

**UNIVERSIDADE BRASIL
CURSO DE GRADUAÇÃO EM ESTATÍSTICA
CAMPUS MOOCA**

LINA IKEJIRI

**APLICAÇÃO DA ANÁLISE DE COMPONENTES PRINCIPAIS NA AVALIAÇÃO DA
QUALIDADE DA ÁGUA DO RESERVATÓRIO TANQUE GRANDE, GUARULHOS
– SP**

São Paulo -SP
2022

LINA IKEJIRI

**APLICAÇÃO DA ANÁLISE DE COMPONENTES PRINCIPAIS NA
AVALIAÇÃO DA QUALIDADE DA ÁGUA DO RESERVATÓRIO
TANQUE GRANDE, GUARULHOS – SP**

Trabalho de Conclusão do Curso de Graduação apresentado à Universidade Brasil, como parte dos requisitos necessários para obtenção do título de Bacharel em Estatística.

Prof. Dr. Antonio Chávez Zena
Orientador

São Paulo – SP
2022

**Ficha catalográfica elaborada pelo Sistema de Bibliotecas da Universidade Brasil,
com os dados fornecidos pelo (a) autor (a).**

I28a IKEJIRI, Lina.

Aplicação da análise de componentes principais na avaliação da qualidade da água do Reservatório Tanque Grande, Guarulhos – SP / Lina Ikejiri – Mooca, SP: Universidade Brasil, 2022.

31 f. il. color.

Trabalho de Conclusão de Curso de Estatística da Universidade Brasil.

Orientação: Prof. Dr. Antonio Chávez Zena.

1. Análise de componentes principais. 2. Qualidade de água. 3. Reservatório Tanque Grande. I. Ikejiri, Lina. II. Título.

CDD 001.422

AGRADECIMENTOS

A Deus, pela grandiosa luz que nos emana, por nos permitir estar aqui nesta jornada da vida e pelas inspirações enviadas nos momentos mais difíceis.

Ao professor Antonio C. Zena pela orientação, pela seriedade e comprometimento singulares, por compartilhar suas valiosas experiências de vida.

Ao professor Paulo L. S. Gomes, pela orientação inicial e sugestões imprescindíveis que deram o formato para o presente trabalho.

Ao professor Paulo C. F. Lima, pela disposição em proporcionar o máximo de conteúdos para as aulas, com quem aprendi ferramentas aplicadas neste trabalho.

Ao professor Koki F. Oikawa, por sua incrível didática e, com isso, trazer leveza a temas tão complexos, por sua paixão por ensinar e enorme generosidade em compartilhar todo o seu conhecimento.

Ao professor Juliano van Melis, por trazer às aulas conteúdos que nos fazem pensar além e "fora da caixa".

Aos meus pais, Hideo e Yoko, pelo amor incondicional, por todos os valores passados, que fazem de mim o que sou, por toda a sabedoria compartilhada.

Ao meu namorado Adônis, meu maior incentivador, por trazer palavras precisas nos momentos mais necessários, por ser compreensivo, estar sempre disposto a ajudar, ouvir meus desabafos e por me apoiar sempre.

À Beatriz pela amizade, pela parceria do dia a dia, pelas divertidas conversas, pelo incentivo nessa jornada da graduação e disposição em ajudar sempre.

Ao Parmigiani, pelo companheirismo durante a graduação, principalmente nas dificuldades enfrentadas.

Ao Odair, por resgatar algo que estava adormecido em mim, o interesse pelas Exatas, ao sugerir alguns cursos que foram cruciais posteriormente na escolha desta graduação.

A todos, meus mais sinceros agradecimentos.

RESUMO

Os estudos sobre a qualidade dos corpos d'água têm se tornado cada vez mais fundamentais a fim de se compreender os possíveis impactos de origem antrópica ao saneamento e ao meio ambiente. Neste sentido, os crescentes desafios decorrentes do aumento da complexidade na gestão dos recursos hídricos exigem ferramentas adequadas de suporte à avaliação dos parâmetros que caracterizam a qualidade de água de um manancial, como condutividade, demanda bioquímica de oxigênio, fósforo total, séries de nitrogênio, oxigênio dissolvido, pH, série de sólidos, temperatura da água, turbidez e coliformes termotolerantes. O presente estudo tem como objetivo a aplicação da Análise de Componentes Principais (PCA) como ferramenta estatística na avaliação dos parâmetros da qualidade da água do reservatório Tanque Grande, manancial localizado no município de Guarulhos, SP. O manancial Tanque Grande é integrante da bacia hidrográfica do Baquirivu-Guaçu e pertence à Unidade de Gerenciamento de Recursos Hídricos do Estado de São Paulo Alto Tietê. A Companhia Ambiental do Estado de São Paulo (CETESB), dentro de suas atribuições, ao monitorar esse reservatório, subsidia o diagnóstico da qualidade das águas para fins de abastecimento público. Desta forma, por meio do sistema Infoáguas, foram obtidos dados correspondentes ao monitoramento do referido reservatório, entre abril de 1987 e março de 2016, anualmente divididos por períodos chuvoso (outubro a março) e seco (abril a setembro). A técnica de Análise de Componentes Principais permite a redução da dimensionalidade das variáveis, por meio de combinações lineares, indicando aquelas que mais contribuem para avaliação da qualidade da água do local estudado. Para ambos os períodos, 5 componentes principais foram considerados os mais relevantes na análise de dados, respondendo por 81,58% da variância total explicada para o período chuvoso e 76,46% para o período seco. Nas duas condições climáticas, as variáveis condutividade, nitrogênio amoniacal e sólidos totais, incluindo o nitrato para o período seco representaram a primeira componente, sendo responsáveis por parte significativa da variação da qualidade da água desse manancial. De modo geral, o estudo foi capaz de caracterizar resumidamente a qualidade da água do reservatório Tanque Grande nesse período, levando-se em consideração as estações seca e chuvosa.

Palavras-chave: Análise de Componentes Principais. Qualidade de água. Reservatório Tanque Grande.

LISTA DE FIGURAS

Figura 1 - Abastecimento de Guarulhos	10
Figura 2 - Sub-bacias hidrográficas que compõem Guarulhos	14
Figura 3 - Mapa de Guarulhos com destaque para a represa Tanque Grande e o ponto de captação TGDE00900.....	15
Figura 4 – CP1 e CP2 incluindo as 12 variáveis e 29 observações do período seco	23
Figura 5 – CP1 e CP2 incluindo as 12 variáveis e 29 observações do período chuvoso	24

LISTA DE TABELAS

Tabela 1 – Percentual da variância total explicada nos períodos seco e chuvoso....	19
Tabela 2 – Componentes Principais – período seco	20
Tabela 3 – Componentes Principais – período chuvoso.....	20
Tabela 4 – Correlação entre as variáveis e as componentes principais – período seco	21
Tabela 5 – Correlação entre as variáveis e as componentes principais – período chuvoso	22

SUMÁRIO

1 INTRODUÇÃO.....	9
2 OBJETIVO.....	13
3 MATERIAL E MÉTODOS	14
4 RESULTADOS E DISCUSSÃO	19
5 CONCLUSÃO.....	27
REFERÊNCIAS	28

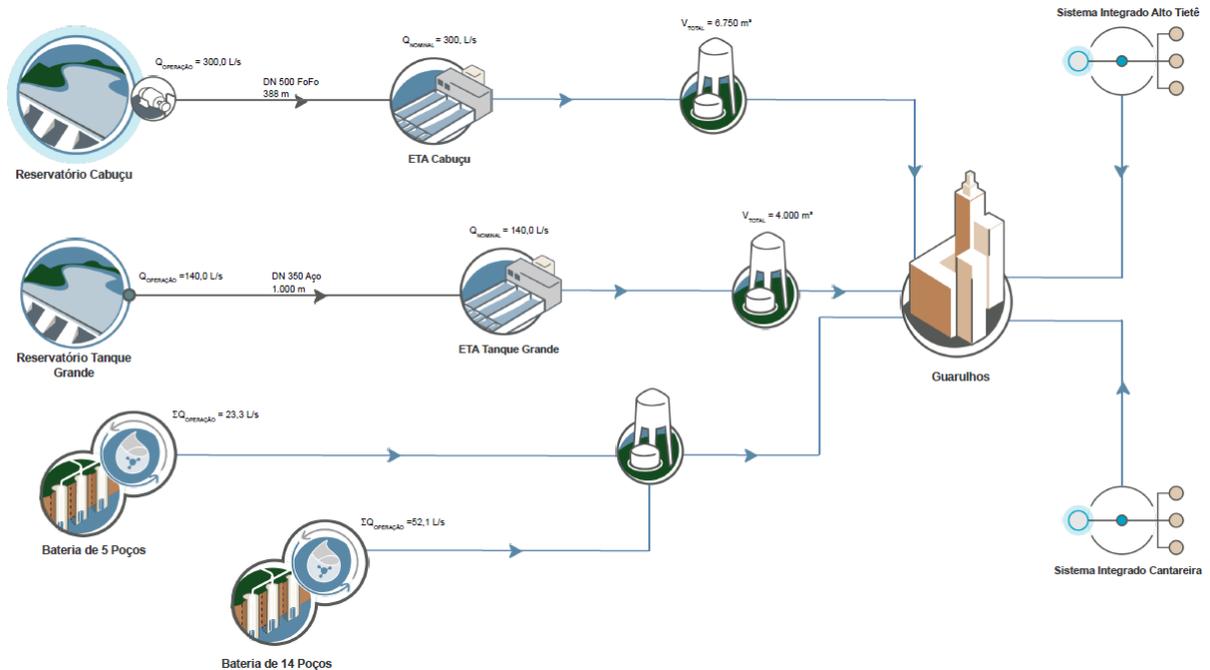
1 INTRODUÇÃO

Segundo projeções da Organização das Nações Unidas (UNITED NATIONS POPULATION FUND, 2022), a população mundial atingiu oito bilhões de pessoas em meados de novembro de 2022. Embora o ritmo de crescimento tenha diminuído nas últimas décadas, o aumento populacional traz alertas sobre os crescentes impactos ao meio ambiente e à saúde como um todo. Tais consequências refletem diretamente na qualidade das águas, seja para fins de consumo humano, fins recreacionais, seja para a comunidade de organismos aquáticos.

O município de Guarulhos - SP é considerado a segunda cidade mais populosa do estado de São Paulo, apresentando 1.372.533 habitantes e 100% urbanizada, conforme dados da Fundação Sistema Estadual de Análise de Dados (SEADE, 2022), evidenciando os enormes desafios à gestão do município, incluindo os serviços de saneamento básico, fundamentais à população.

Com relação ao abastecimento de Guarulhos, a água é captada majoritariamente do sistema integrado Cantareira e sistema integrado Alto Tietê. O abastecimento vindo dos sistemas isolados representados pelos reservatórios Cabuçu e Tanque Grande responde por 8% da distribuição de água tratada (PREFEITURA DE GUARULHOS, 2022), além da captação por poços profundos (ANA, 2021), conforme Figura 1.

Figura 1 - Abastecimento de Guarulhos



Fonte: Agência Nacional das Águas, 2021

Segundo o Conselho Estadual de Saneamento (2018), até 2016, Guarulhos contava com 404.604 unidades cadastradas residenciais ativas de água, correspondendo a um índice de atendimento de 88,4%. Deve ser considerado que, no final de 2018, a prestação de serviços públicos de abastecimento de água e esgotamento sanitário, antes sob responsabilidade da autarquia municipal Serviço Autônomo de Água e Esgoto (SAEE), foi atribuída à Companhia de Saneamento Básico do Estado de São Paulo. Entretanto, com o aumento populacional, tornou-se um grande desafio atender a toda crescente demanda de água necessária (SABESP, 2018).

Além dos possíveis impactos à quantidade de água, outra grande preocupação, não menos importante, recai sobre a qualidade da água. Apenas a partir de 2010, o tratamento de esgoto foi iniciado em Guarulhos, período em que 70% dos esgotos da cidade eram despejados sem tratamento na Bacia hidrográfica do rio Baquirivu-Guaçu (VARGAS et al., 2015), a maior das cinco bacias que compõem o município. Tais impactos, principalmente de origem antrópica, exigem ferramentas adequadas para o gerenciamento cada vez mais complexo dos mananciais. Segundo Carrasco et al. (2019), a qualidade da água pode ser definida por uma série de características físicas,

químicas e biológicas, sendo que a natureza complexa desses estudos exige modelos para identificar as variáveis que mais exercem influência em sua qualidade.

A Portaria GM/MS Nº 888/21 (BRASIL, 2021) que dispõe sobre os procedimentos de controle e de vigilância da qualidade da água para consumo humano e seu padrão de potabilidade, por meio de seu artigo 14, inciso XIV, pode-se extrair a exigência de atendimento dos critérios de qualidade da água dos mananciais de abastecimento de acordo com a respectiva classe de enquadramento, conforme definido pela resolução CONAMA 357/05 (BRASIL, 2005). Esta estabelece os critérios para atendimento à classe a que um corpo hídrico foi enquadrado de acordo com seus usos preponderantes. O manancial Tanque Grande, embora considerado como Classe especial pela Companhia Ambiental do Estado de São Paulo, em seus relatórios anuais, é tratado como classe 1 para fins de estabelecimento de parâmetros para sua avaliação (CUCIO, 2014).

Cecconello, Centeno e Guedes (2018) propuseram como ferramenta de gestão da qualidade da água um índice de qualidade da água modificado que, diferentemente daquele apresentado pela National Sanitation Foundation (NSF) no ano de 1970 e conhecido por índice de qualidade da água - IQA, foi baseado na análise estatística multivariada, demonstrando ser mais apropriada esta abordagem pelo fato de o índice incluir os pesos de cada variável de acordo com seu grau de importância. Ao se utilizar da análise multivariada para classificação ou avaliação do comportamento das variáveis com relação à qualidade dos corpos d'água, muitos pesquisadores utilizam como ferramentas, entre outras, Análise de Componentes Principais (PCA), a Análise de Cluster e Análise Discriminante (AZHAR et al., 2015; BENGRAÏNE; MARHABA, 2003; CARRASCO et al., 2019; CECCONELLO; CENTENO; GUEDES, 2018; MAHAPATRA et al., 2012). Variáveis como condutividade, clorofila-a, demanda bioquímica de oxigênio (DBO), fósforo total, nitrogênio, oxigênio dissolvido (OD), pH, série de sólidos, temperatura da água, transparência, turbidez e coliformes termotolerantes são comumente utilizadas em tais caracterizações. Segundo Mazlum, Özer e Mazlum (1999), ao estudarem a bacia hidrográfica de Sacaria, na Turquia, demonstraram a utilidade da técnica de análise de componentes principais aplicada à análise dos dados de qualidade da água. Os autores ainda citam, baseado em revisão da literatura sobre análise de componentes principais, a adequação do uso da PCA para dados de qualidade da água.

Devido à forte pressão antrópica que sofre e por ser um manancial que abastece apenas Guarulhos, definiu-se como objeto de estudo o sistema isolado Tanque Grande por meio da ferramenta de análise multivariada Análise de Componentes Principais aplicada à avaliação das variáveis de qualidade da água em seu ponto de captação de abastecimento de água.

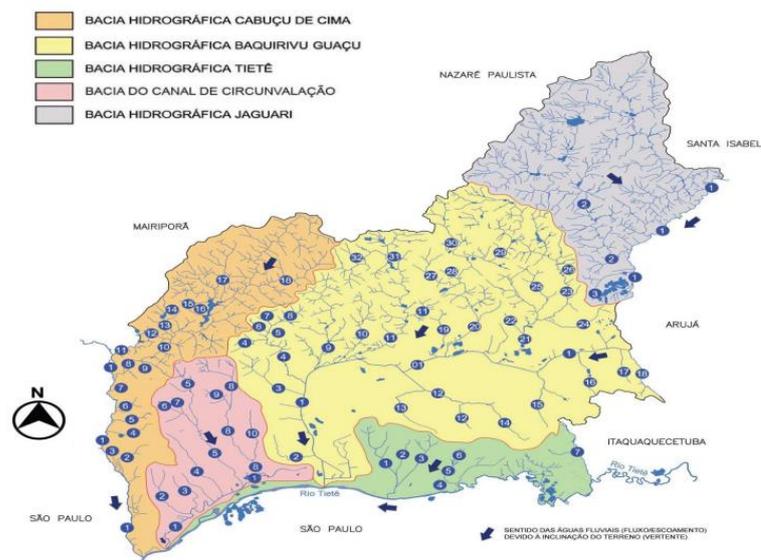
2 OBJETIVO

O presente estudo tem como objetivo a aplicação da Análise de Componentes Principais (PCA) como ferramenta estatística na avaliação dos parâmetros da qualidade da água do reservatório Tanque Grande, manancial localizado no município de Guarulhos, SP, no intervalo de 1987 a 2016, divididos em períodos seco e chuvoso. Desta forma, serão avaliados os parâmetros que melhor explicam a variação na qualidade da água e se tais variáveis são influenciadas pela sazonalidade.

3 MATERIAL E MÉTODOS

Considerando os recursos hídricos de Guarulhos, o município é composto por 5 bacias hidrográficas: Cabuçu de Cima, Baquirivu-Guaçu, Tietê, Canal de Circunvalação e Jaguari (Figura 2). O presente estudo restringiu-se ao manancial Tanque Grande, integrante da bacia hidrográfica do Baquirivu-Guaçu e pertencente à Unidade de Gerenciamento de Recursos Hídricos do Estado de São Paulo Alto Tietê (IPT, 2014).

Figura 2 - Sub-bacias hidrográficas que compõem Guarulhos



Fonte: Guarulhos, 2017 (modificado)

De acordo com Piasentin et al. (2009), cinco sub-bacias tributárias abastecem o reservatório Tanque Grande, a partir do qual ocorre a captação de água da estação de tratamento de água Tanque Grande, localizada na Estrada Saboó, no bairro do Tanque Grande (Figura 3).

Figura 3 - Mapa de Guarulhos com destaque para a represa Tanque Grande e o ponto de captação TGDE00900



Fonte: CETESB, 2022 (modificado)

O ponto de monitoramento da Companhia Ambiental do Estado de São Paulo (CETESB), TGDE00900, localizado na captação do manancial Tanque Grande, foi o utilizado para o presente trabalho. Para tanto, os dados secundários da qualidade da água foram obtidos do Sistema Infoáguas (CETESB, 2022), alimentado pela referida empresa que, dentro de suas atribuições, ao monitorar esse reservatório, subsidia o diagnóstico da qualidade das águas para, entre outros fins, de abastecimento público.

O período de estudo foi definido entre abril de 1987 e março de 2016, anualmente divididos pelos períodos chuvoso (outubro a março) e seco (abril a setembro), divisão esta definida com base nos dados de precipitação apresentados na página eletrônica Weather Spark (2022). Desta forma, foram avaliadas 29 medições para o período chuvoso e 29 medições para o período seco.

Com relação aos parâmetros de qualidade da água utilizados no estudo, foram considerados os seguintes: condutividade, demanda bioquímica de oxigênio, fósforo total, séries de nitrogênio, oxigênio dissolvido, pH, série de sólidos, temperatura da água, turbidez e coliformes termotolerantes.

A condutividade é uma variável física de qualidade da água, medida em microsiemens/cm, e indica a quantidade de sais dissolvidos na água na forma de íons,

de origem geogênica ou antropogênica, podendo ser uma medida indireta da concentração de poluentes (FRANÇA, 2009).

A demanda bioquímica de oxigênio, conhecida como DBO, indica a quantidade de oxigênio necessária para oxidar a matéria orgânica por meio da decomposição microbiana aeróbia para a forma inorgânica estável. Desta forma, a elevação nos valores de DBO na água indicam despejos de origem orgânica principalmente. Deve-se considerar que a DBO representa a quantidade de oxigênio consumida durante 5 dias a 20°C de incubação (CETESB, 2021).

O fósforo é um indicador do despejo de esgotos sanitários na água devido, principalmente, à presença de detergentes, além de material de origem fecal. Faz-se importante considerar ainda a contribuição dos efluentes industriais para o excesso de fósforo nos corpos d'água (FRANÇA, 2009).

A fonte antropogênica de nitrogênio nas águas naturais são os despejos domésticos, industriais, excrementos de animais e fertilizantes. A forma predominante do nitrogênio na água pode indicar o estágio da poluição (VON SPERLING, 1996). Considerando a autodepuração natural em rios, na zona de degradação, identifica-se o nitrogênio orgânico, na zona de decomposição ativa, verifica-se o nitrogênio amoniacal, o nitrito estará na zona de recuperação e o nitrato estará nas zonas de águas limpas. A determinação de nitrogênio pelo método Kjeldahl apresenta como resultado as frações totais de nitrogênio orgânico e amoniacal.

O oxigênio dissolvido (OD) é um importante parâmetro na caracterização dos corpos d'água. Uma vez que sua concentração de saturação é igual a 9,2mg/L, valores superiores a essa concentração podem indicar a presença de algas. Já valores muito baixos podem indicar a presença de atividade bacteriana, que se utiliza do oxigênio em seus processos respiratórios durante a estabilização da matéria orgânica (VON SPERLING, 1996). Dependendo de sua magnitude, pode caracterizar a poluição.

O potencial hidrogeniônico (pH) indica a concentração de íons hidrogênio na água, que pode apresentar condições de acidez ($\text{pH} < 7$), neutralidade ($\text{pH} = 7$) e alcalinidade ($\text{pH} > 7$) (VON SPERLING, 1996). Uma faixa considerada compatível para a sobrevivência de maior parte da comunidade aquática está em pH entre 6 e 9. A presença de efluentes na água e poluentes atmosféricos, como a chuva ácida, podem provocar alterações no pH (FRANÇA, 2009).

Em saneamento, os sólidos nas águas correspondem a toda matéria que permanece como resíduo, após evaporação, secagem ou calcinação da amostra a

uma temperatura pré-estabelecida durante um tempo fixado. Em linhas gerais, as operações de secagem, calcinação e filtração são as que definem as diversas frações de sólidos presentes na água (sólidos totais, em suspensão, dissolvidos, fixos e voláteis). Os métodos empregados para a determinação de sólidos são gravimétricos (utilizando-se balança analítica ou de precisão). Nos estudos de controle de poluição das águas naturais, principalmente nos estudos de caracterização de esgotos sanitários e de efluentes industriais, as determinações dos níveis de concentração das diversas frações de sólidos resultam em um quadro geral da distribuição das partículas com relação ao tamanho (sólidos em suspensão e dissolvidos) e com relação à natureza (fixos ou minerais e voláteis ou orgânicos) (CETESB, 2021).

A temperatura superficial é naturalmente influenciada pelas variações ao longo do dia, sazonais, taxa de fluxo, profundidade da coluna d'água, altitude e latitude. Ações antrópicas, como despejos industriais, podem provocar sua elevação e, conseqüentemente, influenciar as variáveis físico-químicas (CETESB, 2021).

A turbidez, medida em unidades de turbidez (uT), representa o grau de interferência decorrente da presença de sólidos em suspensão, seja de origem natural ou antropogênica à passagem da luz através da água, conferindo uma aparência turva a mesma (VON SPERLING, 1996), assim como detritos orgânicos, que também atenuam a intensidade do feixe de luz. Desta forma, esgotos domésticos e efluentes industriais elevam a turbidez da água (CETESB, 2021).

Os coliformes termotolerantes são micro-organismos do grupo coliforme, tendo como principal representante a *Escherichia coli*, presente em qualidades elevadas em fezes de aves, mamíferos e humanas (CETESB, 2021), sendo de origem unicamente fecal e considerado um organismo indicador de contaminação de origem dessa natureza nos corpos d'água (BRASIL, 2013).

A ferramenta de análise multivariada escolhida para a avaliação dos dados do estudo é a Análise de Componentes Principais (PCA). A análise de componentes principais (*Principal Component Analysis*), conhecida por PCA, e introduzida por Karl Pearson em 1901, utiliza-se de combinações lineares das variáveis originais, combinações estas chamadas de componentes principais, para explicar a estrutura de variância e covariância dos dados (MINGOTI, 2005). É importante ressaltar que as componentes principais obtidas passam a ser combinações lineares não correlacionadas, sendo a primeira componente principal a combinação linear contendo o máximo de variância. Desta forma, considerando que as primeiras

componentes representam a maioria da variância de toda a população, essas componentes poderão “substituir” as variáveis originais, sem perda significativa da informação (JOHNSON; WICHERN, 2007), uma vez que esta estará contida nas componentes principais. Desta maneira, a técnica possibilita a redução da dimensionalidade das variáveis indicando aquelas que mais contribuem para a avaliação da qualidade da água do local estudado.

Conforme definido anteriormente, 58 observações foram divididas em períodos seco e chuvoso, gerando uma PCA por período. Cada observação foi definida como a mediana dos valores obtidos de 3 coletas ocorridas em cada período. Não se optou em utilizar os valores de cada coleta devido a dados faltantes para algumas variáveis, sendo que para a aplicação da PCA, cada observação deve incluir valores para todas as variáveis. Para o período seco, a identificação utilizada para cada observação foi o ano (2 dígitos) seguido da letra S (Seco). O período chuvoso, que abrange o final de um ano e início do ano seguinte, foi identificado pelos dois anos correspondentes (2 dígitos para cada ano) e a letra C (Chuvoso). Para resultados apresentados com o sinal “<”, foi mantido apenas o resultado.

Para o presente estudo, utilizou-se a linguagem R, por meio do R Studio, versão 4.2.1, para o processamento dos dados. A partir das variáveis originais, foi gerada a matriz de correlação e aplicada a técnica de componentes principais. Conforme Mingoti (2005), tal procedimento corresponde a padronização das variáveis e a aplicação da técnica de componentes principais à sua matriz de covariâncias.

4 RESULTADOS E DISCUSSÃO

A partir das 12 variáveis da qualidade da água avaliadas neste estudo, foram obtidas 12 componentes principais. Uma vez que a ideia de se utilizar a ferramenta de PCA é reduzir a dimensionalidade dos dados, concentrando a informação em um número menor de variáveis, adotou-se inicialmente o critério de Kaiser. Este é utilizado para definir as componentes principais que serão selecionadas, ou seja, cujos autovalores são $\lambda \geq 1$, preservando as combinações lineares que conseguiriam explicar, no mínimo, a quantidade de variância de uma variável original padronizada (MINGOTI, 2005). Dessa forma, foram selecionadas 5 componentes principais que representaram 76,46% da variação total para o período seco e 81,58% para o período chuvoso (Tabela 1).

Tabela 1 – Percentual da variância total explicada nos períodos seco e chuvoso

Componente Principal (CP)	Período seco			Período chuvoso		
	Autovalor	Variância explicada (%)	Variância explicada acumulada (%)	Autovalor	Variância explicada (%)	Variância explicada acumulada (%)
CP1	3,1269833	26,06%	26,06%	3,7599426	31,33%	31,33%
CP2	1,9638204	16,36%	42,42%	2,3769148	19,81%	51,14%
CP3	1,5521454	12,93%	55,36%	1,4633525	12,20%	63,33%
CP4	1,4106148	11,76%	67,11%	1,1260207	9,38%	72,72%
CP5	1,1212207	9,34%	76,46%	1,0631079	8,86%	81,58%
CP6	0,7928819	6,61%	83,06%	0,6701521	5,59%	87,16%
CP7	0,7146238	5,96%	89,02%	0,5029931	4,19%	91,35%
CP8	0,4841361	4,03%	93,05%	0,4165676	3,47%	94,83%
CP9	0,346734	2,89%	95,94%	0,2758807	2,30%	97,12%
CP10	0,2092591	1,74%	97,69%	0,1809506	1,51%	98,63%
CP11	0,159737	1,33%	99,02%	0,1320555	1,10%	99,73%
CP12	0,1178434	0,98%	100%	0,0320621	0,27%	100%

Fonte: Autoria própria

Nas tabelas seguintes (Tabelas 2 e 3) são exibidos os resultados obtidos da aplicação da técnica de análise de componentes principais a partir da matriz de correlação dos dados originais, onde constam todas as 12 variáveis e seus pesos nas

diferentes CP. Ainda que se tenha obtido 12 componentes, foram apresentados as cinco selecionadas.

Tabela 2 – Componentes Principais – período seco

Variáveis	CP1	CP2	CP3	CP4	CP5
Coliformes Termotolerantes	0,262	0,004	-0,369	0,285	-0,431
Condutividade	0,412	-0,029	0,266	-0,299	-0,150
DBO	0,121	-0,608	-0,051	-0,130	-0,233
Fósforo Total	-0,141	-0,425	-0,503	-0,136	-0,165
Nitrogênio Amoniacal	0,438	0,096	-0,223	-0,068	0,111
Nitrogênio Kjeldahl	0,159	-0,094	-0,403	-0,432	0,507
Nitrogênio Nitrato	0,398	0,098	0,204	0,084	0,069
Oxigênio Dissolvido	-0,252	0,301	-0,180	-0,079	-0,515
pH	0,049	-0,503	0,435	0,017	-0,106
Sólido Total	0,488	0,118	-0,017	-0,008	-0,293
Temperatura da Água	-0,009	0,250	0,031	-0,657	-0,206
Turbidez	0,220	0,020	-0,248	0,395	0,176

Fonte: Autoria própria

Tabela 3 – Componentes Principais – período chuvoso

Variáveis	CP1	CP2	CP3	CP4	CP5
Coliformes Termotolerantes	0,231	-0,489	0,186	-0,325	-0,070
Condutividade	0,389	0,248	0,213	0,054	0,268
DBO	0,125	0,307	0,381	-0,458	-0,016
Fósforo Total	-0,316	0,106	0,082	-0,270	0,403
Nitrogênio Amoniacal	0,382	0,179	0,150	0,331	0,036
Nitrogênio Kjeldahl	0,329	-0,144	0,093	0,502	-0,019
Nitrogênio Nitrato	0,303	0,233	-0,467	-0,124	-0,177
Oxigênio Dissolvido	-0,227	0,074	0,580	0,290	0,217
pH	0,149	-0,041	-0,304	-0,085	0,782
Sólido Total	0,439	0,113	0,024	-0,166	0,089
Temperatura da Água	0,016	0,522	0,132	-0,177	-0,240
Turbidez	0,260	-0,442	0,269	-0,283	-0,057

Fonte: Autoria própria

Segundo Bengraïne e Marhaba (2003), no estudo de dados ambientais, a distribuição das variáveis entre as componentes tende a se espalhar e, ainda, o fato

de terem sido formados artificialmente por poucos itens, que já devem ter sido destacados nas primeiras componentes, conduzem ao questionamento da possibilidade de se descartar aqueles avaliados como de menor relevância. Desta maneira, tanto para o período seco como chuvoso, as variáveis mais correlacionadas com cada componente principal foram avaliadas com cautela com relação aos possíveis impactos na qualidade da água do reservatório durante o período estudado. Segundo França (2009), a partir dos pesos das variáveis nas componentes principais, pode-se verificar os parâmetros mais importantes em cada componente e constatado pelas correlações calculadas entre as variáveis e as componentes. A autora ainda atribui uma correlação mínima de $|0,7|$ como critério para estabelecer as variáveis relevantes para as componentes principais. Desta forma, a partir das Tabelas 4 e 5, é possível verificar as variáveis que apresentaram correlações significativas com as CP.

Tabela 4 – Correlação entre as variáveis e as componentes principais – período

seco					
Variáveis	PC1	PC2	PC3	PC4	PC5
Coliformes Termotolerantes	0,463	0,005	-0,460	0,339	-0,456
Condutividade	0,728	-0,041	0,331	-0,356	-0,159
DBO	0,214	-0,852	-0,063	-0,155	-0,247
Fósforo Total	-0,249	-0,596	-0,626	-0,162	-0,174
Nitrogênio Amoniacal	0,774	0,135	-0,278	-0,080	0,117
Nitrogênio Kjeldahl	0,282	-0,132	-0,502	-0,513	0,537
Nitrogênio Nitrato	0,703	0,138	0,254	0,100	0,073
Oxigênio Dissolvido	-0,446	0,422	-0,224	-0,094	-0,546
pH	0,087	-0,705	0,542	0,020	-0,113
Sólido Total	0,863	0,165	-0,021	-0,009	-0,310
Temperatura da Água	-0,017	0,351	0,038	-0,781	-0,218
Turbidez	0,389	0,028	-0,309	0,470	0,187

Fonte: Autoria própria

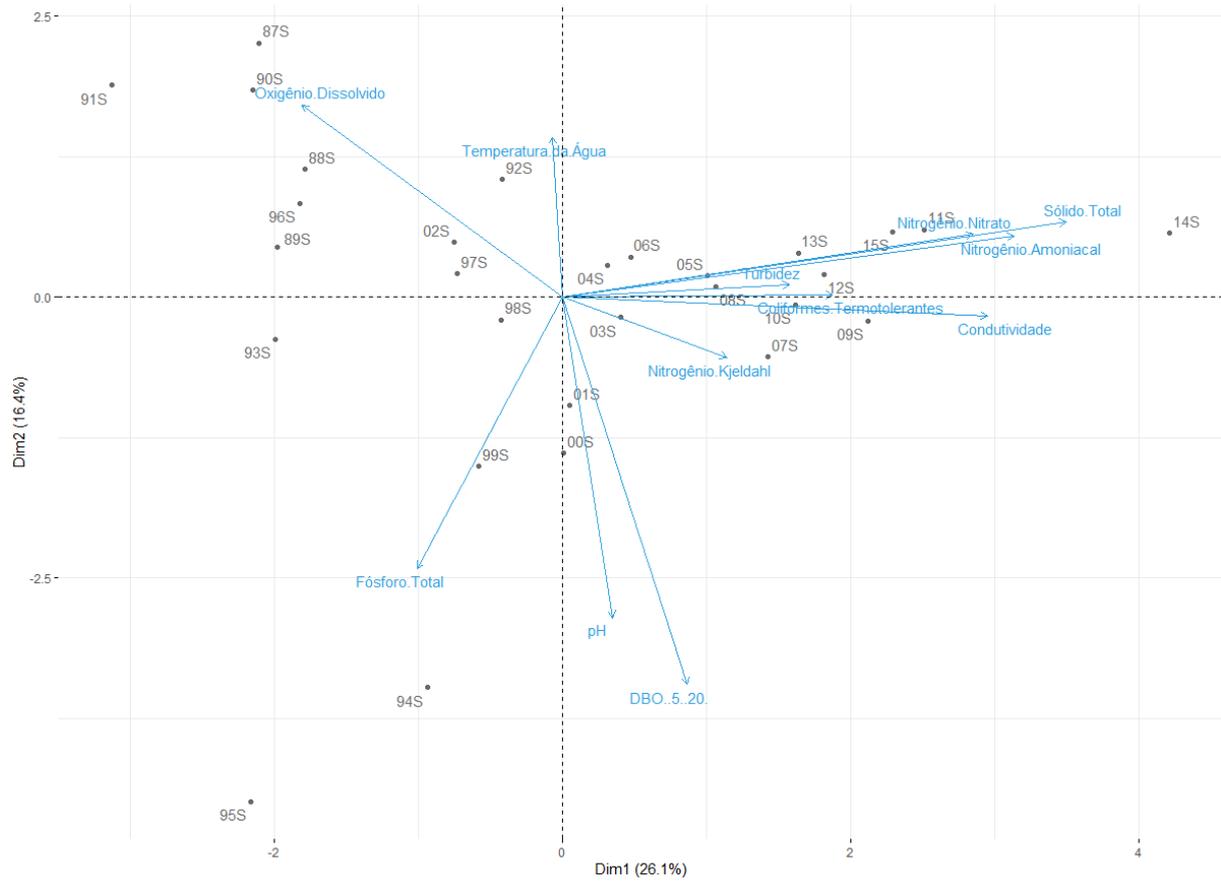
Tabela 5 – Correlação entre as variáveis e as componentes principais – período chuvoso

Variáveis	PC1	PC2	PC3	PC4	PC5
Coliformes Termotolerantes	0,449	-0,753	0,225	-0,345	-0,072
Condutividade	0,754	0,382	0,257	0,057	0,276
DBO	0,242	0,473	0,460	-0,486	-0,016
Fósforo Total	-0,612	0,163	0,099	-0,287	0,416
Nitrogênio Amoniacal	0,741	0,276	0,182	0,352	0,037
Nitrogênio Kjeldahl	0,638	-0,222	0,113	0,533	-0,020
Nitrogênio Nitrato	0,587	0,359	-0,565	-0,132	-0,183
Oxigênio Dissolvido	-0,440	0,114	0,702	0,308	0,223
pH	0,289	-0,063	-0,368	-0,091	0,806
Sólido Total	0,851	0,174	0,029	-0,177	0,091
Temperatura da Água	0,032	0,804	0,160	-0,188	-0,248
Turbidez	0,504	-0,682	0,325	-0,301	-0,059

Fonte: Autoria própria

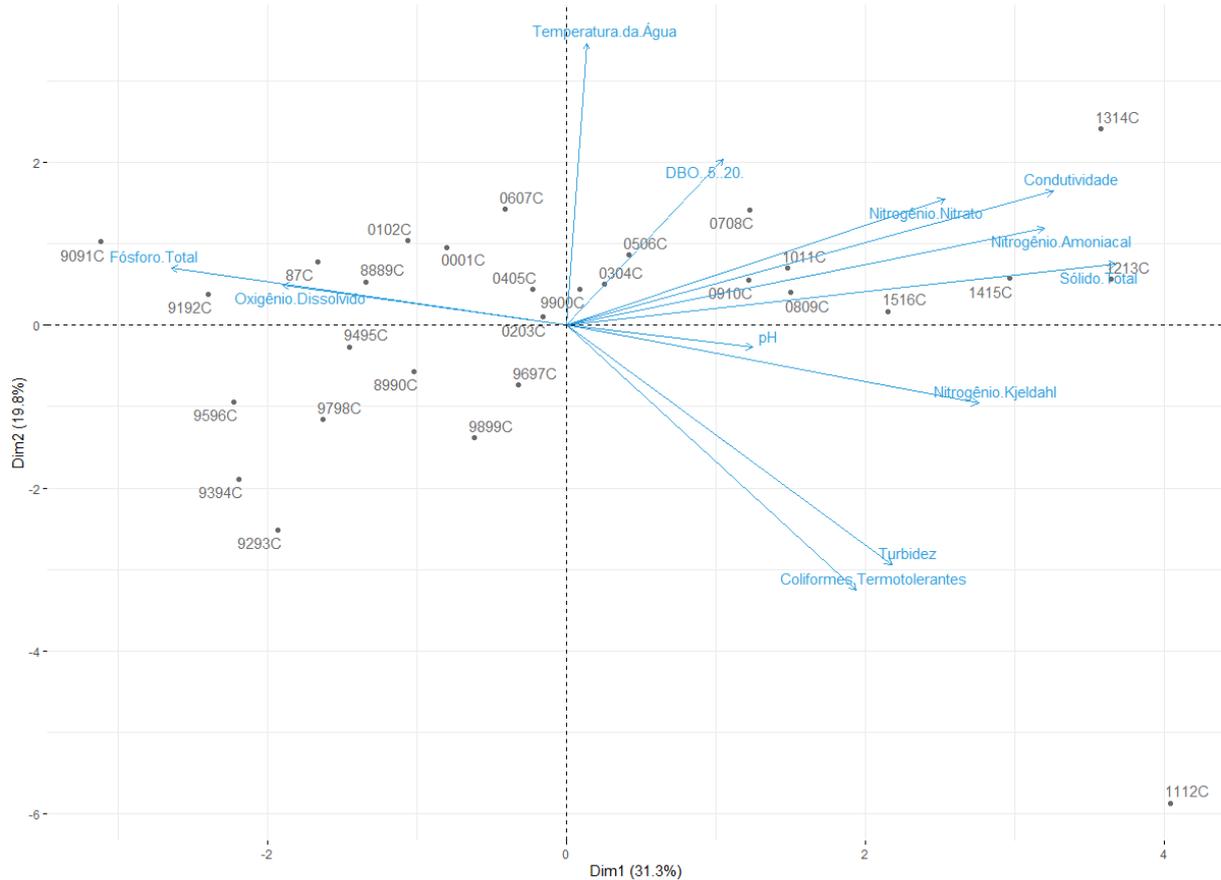
Os gráficos obtidos da PCA (Figuras 4 e 5) apresentam como vetores as variáveis de qualidade da água e os pontos representam as medianas dos períodos seco e chuvoso ao longo do período estudado.

Figura 4 – CP1 e CP2 incluindo as 12 variáveis e 29 observações do período seco



Fonte: Autoria própria

Figura 5 – CP1 e CP2 incluindo as 12 variáveis e 29 observações do período chuvoso



Fonte: Autoria própria

No período seco, para a primeira CP, as variáveis que mais se destacaram foram nitrogênio amoniacal, nitrato, condutividade e sólido total. Verifica-se que tais variações foram mais evidentes nos últimos anos da pesquisa, principalmente o período seco de 2011 e 2014 (Figura 4). Tais variações são indicativos compatíveis com descarga de poluentes, sendo a presença de nitrogênio amoniacal a variável mais evidente para os anos citados, indicando a poluição mais recente e o nitrato, sendo a forma oxidada, indica a poluição mais distante (SPERLING, 1996).

Naturalmente, com a presença de uma carga poluidora, as partículas presentes refletirão em maior quantidade de sólidos totais; sendo que a carga na forma de íons na água será verificada em aumento nos valores de condutividade. Deve-se ressaltar que a análise de componentes principais, embora evidencie a variabilidade da variável nitrato, deve ser interpretada com cautela, uma vez que seus valores estão bem

abaixo do limite da legislação, mas o aumento significativo a partir de 1997, foi suficiente para que a técnica a destacasse na primeira componente principal.

As variáveis DBO e pH apresentaram maior correlação com a segunda CP. A DBO apresentou as maiores variações nos anos de 1994 (5mg/L) e 1995 (6mg/L) (Figura 4), que coincidem com o período em que o pH esteve mais elevado, ultrapassando o limite da legislação, correspondente a 3mg/L. Ainda que a DBO tenha se destacado em uma CP apenas no período seco, não se pode afirmar que se trata de um fenômeno sazonal, uma vez que a variabilidade ocorreu de forma pontual. Embora tenha havido um destaque para o pH, sua variação ao longo do período seco ocorreu dentro da faixa considerada neutra.

A terceira CP, assim como a quinta CP, não apresentaram variáveis com correlações próximas ou acima de $|0,7|$. Tais resultados exigem cautela na interpretação, conforme citado por (FRANÇA, 2009), para que a interpretação não seja artificialmente formulada, não qual não se contribui para novas informações. Faz-se importante considerar a variável Fósforo total que, ainda que não tenha apresentado uma correlação mínima de $|0,7|$ em nenhuma CP, apresentou-se na maioria das coletas do período seco além dos limites do CONAMA. Segundo Cucio (2014), ao estudar a carga máxima admissível de fósforo (P) no reservatório Tanque Grande, calculou o tempo de retenção, que considera a vazão média do corpo hídrico, igual a 18 dias, e considerou o limite de 0,025mg P/L, de acordo com o estabelecido na legislação para classe 1. Essa falta de evidência pode ser considerada devido às variações dos valores de fósforo não terem sido tão significativas, ainda que estivessem em sua maior parte além de 0,025mg/L.

A quarta componente apresentou correlação significativa apenas com a variável temperatura da água, o que pode ser explicado pelo período seco do ano de 2015 ter apresentado uma temperatura atípica para o período estudado, sendo a maior do conjunto de dados.

No período chuvoso, as variáveis que apresentaram correlação mais significativa com a primeira CP foram: condutividade, nitrogênio amoniacal e sólidos totais. Verifica-se que tanto nesse período como no seco, com exceção do nitrato, tais variáveis apresentaram-se de maneira mais evidente em termos de variação conjunta na primeira componente. As correlações com relação às observações abrangem os anos de 2011 a 2014, o que inclui os anos evidenciados para o período seco. Tais resultados, verificados na Figura 5, indicam uma interação dessas variáveis e mais

fortemente relacionadas a um evento temporal do que sazonal, no sentido de queda da qualidade da água do manancial Tanque Grande ao longo do tempo.

Com relação à segunda CP do período chuvoso, embora a variável turbidez tenha apresentado correlação de $|0,682|$, será incluída na análise por ter apresentado elevada correlação inicial com a variável coliformes termotolerantes (0,906). Assim, ambas apresentaram correlação positiva e a temperatura da água indicou uma correlação negativa. As duas primeiras variáveis correlacionaram-se no sentido de que a turbidez pode indicar partículas decorrentes de descargas de efluentes na água acompanhadas da presença de coliformes. Para essa relação evidenciada no período chuvoso, deve-se considerar que a precipitação pode contribuir para o carreamento de matéria orgânica para os corpos d'água. Uma vez que a variável precipitação não foi incluída no conjunto de dados, tal conclusão deve ser realizada com cautela. Embora temperaturas elevadas tendam a propiciar o crescimento bacteriano, a correlação negativa pode ter sido influenciada pelas temperaturas mais amenas em 2011 e 2012 (Figura 5), exatamente quando se evidenciou os maiores valores para coliformes termotolerantes, indicando que a temperatura não influenciou para o controle do crescimento desses organismos.

O oxigênio dissolvido foi correlacionado isoladamente com a terceira CP. Embora faltem elementos para se compreender a causa da variação de oxigênio dissolvido, a sua evidência em uma das componentes se deve ao aumento de OD entre 1990 e 1991 e diminuição abaixo do limite da legislação do CONAMA 357/05 (BRASIL, 2005) entre 2008 e 2010.

A quarta componente não evidenciou nenhuma variável de maneira significativa.

Para a quinta componente do período chuvoso, destaca-se apenas o pH e, da mesma forma que observado para o período seco, o pH apresentou-se dentro da faixa aceitável pela legislação, com maiores variações entre 1992 e 1993.

5 CONCLUSÃO

A primeira componente principal, tanto no período chuvoso como no seco, apresentou as variáveis condutividade, nitrogênio amoniacal e sólidos totais correlacionadas entre si, incluindo o nitrato para o segundo período, abrangendo a maior variabilidade para se avaliar a qualidade da água do manancial Tanque Grande. Tais resultados ainda apontaram para um comportamento temporal e não sazonal para o intervalo de estudo entre 1987 e 2016, sendo os últimos anos mais correlacionados às variáveis mencionadas, indicando para uma queda na qualidade da água ao longo do tempo.

Os resultados obtidos da ACP devem ser interpretados com cautela com relação à qualidade de água do manancial, uma vez que a técnica evidencia a variabilidade dos parâmetros de qualidade, mas não necessariamente atribuindo uma classificação ao manancial, uma vez que o conjunto de dados pode estar de acordo com os limites da legislação.

Deve-se considerar que, para melhor compreensão da qualidade da água do manancial Tanque Grande, mais variáveis poderiam ter sido avaliadas no conjunto, mas por terem passado a ser monitoradas em anos posteriores ao período inicial dos dados utilizados, não puderam ter sido incluídas no presente estudo.

De modo geral, o estudo foi capaz de caracterizar resumidamente a qualidade da água do reservatório Tanque Grande nesse período, levando-se em consideração as estações seca e chuvosa.

REFERÊNCIAS

- À MEDIDA que a população mundial atinge 8 bilhões de pessoas, ONU pede solidariedade no avanço do desenvolvimento sustentável para todos. **United Nations Population Fund**, 2022. Disponível em: <https://brazil.unfpa.org/pt-br/news/a-medida-que-populacao-mundial-atinge-8-bilhoes-de-pessoas-onu-pede-solidariedade-no-avan%C3%A7o-do#:~:text=NOVA%20YORK%2C%20Estados%20Unidos%20%E2%80%93%20De,aumentaram%20a%20expectativa%20de%20vida>. Acesso em: 11 dez. 2022.
- AGÊNCIA NACIONAL DAS ÁGUAS - ANA. **Atlas Águas (2021): segurança hídrica do abastecimento urbano**, 2021. Painel de Indicadores, Investimentos e Croquis. Guarulhos. Disponível em: https://portal1.snirh.gov.br/arquivos/Croquis_SNIRH/3CC_3518800_Guarulhos.pdf. Acesso em: 23 out. 2022
- AZHAR, S. C. *et al.* Classification of river water quality using multivariate analysis. **Procedia Environmental Sciences**, v. 30, p. 79–84, 2015.
- BENGRÄINE, K.; MARHABA, T. F. Using principal component analysis to monitor spatial and temporal changes in water quality. **Journal of Hazardous Materials**, v. B 100, p. 179–195, 2003.
- BRASIL. Fundação Nacional de Saúde – FUNASA. **Manual prático de Análise de água**. Brasília: FUNASA, 4. ed., 2013. 150 p.
- BRASIL. Ministério do Meio Ambiente. Resolução CONAMA nº 357, de 15 de junho de 2005. Alterada pela Resolução 410/2009 e pela 430/2011. 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**, nº 053, de 18 mar. 2005, p. 58-63.
- BRASIL. Portaria GM/MS nº 888, de 4 de maio de 2021. Altera o Anexo XX da Portaria de Consolidação GM/MS nº 5, de 28 de setembro de 2017, para dispor sobre os procedimentos de controle e de vigilância da qualidade da água para consumo humano e seu padrão de potabilidade. **Diário Oficial da União**, ed. 85, Seção 1. p. 127.
- CARRASCO, G. *et al.* Water quality evaluation through a multivariate statistical HJ-Biplot approach. **Journal of Hydrology**, v. 577, 2019.
- CECCONELLO, S. T.; CENTENO, L. N.; GUEDES, H. A. S. Índice de qualidade de água modificado pela análise multivariada: estudo de caso do Arroio Pelotas, RS, Brasil. **Engenharia Sanitária e Ambiental**, v. 23, n. 5, p. 973–978, set. e out. 2018.
- COMPANHIA AMBIENTAL DO ESTADO DE SÃO PAULO - CETESB. **Apêndice C: Significado ambiental e sanitário das variáveis de qualidade das águas e dos sedimentos e metodologias analíticas e de amostragem**, 2021. São Paulo: CETESB, 2022. 57 p.

COMPANHIA AMBIENTAL DO ESTADO DE SÃO PAULO - CETESB. **Infoáguas**, 2022. Disponível em: <https://sistemainfoaguas.cetesb.sp.gov.br/>. Acesso em: 09 ago. 2022.

COMPANHIA DE SANEAMENTO BÁSICO DO ESTADO DE SÃO PAULO - SABESP. Convênio de cooperação município de Guarulhos SSRH 1.747.281/18 assinado em 09/11/18. 2018. Disponível em: https://site.sabesp.com.br/site/uploads/file/investidores/convenio_guarulhos.pdf. Acesso em: 08 out. 2022.

CONSELHO ESTADUAL DE SANEAMENTO – CONESAN. **Plano municipal de saneamento básico de Guarulhos**. Planejamento dos serviços de abastecimento de água e esgotamento sanitário. 2018. Disponível em: <https://www.infraestruturameioambiente.sp.gov.br/conesan/planos-de-saneamento-basico/#pmsb-g>. Acesso em: 07 set. 2022.

CUCIO, M. S. **Análise do estado trófico e carga máxima admissível de fósforo no reservatório Tanque Grande, Guarulhos – SP**. 2014. Dissertação (Mestrado em Engenharia Civil) – Escola Politécnica da Universidade de São Paulo, São Paulo, 2014.

FRANÇA, M. S. **análise estatística multivariada dos dados de monitoramento de qualidade de água da bacia do Alto Iguaçu: uma ferramenta para a gestão de recursos hídricos**. 2009. Dissertação (Mestrado em Engenharia) – Universidade Federal do Paraná, Paraná, 2009.

Fundação Sistema Estadual de Análise de Dados - SEADE. Seade população. Guarulhos: 2022. Disponível em: <https://populacao.seade.gov.br/>. Acesso em: 01 nov. 2022.

GUARULHOS. Prefeitura de Guarulhos. **Plano Diretor de Turismo do Município de Guarulhos (PDTUR)**. Caderno 1: Apresentação e metodologia. Guarulhos, 2017. 60 p. Disponível em: <https://www.guarulhos.sp.gov.br/sites/default/files/file/arquivos/CADERNO%20I.pdf>. Acesso em: 05 dez. 2022.

INSTITUTO DE PESQUISAS TECNOLÓGICAS - IPT. **Relatório Técnico**. Relatório Final. São Paulo: Centro de Tecnologias Geoambientais - CTGeo. Laboratório de Recursos Hídricos e Avaliação Geoambiental – Labgeo, 2014, n. 131, v. 3, p. 578-205.

JOHNSON, R.A.; WICHERN, D.W. Principal Components. In: JOHNSON, R.A.; WICHERN, D.W. **Applied multivariate statistical analysis**. 6. ed. Upper Saddle River: Prentice Hall, 2007. 808 p.

MAHAPATRA, S. S. *et al.* Prediction of water quality using principal component analysis. **Water Quality, Exposure and Health**, v. 4, n. 2, p. 93–104, jun. 2012.

MAZLUM, N.; ÖZER, A.; MAZLUM, S. Interpretation of water quality data by principal components analysis. **Tr. J. of Engineering and Environmental Science**, v. 23, p. 19-26, 1999.

MINGOTI, S. A. Análise de Componentes Principais. In: MINGOTI, S. A. **Análise de dados através de métodos de estatística multivariada**. Uma abordagem aplicada. Belo Horizonte: Editora UFMG, 2005, 281 p.

OBRAS garantem a melhoria contínua do abastecimento de água em Guarulhos. **Prefeitura de Guarulhos**, 2022. Disponível em: <https://www.guarulhos.sp.gov.br/article/obras-garantem-melhoria-continua-do-abastecimento-de-agua-em-guarulhos>. Acesso em: 23 out. 2022.

PIASENTIN, A. M. *et al.* Índice de qualidade da água (IQA) do reservatório Tanque Grande, Guarulhos (SP): análise sazonal e efeitos do uso e ocupação do solo. **Geociências**, v. 8, n. 3, p. 305-317, 2009.

VARGAS, R. R. *et al.* A contribuição do município de Guarulhos para a qualidade da água do rio baquirivu-guaçu. In: SIMPÓSIO BRASILEIRO DE RECURSOS HÍDRICOS, 21., 2015, Brasília: ABRH, 2015.

VON SPERLING, M. Noções de qualidade das águas. In: VON SPERLING, M. **Introdução à qualidade das águas e ao tratamento de esgotos**. 2. ed. Belo Horizonte: Departamento de Engenharia Sanitária e Ambiental, Universidade Federal de Minas Gerais, 1996, v. 1, 243 p.

WEATHER SPARK. **Weather Spark**, 2022. Clima e condições meteorológicas médias em Guarulhos no ano todo. Disponível em: <https://pt.weatherspark.com/y/30292/Clima-caracter%C3%ADstico-em-Guarulhos-Brasil-durante-o-ano#Sections-Humidity>. Acesso em: 23 out. 2022.