação foi detectada, sendo que os somatórios se mantiveram entre seis e oito ocorrências por ensaio. Com a exclusão de correlações fracas, observou-se um destaque para DBO, nitrogênio total, sólidos e coliformes. Fósforo, OD e pH mantiveram número significativo de menções. Mantendo somente correlações fortes, verificou-se que OD apresentou somente uma menção, enquanto temperatura e turbidez não foram citadas em correlações.

Apesar do grande número de correlações apresentadas (162), verificou-se que variáveis com fraca força de relação pouco contribuíram no sentido de sinalizar variáveis representativas. Da mesma forma, considerar somente correlações classificadas como fortes acabaram segregando alguns pontos, não permitindo uma análise adequada dos alvos desse estudo.

Em um trabalho de aplicação de um índice de qualidade específico para o Rio Paraíba do Sul (curso de água que banha os estados de São Paulo, Rio de Janeiro e Minas Gerais), na região sudeste do Brasil, foi apontada relação entre os parâmetros comparados. Os impactos resultantes do lançamento de carga orgânica foram atribuídos aos parâmetros fósforo, nitrogênio amoniacal (fração da série nitrogenada), OD e DBO. Além disso, a presença de organismos como coliformes mostrou-se um importante indicador de contaminação por fezes de animais de sangue quente (Pessoa et al., 2020).

Desta forma, optou-se por considerar como válidas todas as correlações classificadas como fortes e moderadas. Para fins de comparação com a análise de componentes principais, foram consideradas

significativas as variáveis DBO, nitrogênio total, sólidos totais, coliformes e fósforo total. A escolha das variáveis corrobora o papel desempenhado por elas, tanto na presença de correlações fracas como fortes.

Análise de componentes principais (ACP)

A análise de componentes principais (ACP), além de permitir a redução do número de variáveis, tem a vantagem de facilitar a extração de informações relevantes na avaliação da qualidade da água. Com a aplicação desta ferramenta, um pequeno número de variáveis tende a conter informações mais relevantes, enquanto as variáveis restantes adicionam pouco ou nada à interpretação dos dados, em termos qualitativos (Drose et al., 2020).

Esta ferramenta avalia múltiplos dados, reduzindo as informações em dados mais relevantes, sem perda de informações responsáveis pelas maiores variações, denominadas de componentes principais (Cândido et al., 2021). Esta ferramenta busca uma relação natural de independência ou dependência, entre as variáveis (Souza & Santos, 2018).

A ACP vem sendo utilizada na identificação de condições críticas para monitoramento de qualidade de água e é uma das estratégias mais comuns para análise exploratória por facilitar a visualização e o entendimento de problemas com multivariáveis (Pinto et al., 2021).

Além da ACP, outras técnicas como correlações e agrupamentos são utilizadas para auxiliar na análise de grupos com abundância de informações. A ACP é uma técnica estatística que transforma linearmente um conjunto original de variáveis, inicialmente correlacionadas entre si, num conjunto menor de variáveis não correlacionadas que contém a maioria da informação em relação ao conjunto original. Esta técnica está associada à ideia de redução de massa de dados, com menor perda possível da informação. Ela procura explicar a estrutura da variância e covariância de um vetor aleatório, composto por variáveis aleatórias, por meio de combinações lineares dos dados originais. Essas combinações lineares são chamadas de componentes principais e não são correlacionadas entre si (Hongyu et al., 2016).

Desta forma, com o propósito de reduzir o número de variáveis, com a menor perda possível de informações, realizou-se a análise de componentes principais (Figura 3).

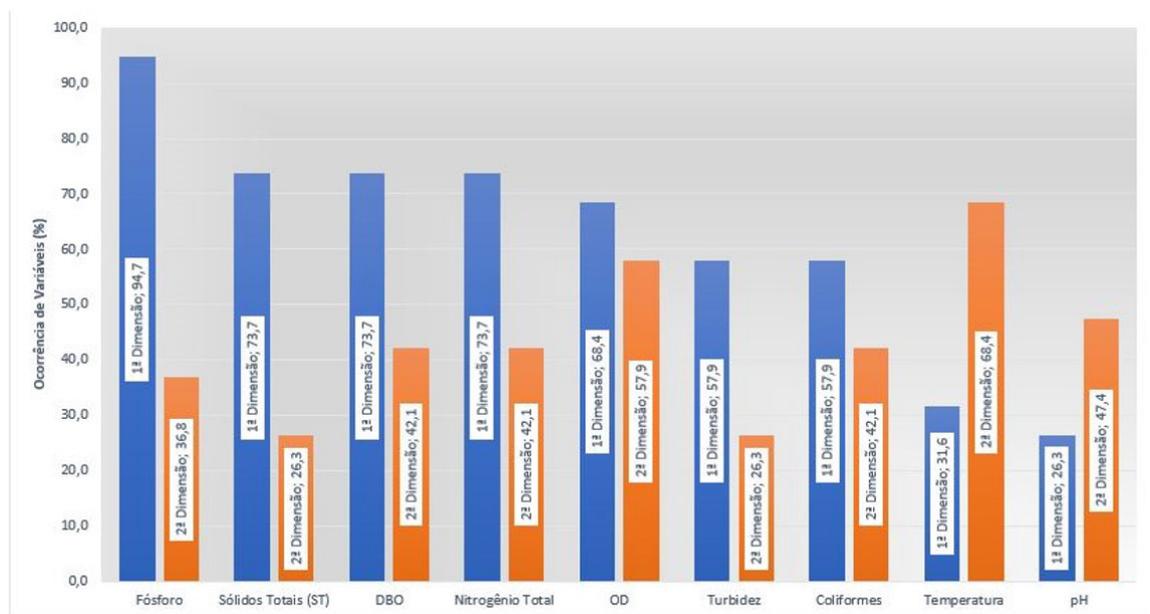


Figura 3- Ocorrência de variáveis em primeira e segunda dimensões (ACP).

Fonte: Autor

Com a obtenção dos parâmetros significativos ($p < 0,05$), verificou-se que fósforo apareceu como um parâmetro significativo em 18 dos 19 rios analisados (94,7%), enquanto DBO, nitrogênio total e sólidos totais foram mencionados em 14 dos 19 rios (73,7%). Oxigênio dissolvido (68,4%), coliformes (57,9%) e turbidez (57,9%) também foram citados como itens significativos em mais da metade dos alvos amostrados. Temperatura e pH, apesar de citados, possuíam baixa ocorrência, sendo mencionados apenas em seis (31,6%) e cinco (26,3%) pontos, respectivamente.

Os parâmetros significativos, ao longo dos pontos monitorados, foram responsáveis por 74,5% da variância total no ponto 19 e 100% no ponto 14. Não foi identificada ligação entre os parâmetros significativos e os valores de rede coletora de esgoto presentes nas regiões monitoradas. Somente o ensaio de sólidos totais apresentou indicação de menor significância em rios com cobertura de esgoto superior aos 90%. Mesmo assim, houve pontos que apresentaram significância com tal porcentagem de cobertura.

Em segunda dimensão, o resumo dos parâmetros mais citados ao longo dos pontos monitorados mostrou que temperatura (68,4%) e OD (57,9%) foram os parâmetros com maior ocorrência, aparecendo em mais da metade dos alvos. Na sequência, aqueles que apresentaram participação em menos da metade dos rios avaliados, com destaque para pH, citado com nove (47,4%) ocorrências, seguido de coliformes, DBO, nitrogênio total, fósforo, turbidez e sólidos.

O percentual de explicação da variância total dos parâmetros sinalizados variou de 79,0% no ponto 17 para 97,7% no ponto 1. Os ensaios variaram ao longo dos pontos, não havendo correlação entre estes. Não foi identificada ligação entre os parâmetros significativos e os valores de rede coletora de esgoto presente nas regiões monitoradas.

Assim como para os testes de correlação apontados, o estudo de componentes principais indicou como significativas as variáveis DBO, nitrogênio total, sólidos totais, coliformes e fósforo total. Oxigênio dissolvido e turbidez também foram citados e merecem atenção por estarem, assim como os demais parâmetros citados, presentes em mais da metade dos pontos monitorados. O parâmetro OD deve ser analisado com cautela, uma vez que sua leitura é inversamente proporcional aos demais parâmetros, para fins de qualidade da água.

O aumento dos valores apresentados para os parâmetros analisados sugeriu que a carga poluidora estaria aumentando, porém os valores de OD também aumentaram. Essa informação gera incoerência pois, ao contrário dos outros parâmetros, o aumento de OD indicaria que a carga poluidora estaria diminuindo, salvo interferências como intrusão de maré, chuvas ou amostragem em pontos com alta turbulência (Santos et al., 2020).

O aumento de OD associado a um aumento do pH poderia sugerir pontualmente o início de um processo de eutrofização. Entretanto, este trabalho avaliou amostras de água em um período aproximado de sete anos (2014 a 2021) e não observou alterações consideráveis nos valores de pH.

Em um estudo que objetivou desenvolver um índice de qualidade de água modificado para o Arroio Pelotas (Rio Grande do Sul, Brasil), por meio de análises multivariadas, a ACP extraiu duas componentes principais que explicaram 100% da variância total dos dados. O estudo apontou DBO, turbidez e sólidos totais como as variáveis mais importantes na construção do índice. Foi indicada diferença para as variáveis mais significativas do IQA CETESB, principalmente em função dos pesos aplicados às variáveis. Coliformes, fósforo, nitrogênio, OD, pH e temperatura apresentaram pesos inferiores, enquanto DBO, turbidez e sólidos totais apresentaram desvios superiores aos utilizados pela CETESB. Por esta razão, foi sugerida uma readequação dos pesos para o ecossistema estudado (Ceconello et al., 2018).

Assim, optou-se por considerar como válidos os ensaios de DBO, nitrogênio total, sólidos totais, coliformes, fósforo total e turbidez. A escolha das variáveis corrobora o papel desempenhado por elas, tanto na presença de correlações fracas quanto na presença de correlações fortes. A variável temperatura é indicada como não significativa para fins práticos de monitoramento em função de pequena variação que não seja sazonal (Rabelo et al., 2020).

Análise de similaridade (*cluster*)

Além da ACP, outra ferramenta que busca obter a maior homogeneidade possível, dentro de grupos de informações e, por conseguinte, alta heterogeneidade entre os grupos formados, é a análise de agrupamento. Esta técnica, que também é conhecida como *cluster* ou análise hierárquica de agrupamentos, é utilizada com o propósito de agrupar similaridades, ou seja, semelhanças entre variáveis por proximidade (Souza et al., 2021).

Testes de agrupamento são utilizados para reunir, por algum critério de classificação, os objetos ou parcelas de uma amostra da população estudada. A técnica apresenta uma série de benefícios como a rápida convergência das informações. Isto porque o número máximo de interações equivalerá, no pior caso, ao número de observações já existentes. Além disso, a determinação do número de grupos, independe do pesquisador, não permite influência da subjetividade. Outra vantagem está no fato do teste fornecer sempre as mesmas soluções, sendo, portanto, determinístico. Assim, para uma mesma massa de dados, a técnica só necessita ser executada uma vez, ou seja, fornecerá apenas um resultado

de agrupamento. Cabe ressaltar que a técnica apresenta algumas limitações como a influência da presença de *outliers* na qualidade das soluções geradas (Albuquerque & Barros, 2020).

A busca por associação entre as variáveis significativas por similaridade revelou que coliformes apareceu em 100% das observações como variável significativa ($p < 0,05$). Com menor incidência, fósforo, DBO e turbidez ocorreram em sete dos 19 pontos monitorados. Nitrogênio total, sólidos totais e oxigênio dissolvido também compõem a lista de variáveis significativas com seis, cinco e quatro ocorrências, respectivamente.

Por outro lado, as análises de pH e temperatura não apresentaram grau de significância para similaridade nos alvos estudados, indicando que estes parâmetros deixam de ser opções válidas para uso no monitoramento dos rios utilizados neste estudo.

Dados de monitoramento foram aplicados em um estudo para verificar a variação espacial da qualidade de águas superficiais por meio de testes estatísticos como a análise de *cluster*. Foram analisados 14 parâmetros em 65 estações de monitoramento, no período de 2008 a 2016. As estações monitoradas pertencem à sub-bacia do Rio das Velhas, inserida na Bacia Hidrográfica do Rio São Francisco, Brasil. A Análise de *cluster* foi utilizada para identificar os principais parâmetros responsáveis pela diferença entre as estações estudadas. Foram formados dez grupos de estações, conforme a similaridade da qualidade da água, sendo justificados principalmente por pressões antrópicas similares. Coliformes, manganês e fósforo foram indicados como os parâmetros mais relevantes da bacia do rio das Velhas. Os resultados sugerem que os parâmetros considerados críticos seriam responsáveis pela degradação da qualidade da água e poderiam estar relacionados aos despejos de esgotos domésticos, de efluentes industriais, de atividades de mineração, de drenagem de áreas agrícolas e de processos de erosão e assoreamento (Soares & Pinto, 2017).

Para selecionar as características físico-químicas mais importantes para explicar a variabilidade da qualidade das águas de uma sub-bacia hidrográfica rural no Sul do Estado do Espírito Santo (Brasil), empregaram-se técnicas de estatística multivariada. Com o uso de ACP e agrupamento hierárquico, foram considerados como parâmetros mais representativos da variabilidade da qualidade das águas superficiais e subterrâneas (sub-bacia hidrográfica do Córrego Horizonte) sólidos totais, turbidez, DBO, entre outros (não aplicáveis ao IQA). Importante ressaltar que os períodos de coleta (chuva e estiagem) não influenciaram na segregação dos grupos formados (Cecílio et al., 2013).

Regressão linear múltipla

Desenvolvida na primeira década do século XIV para o cálculo de órbitas de corpos celestes, o método dos mínimos quadrados foi utilizado como base matemática para elaboração da análise de regressão. A regressão é uma classe de ferramentas estatísticas utilizada para quantificar as relações entre uma ou mais variáveis independentes e uma variável dependente. Com uso de programas computacionais e modelos estatísticos, essas relações são determinadas para abordar vários questionamentos em uma pesquisa (Pisicã et al., 2022).

Diferentemente da regressão linear simples (ou bivariada), usada para prever o valor de uma variável a partir de outra, a regressão linear múltipla é um método que utiliza mais de uma variável (Bazdaric et al., 2021).

Entre os três métodos mais utilizados como critério de seleção em testes de regressão, o critério de informação Akaike (AIC) foi proposto pela primeira vez como um método para comparar diferentes modelos sobre um determinado resultado. Como regra para escolha do melhor modelo matemático tem-se que quanto menor o valor de AIC, melhor será o modelo (Pho et al., 2019).

Portanto, torna-se necessária a realização de estudos que permitam avaliar a qualidade de rios com o menor número possível de variáveis, mantendo resultados de fácil interpretação, com redução de custos e tempo para comunicação dos resultados.

Na análise de regressão, a preocupação é sempre com a dependência estatística entre variáveis. Trabalha-se com variáveis aleatórias, com uma distribuição de probabilidade. Em resumo, estes modelos são utilizados para analisar, qualitativamente e quantitativamente, relações entre variáveis. Chama-se de variável dependente ou variável endógena aquela cujo comportamento será explicado pela variável chamada de variável explicativa. Assim, pode-se considerar que a ideia do modelo de regressão linear é estimar uma reta que melhor descreva a relação entre variáveis (Chein, 2019).

A regressão é uma classe de ferramentas estatísticas para quantificar as relações entre uma ou mais variáveis independentes e uma variável dependente. Com uso de programas computacionais e modelos estatísticos, essas relações são determinadas para abordar vários questionamentos em uma pesquisa (Pisicã et al., 2022). Diferentemente da regressão linear simples (ou bivariada), que é usada

para prever o valor de uma variável a partir de outra, a regressão linear múltipla é um método que usa mais de uma variável (Bazdaric et al., 2021).

O banco de dados de cada um dos pontos de amostragem foi submetido ao teste de *stepwise* (regressão linear múltipla) em busca de modelos matemáticos que pudessem explicar o comportamento do IQA com o menor número possível de variáveis. Partindo-se de um conjunto composto pelos nove parâmetros originais (OD, pH, temperatura, coliformes, DBO, fósforo, nitrogênio, turbidez e sólidos totais) houve redução, proposta pelo teste, em todos os pontos avaliados.

Em média, os pontos monitorados apresentaram redução de nove variáveis iniciais para seis variáveis significativas. As menores reduções ocorreram nos rios 5, 6, 12, 18 e 19. Nestes pontos houve remoção somente de uma variável, mantendo-se oito parâmetros ainda como significativos. Temperatura foi excluída dos modelos matemáticos para os pontos 5, 6 e 19. Para os pontos 12 e 18, nitrogênio deixa de fazer parte dos modelos sugeridos.

As maiores reduções foram verificadas nos rios 4, 13 e 14. Os três modelos propostos mantiveram quatro variáveis como sendo significativas. A variável fósforo foi mantida para os três pontos, seguida de DBO e OD que foram mencionadas nos pontos 13 e 14. Para o rio 4, pH e nitrogênio foram mantidos para composição do modelo matemático. Para o ponto 13, além de OD, DBO e fósforo, a análise de sólidos totais foi mantida como significativa. Por fim, para o ponto 14, além de OD e fósforo, as variáveis temperatura e turbidez complementam o modelo matemático final.

De maneira geral, as variáveis OD (95%), coliformes (79%), DBO (89%), fósforo (95%), turbidez (68%) e sólidos (58%) são sugeridas como significativas em mais da metade dos rios verificados. Não foi verificada relação entre as variáveis mais significativas com a presença de rede coletora de esgoto ao longo dos rios monitorados.

Utilizando um modelo desenvolvido para auxiliar na avaliação rápida e de baixo custo da qualidade da água do Rio Tigre (Bagdá, Iraque), foram selecionadas entre 11 parâmetros as variáveis mais significativas após uso de regressão linear múltipla. Entre as variáveis apontadas, foram sugeridas as análises de turbidez, coliformes e pH (Ewaid et al., 2018).

Variáveis significativas

Com o intuito de selecionar os parâmetros significativos para o monitoramento dos rios estudados, vários testes estatísticos foram utilizados (Quadro 3).

Quadro 3 - Variáveis significativas por teste realizado.

Teste	Variáveis Significativas Apontadas Para Cada Teste Estatístico						
	Fósforo	Sólidos	DBO	Coliformes	Nitrogênio	-	-
Correlação	Fósforo	Sólidos	DBO	Coliformes	Nitrogênio	-	-
ACP	Fósforo	Sólidos	DBO	Coliformes	Nitrogênio	Turbidez	-
Cluster	Fósforo	Sólidos	DBO	Coliformes	Nitrogênio	Turbidez	OD
Stepwise	Fósforo	Sólidos	DBO	Coliformes	-	Turbidez	OD

Fonte: Autor

Em todos os testes realizados, as variáveis sólidos totais, fósforo, DBO e coliformes foram sugeridas como significativas para manutenção do monitoramento. Apesar das análises de nitrogênio, turbidez e OD também serem sugeridas, é importante considerar algumas ressalvas quanto à manutenção delas nos modelos matemáticos.

A manutenção da utilização da variável OD implica em maiores cuidados na amostragem para evitar possíveis interferências como a intrusão de maré, chuvas ou amostragem em pontos com alta turbulência. De acordo com Santos et al. (2020), oscilações como cheias dos rios e períodos chuvosos, podem alterar significativamente os valores de OD e, conseqüentemente, afetar o IQA, uma vez que este parâmetro corresponde a 17% do peso deste índice.

A análise de nitrogênio deixou de ser considerada no estudo de regressão linear múltipla por aparecer em menos da metade dos pontos avaliados. Ainda assim, é importante considerar que esta variável ocorre como significativa em 47% dos alvos de amostragem.

Para o teste de correlação, a análise de turbidez não foi apontada como significativa por possuir relação e frequência baixas. A única relação apontada ocorre com a análise de sólidos.

Das variáveis estudadas, DBO e sólidos totais estão de acordo com estudos que buscaram a redução de variáveis em índices de qualidade como o IQA modificado pelo CETESB (Ceconello et al., 2018).

Um novo teste de regressão linear considerando somente as variáveis sugeridas no Quadro 3 indicou que os ensaios de sólidos totais e nitrogênio, quando adicionadas ao modelo matemático, apresentaram valores de $p > 0,001$ (intervalo de confiança de 95%). Verificou-se que a manutenção destas variáveis, além de não contribuir para o modelo, também pouco contribuíram para a obtenção de um R e R² maiores. Desta forma, optou-se pela elaboração de um novo teste com as variáveis OD, DBO, fósforo, coliformes e turbidez.

No IQA, as variáveis sugeridas como significativas neste estudo (OD, coliformes, DBO, fósforo e turbidez) representam 60% do peso total do indicador (Brasil, 2004).

Validade do modelo adotado

Buscou-se a obtenção de um único modelo matemático que pudesse representar os pontos de amostragem, conforme expresso a seguir:

$$IQA = 36,5553 + (OD \times 2,6992) - (DBO \times 0,0522) - (P \times 9,2080) - (Col. \times 6,54^{-6}) - (TUR \times 0,0302) \quad (1)$$

O nível de assertividade médio apresentou capacidade de explicação de apenas 61,0% ao longo dos 19 pontos testados. Em função do baixo nível de assertividade, optou-se por gerar modelos matemáticos para cada um dos pontos analisados, conforme descrito no Quadro 4.

Quadro 4 – Modelos matemáticos, pesos e assertividade (%) gerados para cada um dos 19 pontos monitorados. P, fósforo total; COL., coliformes e TUR, turbidez.

Ponto	Modelo Matemático	Assertividade
1	$IQA = 40,9466 + (OD \times 2,8553) - (COL.x0,0000493) - (DBO \times 0,4554) - (P \times 6,5249) - (TUR \times 0,0629)$	83
2	$IQA = 42,365 - (COL.x0,000247) - (DBO \times 0,517) + (OD \times 3,660) - (P \times 13,939) - (TUR \times 0,104)$	89
3	$IQA = 43,477 - (DBO \times 0,4426) - (COL.x0,000025) + (OD \times 2,0542) - (P \times 6,9321) - (TUR \times 0,0562)$	77
4	$IQA = 45,761 - (COL.x0,00000335) - (DBO \times 0,446) - (P \times 5,512) + (OD \times 1,572) - (TUR \times 0,104)$	70
5	$IQA = 51,0223 - (DBO \times 0,2573) - (COL.x0,00000322) - (P \times 4,3115) - (TUR \times 0,0937) - (OD \times 0,0464)$	80
6	$IQA = 40,273 + (OD \times 3,143) - (COL.x0,0000171) - (DBO \times 0,321) - (P \times 9,796) - (TUR \times 0,211)$	70
7	$IQA = 42,19931 + (OD \times 2,59284) - (COL.x0,0000455) - (DBO \times 0,48464) - (P \times 12,06882) + (TUR \times 0,00468)$	80
8	$IQA = 42,5315 + (OD \times 2,5970) - (COL.x0,0000536) - (DBO \times 0,3985) - (P \times 5,1824) - (TUR \times 0,0534)$	80
9	$IQA = 39,1335 + (OD \times 3,2235) - (COL.x0,0000130) - (DBO \times 0,5651) - (P \times 5,6849) - (TUR \times 0,0769)$	62
10	$IQA = 37,6876 + (OD \times 2,9377) + (COL.x0,0000131) - (DBO \times 0,4726) - (P \times 6,3744) - (TUR \times 0,0772)$	80
11	$IQA = 39,1823 + (OD \times 2,5671) - (COL.x0,000024) - (DBO \times 0,3868) - (P \times 8,0049) + (TUR \times 0,0173)$	73
12	$IQA = 28,71508 + (OD \times 3,51345) - (COL.x0,00000864) - (DBO \times 0,16839) - (P \times 3,79709) - (TUR \times 0,00105)$	88
13	$IQA = 37,0498 + (OD \times 2,7619) - (COL.x0,00000842) - (DBO \times 0,0923) - (P \times 9,0613) - (TUR \times 0,0383)$	73
14	$IQA = 29,9583 + (OD \times 3,6583) - (COL.x0,0000739) - (DBO \times 0,0174) - (P \times 6,0711) - (TUR \times 0,0389)$	78
15	$IQA = 30,5773 + (OD \times 3,0247) - (COL.x0,0000782) - (DBO \times 0,0306) - (P \times 2,8049) - (TUR \times 0,0493)$	72
16	$IQA = 39,2673 + (OD \times 3,5892) - (COL.x0,000240) - (DBO \times 0,3462) - (P \times 5,9270) - (TUR \times 0,0861)$	94
17	$IQA = 57,6148 - (OD \times 0,0525) - (COL.x0,000464) - (DBO \times 0,3222) - (P \times 17,4648) - (TUR \times 0,0111)$	89
18	$IQA = 44,400 + (OD \times 3,316) - (COL.x0,000419) - (DBO \times 0,161) - (P \times 27,101) - (TUR \times 0,227)$	94
19	$IQA = 40,1287 + (OD \times 3,1475) - (COL.x0,000356) - (DBO \times 0,2478) - (P \times 13,8322) - (TUR \times 0,0392)$	100

Fonte: Autor

Inicialmente, os índices gerados foram comparados na forma de desvios, em que os valores obtidos no cálculo original do IQA foram comparados aos valores obtidos pelos modelos matemáticos. Posteriormente, esses valores foram classificados conforme suas capacidades de explicação, ou seja, os resultados obtidos tanto pelo cálculo do IQA como pelos modelos matemáticos foram classificados em péssimo, ruim, regular, bom e ótimo. A calibração para cada um dos 19 pontos foi realizada para validar os modelos propostos (dados não mostrados).

A utilização das variáveis coliformes, DBO, OD, fósforo e turbidez produziram modelos matemáticos com assertividade média de 81,0%. Destes, o ponto com pior capacidade de explicação ocorreu no ponto 9 com 62,0%, enquanto o melhor desempenho ocorreu no ponto 19 com 100,0% de explicação. O desvio apresentado para o ponto 9 foi verificado e a causa provável para esta diferença está atrelada à diferença entre os pesos adotados para IQA e para o modelo proposto para o alvo de amostragem em questão. As variáveis coliformes e fósforo apresentam peso de 15,0% e 10,0%, respectivamente, para o IQA. No modelo matemático proposto, essa ordem se inverte. O parâmetro fósforo apresenta peso superior ao valor de coliformes. Os pesos para cada uma das variáveis estão apresentados no Quadro 4, conforme modelo matemático calculado para cada um dos pontos monitorados.

Neste caso, com a diminuição de coliformes e aumento de fósforo na penúltima campanha de amostragem, o IQA indica valor de 60 enquanto o novo modelo proposto indica somente 43. Novamente, atenção especial é dada para o ensaio de OD. Na maioria dos modelos matemáticos gerados, há sinalização positiva (incremento gerado pela variável), em consonância com os demais parâmetros considerados significativos. Esta incoerência sugere a retirada de OD, mas por se tratar de modelos matemáticos com boa capacidade de explicação dos dados, optou-se por não retirar do ensaio, principalmente pelo baixo impacto causado pela variável no resultado.

Um estudo realizado no Arroio Pelotas (Rio Grande do Sul, Brasil) teve como objetivo desenvolver um índice para comparar com o IQA CETESB, entretanto, divergência entre as variáveis significativas apontadas foi verificada em função do peso obtido para cada uma delas. As constatações evidenciaram necessidade de calibração entre os modelos, principalmente na readequação dos pesos para cada manancial. A proposta teve o intuito de obter variáveis mais importantes e pesos que permitiam representar as reais características de cada local (Ceconello et al., 2018).

Independente do modelo matemático a ser adotado, ficou evidente que os pesos dos parâmetros utilizados possuem variações nos pesos de cada ensaio e que estes não se aproximam dos pesos definidos no cálculo do IQA. Foi evidenciado que os pontos possuem particularidades quanto às suas características e usos.

Um estudo conduzido na região central do país (Distrito Federal, Brasil) indicou, entre outras análises, que OD, coliformes e fósforo podem ser utilizadas como variáveis significativas. Ao longo de três rios avaliados (Rio Jardim, Córrego Sobradinho e Córrego Sarandi), foi verificado que a aplicação de indicadores de forma regionalizada poderá auxiliar no gerenciamento dos recursos hídricos (Passos et al., 2019).

Redução de custos

Os custos para realização de um plano de monitoramento contemplam força de trabalho para coleta das amostras, realização de análises em campo, preservação das amostras e transporte até o laboratório para execução dos demais ensaios. Apesar da facilidade para execução de ensaios como pH e temperatura, deve-se contemplar que os custos relacionados (manutenção, calibração, materiais de referência certificados, tempo de analista, por exemplo) precisam ser contemplados em um plano de monitoramento, seja para pesquisa ou para iniciativa privada (empresas de saneamento).

Um levantamento dos gastos envolvidos nesse processo indicou um total de R\$ 151,58 ou US\$ 27,58 (1 USD = 5,50 BRL) por amostra. Considerando que foram 19 pontos de análise por campanha por mês, tem-se um gasto mensal de R\$ 2.880,03 (US\$ 523,96). O custo anual para realização do monitoramento ao longo de um ano demanda um total aproximado de R\$ 34.560,37 (US\$ 6.287,51).

Com a utilização apenas das variáveis propostas neste trabalho, o plano de monitoramento passou a ser realizado com um custo aproximado de R\$ 92,26 (US\$ 16,78) por ponto, R\$ 1.752,99 (US\$ 318,92) para 19 pontos por mês e R\$ 21.035,87 (US\$ 3.827,02) por ano. A redução chega a 40%, ou seja, para as 30 campanhas realizadas, haveria a redução dos valores estimados para R\$ 33.811,26 (US\$ 6.151,22). Com a proposta, observa-se possibilidade de economicidade do processo e, conseqüentemente, sustentabilidade para a manutenção de um monitoramento representativo.

É importante ressaltar que a redução de custos poderia ser ainda maior quando se parte de um plano de monitoramento que não contempla, desde o início, a compra de equipamentos, materiais e reagentes necessários para a realização das análises, que foram retiradas do plano de monitoramento deste trabalho.

CONCLUSÃO

O trabalho procurou, por meio de ferramentas estatísticas, sugerir variáveis significativas como condição mínima para o monitoramento e avaliação da qualidade da água em uma bacia hidrográfica influenciada por intrusão de maré no sul do Brasil.

O teste de associação de Spearman indicou a presença de correlações fortes e moderadas, sinalizando como significativas as variáveis DBO, nitrogênio total, sólidos totais, coliformes e fósforo total. Correlações fortes entre os ensaios de fósforo total e sólidos foram evidenciadas na maioria dos pontos analisados. Estes ensaios também foram elencados como significativos após a realização da análise de componentes principais.

Ensaio de agrupamento sugerem que coliformes fecais, DBO, fósforo total, sólidos totais, nitrogênio total e turbidez compõem o conjunto de variáveis representativas e que possuem maior significância para realização de monitoramento nos corpos hídricos estudados neste trabalho.

Após utilização de regressão linear foram mantidas como significativas cinco variáveis: OD, DBO, coliformes, fósforo e turbidez.

Os dados gerados ao longo do estudo evidenciaram que, mesmo com as diferenças entre os alvos amostrados, é possível realizar um monitoramento de maneira eficaz, com redução de tempo e principalmente de custos. Além destes parâmetros permitirem uma avaliação do nível de poluição antrópica, como despejo de efluentes, o uso destas variáveis permitirá acompanhar a evolução de rios cuja cobertura de saneamento básico ainda é considerada baixa.

O estudo possui relevância por considerar dados de pontos de amostragem com características totalmente distintas como vazão, velocidade de escoamento da água, declividade, presença de áreas urbanizadas e intrusão de maré. Além disso, é significativo por ter sido realizado com dados obtidos em diferentes anos, levando em consideração características como sazonalidade, por exemplo.

Como perspectiva de estudos futuros, sugere-se a realização de experimentos semelhantes buscando a elaboração de modelos que auxiliem de maneira mais assertiva, com tempo e custo reduzidos, no monitoramento de outros compartimentos, como sedimentos, além de investigações para avaliação da eficiência do tratamento de efluentes domésticos.

AGRADECIMENTOS

Agradecemos à Companhia Águas de Joinville e ao Instituto Federal Catarinense (IFC-Araquari) pelo apoio cedido para realização desta pesquisa.

REFERÊNCIAS

- Abbasi, T., & Abbasi, S. A. (2012). *Water quality indices*. Amsterdã: Elsevier. <https://doi.org/10.1016/C2010-0-69472-7>.
- Albuquerque, M. A., & Barros, K. N. N. O. (2020). Determining the number of groups in cluster analysis via radius of influence. *Brazilian Journal of Development*, 6(6), 38342-38355. Recuperado em 23 de junho de 2023, de <https://ojs.brazilianjournals.com.br/ojs/index.php/BRJD/article/view/11795>
- American Public Health Association – APHA. American Water Works Association. Water Environment Federation. (2017). *Standard methods for the examination of water and wastewater* (23. ed.). Washington, D.C.: American Public Health Association. Recuperado em 23 de junho de 2023, de <https://www.standardmethods.org/>
- Araújo, N. M. S., & Graças e Silva, M. (2021). A atual política de saneamento básico no contexto da gestão das águas no Brasil: apontamentos iniciais. In M. F. F. Silva (Org.), *Políticas públicas e mobilidade urbana* (pp. 13-28). Guarujá: Científica Digital. Recuperado em 23 de junho de 2023, de <http://www.editoracientifica.org/articles/code/201202601>
- Bazdaric, K., Sverko, D., Salaric, I., Martinovic, A., & Lucijanac, M. (2021). The ABC of linear regression analysis: what every author and editor should know. *European Science Editing*, 47, e63780. <http://dx.doi.org/10.3897/ese.2021.e63780>.
- Bilgin, A. (2015). An assessment of water quality in the Coruh Basin (Turkey) using multivariate statistical techniques. *Environmental Monitoring and Assessment*, 187, 721. <https://doi.org/10.1007/S10661-015-4904-9>.
- Brasil. Ministério do Desenvolvimento Regional. Agência Nacional de Águas e Saneamento Básico – ANA. (2004). *Indicadores de qualidade - Índice de Qualidade das Águas (IQA)*. Recuperado em 23 de junho de 2023, de <http://pnqa.ana.gov.br/indicadores-indice-aguas.aspx>

- Camargo, D. M., Souza, N. M., Pacheco, T. C. K. F., Alcântara, G. C., & Dota, E. M. (2017). Modelagem geoespacial para identificação de áreas vulneráveis ao contágio por doenças relacionadas à falta de saneamento: o caso da Região Metropolitana de Campinas. *Revista Brasileira de Cartografia*, 69(3), 561-573. Recuperado em 23 de junho de 2023, de <https://seer.ufu.br/index.php/revistabrasileiracartografia/article/view/44349>
- Cândido, T. R. S., Santos, A. P. C., Mendonça, R. S., Donato, M. E. A. A. L., Oliveira, V. B., Lemos, D. G. L., Silva, I. L., & Santos, V. A. (2021). Aplicação da ferramenta estatística Análise de Componentes Principais (ACP) na identificação e análise da qualidade de estação de tratamento de água para hemodiálise. *Brazilian Journal of Development*, 7(12), 111171 – 111183. Recuperado em 23 de junho de 2023, de <https://www.brazilianjournals.com/index.php/BRJD/article/view/40588>
- Capellari, A., & Capellari, M. B. (2018). A água como bem jurídico, econômico e social. A necessidade de proteção das nascentes. *Cidades*, 36, 1-16. Recuperado em 23 de junho de 2023, de <http://journals.openedition.org/cidades/657>
- Castro, J. D. B., & Venâncio, V. (2019). Índice de qualidade da água: uma aplicação para o Centro-Oeste brasileiro. *Revista de Estudos Sociais*, 21(42), 122-137. Recuperado em 23 de junho de 2023, de <https://dialnet.unirioja.es/servlet/articulo?codigo=7055646>
- Ceconello, S. T., Centeno, L. N., Guedes, H. A. S., Ceconello, S. T., Centeno, L. N., & Guedes, H. A. S. (2018). Water quality index modified by using multivariate analysis: a case study of Pelotas Stream, RS, Brazil. *Engenharia Sanitaria e Ambiental*, 23(5), 973-978. Recuperado em 23 de junho de 2023, de http://www.scielo.br/scielo.php?script=sci_abstract&pid=S1413-41522018000500973&lng=en&nrm=iso&tng=pt
- Cecílio, R. A., Garcia, G. O., Neves, M. A., Bertossi, A. P. A., & Menezes, J. P. C. (2013). Selection and grouping of indicators of water quality using Multivariate Statistics. *Semina: Ciências Agrárias*, 34(5), 2025-2036. <https://doi.org/10.5433/1679-0359.2013v34n5p10862>.
- Chein, F. (2019). *Introdução aos modelos de regressão linear: um passo inicial para compreensão da econometria como uma ferramenta de avaliação de políticas públicas*. Brasília: Escola Nacional de Administração Pública. Recuperado em 23 de junho de 2023, de <http://repositorio.enap.gov.br/jspui/handle/1/4788>
- Comitê de Gerenciamento das Bacias Hidrográficas da Bacia da Babitonga e Bacias Contíguas – CCJ. (2016). *Avaliação do Índice de Qualidade da Água (IQA)*. Recuperado em 23 de junho de 2023, de <https://www.aguas.sc.gov.br/base-documental-rio-cubatao-norte/noticias-rio-cubatao-norte/item/3969-avaliacao-do-indice-de-qualidade-da-agua-iqa>
- Companhia Águas de Joinville. (2021). Recuperado em 23 de junho de 2023, de www.aguasdejoinville.com.br
- Drose, A., Valentini, M., Duarte, V., Santos, G. B., Nadaleti, W., & Vieira, B. M. (2020). Utilização de métodos estatísticos multivariados no monitoramento de qualidade da água da Lagoa Mirim. *Meio Ambiente*, 2(4), 58-67. Recuperado em 23 de junho de 2023, de <http://www.meioambientebrasil.com.br/index.php/MABRA/article/view/69>
- Ewaid, S. H., Abed, S. A., & Kadhum, S. A. (2018). Predicting the Tigris River water quality within Baghdad, Iraq by using water quality index and regression analysis. *Environmental Technology & Innovation*, 11, 390-398. <http://dx.doi.org/10.1016/j.eti.2018.06.013>.
- Hongyu, K., Sandanielo, V. L. M., & Oliveira Junior, G. J. (2016). Análise de componentes principais: resumo teórico, aplicação e interpretação. *Engineering and Science*, 5(1), 83-90. Recuperado em 23 de junho de 2023, de <https://periodicoscientificos.ufmt.br/ojs/index.php/eng/article/view/3398>
- Instituto Trata Brasil. (2021). *Ranking do saneamento 2021*. São Paulo: Instituto Trata Brasil. Recuperado em 23 de junho de 2023, de <https://tratabrasil.org.br/ranking-do-saneamento-2021/>
- Jamovi. (2021). *The Jamovi project [computer software]*. Recuperado em 23 de junho de 2023, de <https://www.jamovi.org>
- Makowski, D., Ben-Shachar, M., Patil, I., & Lüdecke, D. (2020). Methods and algorithms for correlation analysis in R. *The Journal of Open Source Software*, 5(51), 2306. <https://doi.org/10.21105/joss.02306>.
- Mancuso, A. C. B., Castro, S. M. J., Guimarães, L. S. P., Leotti, V. B., Hirakata, V. N., & Camey, S. A. (2018). Estatística descritiva: perguntas que você sempre quis fazer, mas nunca teve coragem. *Clinical and Biomedical Research*, 38(4), 414-418. Recuperado em 23 de junho de 2023, de <https://www.seer.ufrgs.br/index.php/hcpa/article/view/89242>.
- Molla, M. H., Chowdhury, M. A. T., Bhuiyan, M. H. R., Das, S., Morshed, A. J. M., Das, J., & Islam, S. (2022). Seasonal variation of drinking water quality in urban water bodies (UWBs) of Chittagong Metropolitan City, Bangladesh: implications of higher water quality index (WQI) for the urban environment. *Water Supply*, 22(5), 4934-4950. <http://dx.doi.org/10.2166/WS.2022.151>.
- Nong, X., Shao, D., Zhong, H., & Liang, J. (2020). Evaluation of water quality in the South-to-North Water Diversion Project of China using the water quality index (WQI) method. *Water Research*, 178, 115781. <http://dx.doi.org/10.1016/j.WATRES.2020.115781>.

- Oliveira, C. N. (2018). *Aplicação de índice de qualidade de água subterrânea em ambiente de percolados orgânicos na região de Camaçari – BA* (Tese de doutorado). Universidade Federal da Bahia, Salvador. Recuperado em 23 de junho de 2023, de <https://repositorio.ufba.br/handle/ri/28658>
- Oliveira, M., Hora, M., & Hora, A. (2020). Metodologia para estimativa do Índice de Qualidade das Águas. Estudo de caso dos rios tributários ao Reservatório de Juturnaíba, estado do Rio de Janeiro. *Anuário do Instituto de Geociências*, 43(3), 386-401. https://doi.org/10.11137/2020_3_386_401.
- Oliveira, T. M. N., Steffen, Á. H., Cidral, A., Garcia, B. R. Z., Mougá, D. M. D. S., Scaini, F., Pereira, L., Westrupp, M. F., & Moraes, T. M. R. (2017). *Bacias hidrográficas da Região de Joinville: gestão e dados*. Joinville: Editora Univille. Recuperado em 23 de junho de 2023, de https://www.univille.edu.br/account/editora/VirtualDisk.html/downloadDirect/1145899/Bacias_hidrograficas_2017.pdf
- Orcan, F. (2020). Parametric or non-parametric: skewness to test normality for mean comparison. *International Journal of Assessment Tools in Education*, 7(2), 255-265. <http://dx.doi.org/10.21449/ijate.656077>.
- Passos, A. L. L., Muniz, D. H. F., Rangel, L. P., Oliveira-Filho, E. C., Passos, A. L. L., Muniz, D. H. F., Rangel, L. P., & Oliveira-Filho, E. C. (2019). Selection of variables in the definition of a water quality index for the Brazilian Federal District. *Revista Ambiente & Água*, 14(4), e2385. <https://doi.org/10.4136/ambi-agua.2385>.
- Pessoa, M. A. R., Souza, F. J., Domingos, P., Azevedo, J. P. S., Pessoa, M. A. R., Souza, F. J., Domingos, P., & Azevedo, J. P. S. (2020). Índice fuzzy de qualidade de água para ambiente lótico - IQA_{FAL}. *Engenharia Sanitaria e Ambiental*, 25(1), 21-30. <http://dx.doi.org/10.1590/s1413-41522020147587>.
- Pho, K.-H., Ly, S., Ly, S., & Lukusa, T. M. (2019). Comparison among Akaike Information Criterion, Bayesian Information Criterion and Vuong's test in model selection: a case study of violated speed regulation in Taiwan. *Journal of Advanced Engineering and Computation*, 3(1), 293-303. Recuperado em 23 de junho de 2023, de <https://jaec.vn/index.php/JAEC/article/view/220>
- Pinilla, J. O., & Rico, A. F. O. (2021). Pearson y Spearman, coeficientes intercambiables? *Comunicaciones en Estadística*, 14(1), 53-63. Recuperado em 23 de junho de 2023, de <https://revistas.usantotomas.edu.co/index.php/estadistica/article/view/6769>
- Pinto, C. C., Faria, F. P. C. P., & Almeida, G. M. (2021). Utilização de modelo baseado em análise por componentes principais para identificação de condições críticas de qualidade de água superficial. *Revista Ibero-Americana de Ciências Ambientais*, 12(4), 288-305. Recuperado em 23 de junho de 2023, de <http://www.sustenere.co/index.php/rica/article/view/CBPC2179-6858.2021.004.0024>
- Pisicã, D., Dammers, R., Boersma, E., & Volovici, V. (2022). Tenets of good practice in regression analysis. A brief tutorial. *World Neurosurgery*, 161, 230-239.e6. <http://dx.doi.org/10.1016/j.wneu.2022.02.112>.
- Prado, R. B., Parron, L. M., Campanha, M. M., Silva, M. S. L., Matthiensen, A., & Lima, J. E. F. W. (2018). *Água e saneamento: contribuições da Embrapa*. Brasília: Embrapa. Recuperado em 23 de junho de 2023, de <http://www.alice.cnptia.embrapa.br/handle/doc/1098478>
- Rabelo, G. C., Biscaíno Neto, J., & Freire, R. (2020). Análise dos dados de monitoramento de qualidade de água de uma bacia hidrográfica na proposta de um índice qualidade de água urbano (IQAu). *Revista AIDIS de Ingeniería y Ciencias Ambientales: Investigación, Desarrollo y Práctica*, 13(3), 806-821. <http://dx.doi.org/10.22201/iingen.0718378xe.2020.13.3.68355>.
- Reis, I., Fernandes, C. E., & Fernandes, L. I. F. A. (2020). Urban occupation: miscellaneous use of water resources in João Cesario Stream in Anápolis-GO. *Brazilian Journal of Development*, 6(1), 1116-1137. <http://dx.doi.org/10.34117/bjdv6n1-079>.
- Rezaie-Balf, M., Attar, N. F., Mohammadzadeh, A., Murti, M. A., Ahmed, A. N., Fai, C. M., Nabipour, N., Alaghmand, S., & El-Shafie, A. (2020). Physicochemical parameters data assimilation for efficient improvement of water quality index prediction: comparative assessment of a noise suppression hybridization approach. *Journal of Cleaner Production*, 271, 122576. Recuperado em 23 de junho de 2023, de <http://www.sciencedirect.com/science/article/pii/S0959652620326238>
- Rodrigues, R. S. S., Fernandes, L. L., Blanco, C. J. C., Teixeira, L. C. G. M., & Bittencourt, G. M. (2020). Abordagem sobre o comportamento do iqa a partir da degradação dos recursos hídricos na amazônia. *Revista AIDIS de Ingeniería y Ciencias Ambientales: Investigación, Desarrollo y Práctica*, 13(3), 885-904. <http://dx.doi.org/10.22201/iingen.0718378xe.2020.13.3.67814>.
- RStudio Team. (2020). *RStudio: integrated development for R (2022.02.3)*. Recuperado em 23 de junho de 2023, de <https://rstudio.com/>
- Santos, G. B., Valentini, M. H. K., Silva, L. A., Franz, H. S., Corrêa, B. L., Duarte, V. H., Silva, M. A., Corrêa, M. G., Vieira, B. M., Nadaleti, W. C., & Vieira, B. M. (2020). Análise da qualidade das águas do Arroio

- Moreira/Fragata (RS) através de métodos estatísticos. *Revista Ibero-Americana de Ciências Ambientais*, 11(4), 217-226. <http://dx.doi.org/10.6008/CBPC2179-6858.2020.004.0019>.
- Soares, A. L. C., & Pinto, C. C. (2017). IV-083 - Avaliação da qualidade da água da sub-bacia do Rio das Velhas, inserida na Bacia Hidrográfica do Rio São Francisco. In *Anais do Congresso ABES FENASAN 2017* (pp. 1-8). Rio de Janeiro: ABES.
- Souza, A., & Santos, D. A. S. (2018). Análise das componentes principais no processo de monitoramento ambiental. *Nativa*, 6(6), 639-647. Recuperado em 23 de junho de 2023, de <https://periodicoscientificos.ufmt.br/ojs/index.php/nativa/article/view/6453>
- Souza, G. S., Ceconello, S. T., & Centeno, L. N. (2021). Análise espacial de similaridade da qualidade da água de um rio localizado no sul do Rio Grande do Sul. *Revista Ambientale*, 13(4), 38-49. <https://doi.org/10.48180/ambientale.v13i4.319>.
- Souza, S. D., & Ponzilacqua, M. H. P. (2020). Acesso à água de qualidade como direito humano fundamental: a garantia do mínimo existencial. *Humanidades & Inovação*, 7(20), 537-551. Recuperado em 23 de junho de 2023, de <https://revista.unitins.br/index.php/humanidadesinovacao/article/view/3832>
- Sudarshan, P., Mahesh, M. K., & Ramachandra, T. V. (2019). Assessment of seasonal variation in water quality and Water Quality Index (WQI) of Hebbal Lake, Bangalore, India. *Environment and Ecology*, 37(1B), 309-317. Recuperado em 23 de junho de 2023, de http://wgbis.ces.iisc.ernet.in/energy/water/paper/Hebbal_environment%20and%20Ecology/MS1.pdf
- Taloor, A. K., Pir, R. A., Adimalla, N., Ali, S., Manhas, D. S., Roy, S., & Singh, A. K. (2020). Spring water quality and discharge assessment in the Basantar watershed of Jammu Himalaya using geographic information system (GIS) and water quality Index (WQI). *Groundwater for Sustainable Development*, 10, 100364. <http://dx.doi.org/10.1016/J.GSD.2020.100364>.
- Zancul, J. S. (2015). Direitos humanos à água e ao saneamento e a política de saneamento básico no Brasil. *Cadernos Ibero-Americanos de Direito Sanitário*, 4(2), 23-46. <http://dx.doi.org/10.17566/ciads.v4i2.151>.

Contribuições dos autores:

Glauber Rover Cadorin: investigação, análise formal, redação – rascunho original, revisão e edição.

Anelise Destefani: visualização, análise formal.

Carlos Eduardo Nogueira Martins: supervisão, validação.

André Luis Fachini de Souza: supervisão, redação - revisão e edição.