

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

Tiago Antônio Figueiredo

**DETERMINAÇÃO DE CLOROFILA-A UTILIZANDO SENSOR HIPERESPECTRAL E  
ANÁLISE DA QUALIDADE DA ÁGUA DA LAGOA DA PAMPULHA-MG**

Belo Horizonte - MG  
2023

Tiago Antônio Figueiredo

**DETERMINAÇÃO DE CLOROFILA-A UTILIZANDO SENSOR HIPERESPECTRAL E  
ANÁLISE DA QUALIDADE DA ÁGUA DA LAGOA DA PAMPULHA-MG**

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

Área de concentração: Meio Ambiente

Linha de pesquisa: Caracterização, prevenção e controle da poluição

Orientador: Camila Costa Amorim

F475d	<p>Figueiredo, Tiago Antônio.  Determinação de clorofila-a utilizando sensor hiperespectral e análise da qualidade da água da Lagoa da Pampulha - MG [recurso eletrônico] / Tiago Antônio Figueiredo. – 2023.  1 recurso online (132 f. : il., color.) : pdf.</p> <p>Orientadora: Camila Costa de Amorim Amaral.</p> <p>Dissertação (mestrado) - Universidade Federal de Minas Gerais, Escola de Engenharia.</p> <p>Apêndices: f. 121-132.</p> <p>Bibliografia: f. 111-120.  Exigências do sistema: Adobe Acrobat Reader.</p> <p>1. Engenharia sanitária - Teses. 2. Meio ambiente - Teses.  3. Água - Poluição - Teses. 4. Eutrofização - Teses. 5. Lagos - Teses.  6. Modelagem de dados - Teses. 7. Sensoriamento remoto - Teses.  I. Amaral, Camila Costa de Amorim. II. Universidade Federal de Minas Gerais. Escola de Engenharia. III. Título.</p>
	CDU: 628(043)



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

### **FOLHA DE APROVAÇÃO**

"Determinação de Clorofila-a Utilizando Sensor Hiperespectral e Análise da Qualidade da Água da Lagoa da Pampulha-MG"

**TIAGO ANTÔNIO FIGUEIREDO**

Dissertação defendida e aprovada pela banca examinadora constituída pelos Senhores:

Profa Camila Costa de Amorim Amaral

Profa Talita Fernanda das Graças Silva

Profa Silvia Maria Alves Corrêa Oliveira

Proif. Lino Augusto Sander de Carvalho

Aprovada pelo Colegiado do PG SMARH

Versão Final aprovada por

Profa. Priscilla Macedo Moura

Prof<sup>ª</sup>. Camila Costa de Amorim Amaral

Coordenadora

Orientadora

Belo Horizonte, 02 de junho de 2023.



Documento assinado eletronicamente por **Silvia Maria Alves Correa Oliveira, Professora do Magistério Superior**, em 06/06/2023, às 07:03, conforme horário oficial de Brasília, com fundamento no art. 5º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Talita Fernanda das Graças Silva, Professora do Magistério Superior**, em 06/06/2023, às 10:25, conforme horário oficial de Brasília, com fundamento no art. 5º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Lino Augusto Sander de Carvalho, Usuário Externo**, em 06/06/2023, às 13:27, conforme horário oficial de Brasília, com fundamento no art. 5º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Camila Costa de Amorim Amaral, Professora do Magistério Superior**, em 06/06/2023, às 13:55, conforme horário oficial de Brasília, com fundamento no art. 5º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Priscilla Macedo Moura, Coordenador(a) de curso de pós-graduação**, em 23/10/2023, às 09:57, conforme horário oficial de Brasília, com fundamento no art. 5º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



A autenticidade deste documento pode ser conferida no site [https://sei.ufmg.br/sei/controlador\\_externo.php?acao=documento\\_conferir&id\\_orgao\\_acesso\\_externo=0](https://sei.ufmg.br/sei/controlador_externo.php?acao=documento_conferir&id_orgao_acesso_externo=0), informando o código verificador **2352792** e o código CRC **1F95A384**.

## **AGRADECIMENTOS**

Agradeço, primeiramente, aos meus familiares pelo apoio e pelos momentos de distração em meio a esta caminhada.

Minha querida mãe Luciana Gonçalves de Souza que sempre esteve ao meu lado, quase sem reclamar e incondicionalmente, mesmo diante de todas as dificuldades, prestando todo carinho que só uma mãe pode dar. Prometo arrumar tudo que você quer na casa a partir de agora.

Ao meu pai Antônio Rosa Figueiredo por todos os milhares de casos contados nos momentos de pausa ao longo de todo este trabalho

Aos meus irmãos Tatiani Gonçalves Figueiredo e Thales Pedro de Figueiredo e ao meu cunhado Brendon Henrique Barbosa, por serem bem chatos. À minha sobrinha, Alana Gonçalves Barbosa, que só por sua existência deixa alegria por onde passa, pode me agradecer por isso quando aprender a ler.

À minha orientadora Prof. Dra. Camila Costa Amorim por toda a paciência, dedicação, paciência e apoio para conclusão deste trabalho.

À minha coorientadora Dra. Gabriela Rabelo Andrade, que mesmo de forma não oficial, foi essencial para a conclusão deste trabalho. Muito obrigado por todos os ensinamentos, paciência e dedicação que ficarão guardados comigo para sempre.

Aos colegas do GruPOA e do GT-607, que também me auxiliaram nesta caminhada. Agradeço também à CEMIG e à prefeitura de Belo Horizonte que forneceram meios para a realização deste projeto de pesquisa.

À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) por todo o suporte e amparo concedido para o desenvolvimento desta pesquisa.

À Universidade Federal de Minas Gerais e ao Programa de Pós-Graduação de Saneamento, Meio Ambiente e Recursos Hídricos, por todo o apoio e conhecimento oferecidos.

E por fim, a todos que torceram por mim durante todo este processo!

## RESUMO

Monitorar ambientes aquáticos em bacias hidrográficas urbanas é essencial para se conhecer a dinâmica da qualidade da água e os efeitos da poluição incidente. O objetivo deste estudo foi analisar a utilização de um sensor hiperespectral para monitorar clorofila-a (chl-a) na Lagoa da Pampulha, onde o aporte de sedimentos e de efluentes industriais e domésticos provocam significativos impactos. Foi utilizada a metodologia empírica, onde os algoritmos *Support Vector Regressor* (SVR), *Random Forest Regressor* (RFR) e *XGBooster* (XGB) e regressão linear (RL) foram testados na construção do modelo para estimar as concentrações de chl-a no reservatório. Experimentos em laboratório foram realizados para validar a metodologia de funcionamento do sensor, que foi utilizado para estimativa da concentração de chl-a, através da alga *Raphidocelis subcapitata*. Análises estatísticas uni e multivariadas foram aplicadas aos dados secundários e primários para caracterizar, de forma espacial e temporal, a qualidade das águas do reservatório. A partir das análises estatísticas dos dados secundários, entre 2013 e 2021, constatou-se que chl-a, densidade de cianobactérias e fósforo total ultrapassaram significativamente o limite estabelecido pela legislação, principalmente no ponto amostral a montante do reservatório onde deságuam os principais afluentes, os córregos Ressaca e Sarandi. Mesmo com as intervenções com os produtos *Phoslock®* e *Enzilimp®*, a partir de 2016, não houve redução dos níveis de fósforo, principal foco do tratamento, parâmetro este, que é apontado como a principal influência para a alta taxa de produção primária na lagoa. Os dados primários coletados mensalmente, entre março de 2022 e fevereiro de 2023, revelaram que a região próxima aos principais tributários detém os piores índices de qualidade da água. A cortina retentora de resíduos, instalada em 2016, também apresentou influência, com a região a montante expondo piores índices de poluição. A proximidade dos principais tributários e a menor profundidade da coluna da água na região são as possíveis causas destes resultados. A análise de cluster comprovou tais resultados e indicou a presença de um quarto compartimento na lagoa, além das três regiões representadas pelos pontos de monitoramento do IGAM, que se encontra logo após a cortina, o que pode denotar a necessidade de ampliação dos pontos monitorados no reservatório. Na análise de laboratório foi observada alta relação entre as imagens captadas pelo sensor e a chl-a ( $R^2 = 0,9942$  e  $RMSE = 3,42 \mu\text{g/L}$ ), utilizando a regressão de mínimos quadrados parciais. Na análise de campo, dentre os modelos observados, o algoritmo que melhor se adaptou aos pares de dados, imagens hiperespectrais e concentrações de chl-a, foi o SVR ( $R^2 = 0,81$  e  $RMSE = 6,87 \mu\text{g/L}$ ). O sensor hiperespectral apresentou boa performance para estimar tanto as amostras mono específicas de algas, quanto para prever chl-a em um ambiente com alto dinamismo de suas concentrações de sedimentos e compostos dissolvidos, representando uma alternativa para monitoramento de chl-a na Lagoa da Pampulha. Entretanto, são necessários estudos complementares para analisar se os resultados ocorreram em função das ações do poder público, pois a dinâmica dos parâmetros, principalmente do fitoplâncton, pode ser influenciada por diversos fatores, como as alterações climáticas.

Palavras-chave: Modelagem de dados. Lagos urbanos. Eutrofização. Sensoriamento remoto. Poluição hídrica.

## ABSTRACT

Monitoring aquatic environments in urban watersheds is essential to understand water quality dynamics and the effects of incident pollution. The objective of this study was to analyze the use of a hyperspectral sensor to monitor chlorophyll-a (chl-a) in Pampulha Lagoon, where sediment inputs and industrial and domestic effluents cause significant impacts. The empirical methodology was employed, where the Support Vector Regressor (SVR), Random Forest Regressor (RFR), XGBoost (XGB), and linear regression (LR) algorithms were tested in constructing a model to estimate chl-a concentrations in the reservoir. Laboratory experiments were conducted to validate the sensor's methodology, which was used to estimate chl-a concentration through the *Raphidocelis subcapitata* algae. Uni- and multivariate statistical analyses were applied to secondary and primary data to characterize the spatial and temporal water quality of the reservoir. Based on the statistical analyses of secondary data between 2013 and 2021, it was found that chl-a, cyanobacteria density, and total phosphorus significantly exceeded the limits established by legislation, especially at the upstream sampling point where the main tributaries, Ressaca and Sarandi streams, discharge into the reservoir. Despite interventions with Phoslock® and Enzilimp® products since 2016, there was no reduction in phosphorus levels, the main focus of the treatment, which is considered the main influence on the high rate of primary production in the lagoon. The monthly primary data collected between March 2022 and February 2023 revealed that the region near the main tributaries had the worst water quality indices. The waste retention curtain installed in 2016 also had an influence, with the upstream region exhibiting worse pollution indices. The proximity of the main tributaries and the shallower water column in the region are possible causes of these results. Cluster analysis confirmed these findings and indicated the presence of a fourth compartment in the lagoon, in addition to the three regions represented by the IGAM monitoring points, just after the curtain, which may indicate the need to expand the monitored points in the reservoir. In the laboratory analysis, a high correlation was observed between the images captured by the sensor and chl-a ( $R^2 = 0.9942$  and  $RMSE = 3.42 \mu\text{g/L}$ ) using partial least squares regression. In the field analysis, among the observed models, the algorithm that best adapted to the data pairs, hyperspectral images, and chl-a concentrations, was SVR ( $R^2 = 0.81$  and  $RMSE = 6.87 \mu\text{g/L}$ ). The hyperspectral sensor demonstrated good performance in estimating both monospecific algal samples and predicting chl-a in an environment with high dynamics of sediment and dissolved compound concentrations, representing an alternative for chl-a monitoring in Pampulha Lagoon. However, further studies are needed to analyze whether the results occurred due to the actions of the government, as the dynamics of parameters, especially phytoplankton, can be influenced by various factors such as climate change.

Keywords: Data modeling. Urban lakes. Eutrophication. Remote sensing. Water pollution.

## LISTA DE FIGURAS

<b>Figura 1</b> - Trajetória da luz incidente (solar) em um corpo hídrico até atingir o sensor. Demonstrando o espalhamento pela atmosfera (p1), skylint (p2), sunglint (p3), radiância da coluna de água (p4), absorção e espalhamento pelo fitoplâncton (p5), espalhamento pelo sedimento (p6) e absorção pelo CDOM (p7).....	32
<b>Figura 2</b> - Comparação entre um espectro contínuo retirado de uma imagem hiperespectral (a) e o espectro de uma imagem multiespectral (b).....	38
<b>Figura 3</b> - Localização geográfica da Lagoa da Pampulha e subdivisão de suas sub-bacias hidrográficas .....	45
<b>Figura 4</b> - Fluxograma da metodologia utilizada no presente estudo .....	49
<b>Figura 5</b> - Pontos amostrais monitorados pelo IGAM, pontos utilizados para o monitoramento com a sonda multiparâmetros e os pontos selecionados para o monitoramento com o sensor hiperespectral. Além destes, observa-se a localização da cortina retentora de resíduos, na região de transição da lagoa.....	53
<b>Figura 6</b> - Algas <i>Raphidocelis supcapitata</i> utilizadas no experimento de laboratório para determinação de clorofila-a.....	56
<b>Figura 7</b> - Aparato experimental para as coletas das curvas espectrais realizadas em laboratório (a) e resultado do teste de irradiância da lâmpada de halogênio (b).....	58
<b>Figura 8</b> - Disposição do sensor na embarcação para coleta das imagens hiperespectrais.....	60
<b>Figura 9</b> - Box-plot com a distribuição (primeiro quartil, segundo quartil, mediana, terceiro quartil e outliers) dos dados analisados no presente estudo para os pontos PV110, PV115, PV210, PV230, PV235 e PV240 para o período analisado (2013 a 2021), subdivididos em período chuvoso e seco. Destaca-se ainda os resultados dos testes de Kruskal-Wallis, onde diferentes letras na representação de cada ponto representam diferença significativamente estatística entre as medianas dos pontos para o parâmetro.....	66
<b>Figura 10</b> - Gráfico temporal das concentrações de clorofila-a (a), densidade de cianobactérias (b) e fósforo total (c) nos pontos PV230, PV235 e PV240 para o período analisado (2013 a 2021), subdivididos em período chuvoso e seco .....	73
<b>Figura 11</b> - <i>Box-plot</i> com a distribuição dos dados de chl-a nos 30 pontos amostrais (entre março/2022 e fevereiro/2023), com o limite estabelecido pela Resolução CONAMA 357/05 e a representação da cortina retentora de resíduos (a).	

Espacialização dos valores medianos de chl-a realizada por meio de interpolação com Krigagem ordinária, subdivididos em período chuvoso (b) e seco (c) .....	79
<b>Figura 12</b> - Box-plot com a distribuição dos dados de ficocianina nos 30 pontos amostrais (entre março/2022 e fevereiro/2023), com a representação da cortina retentora de resíduos (a). Espacialização dos valores medianos de ficocianina realizada por meio de interpolação com Krigagem ordinária, subdivididos em período chuvoso (b) e seco (c) .....	81
<b>Figura 13</b> - <i>Box-plot</i> com a distribuição dos dados de pH nos 30 pontos amostrais (entre março/2022 e fevereiro/2023), com o limite estabelecido pela Resolução CONAMA 357/05 e a representação da cortina retentora de resíduos (a). Espacialização dos valores medianos de pH realizada por meio de interpolação com Krigagem ordinária, subdivididos em período chuvoso (b) e seco (c) .....	83
<b>Figura 14</b> - <i>Box-plot</i> com a distribuição dos dados de OD nos 30 pontos amostrais (entre março/2022 e fevereiro/2023), com o limite estabelecido pela Resolução CONAMA 357/05 e a representação da cortina retentora de resíduos (a). Espacialização dos valores medianos de OD realizada por meio de interpolação com Krigagem ordinária, subdivididos em período chuvoso (b) e seco (c) .....	85
<b>Figura 15</b> - <i>Box-plot</i> com a distribuição dos dados de turbidez nos 30 pontos amostrais (entre março/2022 e fevereiro/2023), com o limite estabelecido pela Resolução CONAMA 357/05 e a representação da cortina retentora de resíduos (a). Espacialização dos valores medianos de turbidez realizada por meio de interpolação com Krigagem ordinária, subdivididos em período chuvoso (b) e seco (c) .....	86
<b>Figura 16</b> - <i>Box-plot</i> com a distribuição dos dados de disco de Secchi nos 30 pontos amostrais (entre março/2022 e fevereiro/2023), com a representação da cortina retentora de resíduos (a). Espacialização dos valores medianos de disco de Secchi realizada por meio de interpolação com Krigagem ordinária, subdivididos em período chuvoso (b) e seco (c) .....	88
<b>Figura 17</b> - Dendrograma da análise de agrupamento dos parâmetros coletados <i>in situ</i> na lagoa da Pampulha .....	90
<b>Figura 18</b> - Espectros de reflectância obtidos através das imagens hiperespectrais das diferentes concentrações de chl-a e da água destilada, captadas no experimento de laboratório .....	92

<b>Figura 19</b> - Exemplos de 4 pixels extraídos de uma imagem do experimento de laboratório, expondo o elevado índice de ruídos nos primeiros comprimentos de ondas captados pelo sensor .....	92
<b>Figura 20</b> - Espectro de reflectância hiperespectral da alga <i>T. Raphidocelis</i> , captada em experimentos de laboratório .....	94
<b>Figura 21</b> - Visualização dos resultados da validação cruzada mensurando o número de componentes ótimos encontrados na análise de PLS para o modelo (a), o gráfico com os resultados de validação cruzada da calibração (b) e para o teste do modelo utilizando apenas as componentes principais selecionadas na PLS (c). .....	96
<b>Figura 22</b> - Espectros de reflectância hiperespectral obtidos através das imagens coletadas em campo na lagoa da Pampulha (a), dados de reflectância obtidos após a retirada dos comprimentos de onda afetados por ruídos (b) e reflectância espectral final, obtida após retirada de dados anormais (c).....	97
<b>Figura 23</b> - Gráfico de dispersão entre os valores mensurados e estimados e os resultados das métricas de validação para o modelo SVR .....	103
<b>Figura 24</b> - Reflectância média dos espectros utilizados nos algoritmos (Figura 22c) e a representação das bandas preditoras selecionadas pelo método <i>forward</i>	105

## LISTA DE TABELAS

<b>Tabela 1</b> - Relação dos parâmetros avaliados no presente estudo e seus limites para as classes 1, 2 e 3, segundo a resolução CONAMA 357/2005 e portaria 05/2017 do Ministério da Saúde (microcistina e saxitoxina).....	<b>26</b>
<b>Tabela 2</b> - Divisão do espectro eletromagnético, com destaque para os comprimentos de ondas da luz visível .....	<b>30</b>
<b>Tabela 3</b> - Estudos utilizando sensores hiperespectrais, detalhando o sensor, a faixa do espectro imageada, número de bandas, parâmetros utilizados, modo de aquisição e resultados obtidos .....	<b>39</b>
<b>Tabela 4</b> - Estudos relatando as faixas espectrais de absorção e reflectância da clorofila-a em ambientes aquáticos e as formas de aquisição das imagens .....	<b>41</b>
<b>Tabela 5</b> - Relação e descrição dos modelos e suas descrições utilizados na etapa de estimativa da concentração de clorofila-a com dados hiperespectrais de campo .....	<b>44</b>
<b>Tabela 6</b> - Concentrações de chl-a utilizadas na captura das imagens hiperespectrais e posterior construção do modelo de regressão .....	<b>57</b>
<b>Tabela 7</b> - Resultado para o teste de Mann-Whitney comparando as concentrações dos parâmetros observados antes e depois da cortina retentora de resíduos na Lagoa da Pampulha .....	<b>78</b>
<b>Tabela 8</b> - Estatística descritiva para os dados de chl-a em µg/L coletados simultaneamente com as imagens hiperespectrais para construção dos modelos .....	<b>101</b>
<b>Tabela 9</b> - Resultados dos modelos para o teste C, que utilizou todos os meses, demonstrando os valores de R <sup>2</sup> para o treino, teste e validação e RMSE de validação para cada tipo de entrada dos dados utilizada, com destaque para o melhor resultado observado.....	<b>102</b>

## LISTA DE ABREVIATURAS E SIGLAS

ABNT - Associação Brasileira de Normas Técnicas

ABNT - Associação Brasileira de Normas Técnicas

AC - Análise de Cluster

CCC - Coeficiente de Correlação Cofenética

CDOM - Matéria Orgânica Dissolvida Colorida

CEMIG - Companhia Energética de Minas Gerais

CERH-MG - Conselho Estadual de Recursos Hídricos do Estado de Minas Gerais

chl-a - Clorofila-a

cm - Centímetro

CO<sub>2</sub> - Dióxido de Carbono

COAs - Componentes opticamente ativos da água

CONAMA - Conselho Nacional do Meio Ambiente

COPAM-MG - Conselho Estadual de Política Ambiental do Estado de Minas Gerais

DBO - Demanda Bioquímica de Oxigênio

DC - Densidade de Cianobactérias

DESA-UFMG - Departamento de Engenharia Sanitária e Ambiental da Universidade Federal de Minas Gerais

DQO - Demanda Química de Oxigênio

ETAf - Estação de Tratamento de Águas Fluviais

IGAM - Instituto Mineiro de Gestão das Águas

km - Quilômetro

L - Litro

LabNS-UFMG - Laboratório de Nano-Espectroscopia da Universidade Federal de Minas Gerais

LANDSAT - Land Remote Sensing Satellite

LaPO<sub>4</sub> - Rabdofano

m - Metro

MAPE

mg - Miligrama

MODIS - Moderate Resolution Imaging Spectroradiometer

MS - Ministério da Saúde

N - Nitrogênio

N. Am. - Nitrogênio Amoniacal

NBR - Norma Brasileira

ND - Número Digital

NH<sub>4</sub>-N - Amônio

nm - Nanometros

NO<sub>3</sub>-N - Nitrato

O<sub>2</sub> - Oxigênio

OD - Oxigênio Dissolvido

pH - Potencial Hidrogeniônico

PLS - Partial Least Squares Regression

PT - Fósforo Total

R<sup>2</sup> - Coeficiente de determinação

REM - Radiação Eletromagnética

RFR - Random Forest Regressor

RL - Regressão Linear

RMSE - Root Mean Squared Error

Rrs - Reflectância de Sensoriamento Remoto

SR - Sensoriamento Remoto

SS - Sólidos Suspensos

SVR - Support Vector Regressor

UNESCO -United Nations Educational, Scientific and Cultural Organization

UNT - Unidade Nefelométricas de Turbidez

W - Watts

XGB - eXtreme Gradient Boosting

µg - Micrograma

µm - Micrômetro

## SUMÁRIO

1	INTRODUÇÃO .....	17
2	OBJETIVOS .....	22
2.1	Objetivo Geral .....	22
2.2	Objetivos Específicos .....	22
3	REVISÃO DA LITERATURA .....	23
3.1	Recursos hídricos em bacias hidrográficas urbanas .....	23
3.2	Monitoramento de recursos hídricos .....	24
3.2.1	Parâmetros de qualidade da água .....	25
3.3	Princípios do sensoriamento remoto .....	29
3.3.1	Sensoriamento remoto em ambientes aquáticos .....	31
3.4	Sensoriamento remoto aplicado ao monitoramento de recursos hídricos ....	33
3.5	Métodos de aquisição de dados no sensoriamento remoto .....	35
3.6	Sensoriamento remoto hiperespectral.....	37
3.6.1	Propriedades ópticas da clorofila-a .....	41
3.7	Processamento de dados.....	42
3.8	Estudo de caso: Lagoa da Pampulha.....	44
4	METODOLOGIA .....	49
4.1	Materiais.....	49
4.1.1	Sensor hiperespectral .....	49
4.1.2	Sonda multiparâmetros .....	51
4.2	Caracterização da área de estudo com dados históricos de monitoramento .....	51
4.3	Caracterização da qualidade da água na Lagoa da Pampulha através de medições obtidas pela sonda multiparâmetros .....	53
4.4	Desenvolvimento da metodologia de calibração e utilização do sensor hiperespectral em laboratório.....	56
4.4.1	Desenvolvimento do algoritmo e validação dos dados para o experimento em laboratório .....	58
4.5	Estimativa da concentração de clorofila-a através do sensor hiperespectral em campo e do desenvolvimento de modelos empíricos .....	60
4.5.1	Aquisição das imagens .....	60
4.5.2	Processamento das imagens coletadas em campo .....	61
4.5.3	Desenvolvimento dos algoritmos empíricos e validação dos dados para as imagens coletadas <i>in situ</i> .....	62
5	RESULTADOS E DISCUSSÃO .....	64
5.1	Caracterização da área de estudo com dados históricos de monitoramento .....	64

5.2	Caracterização da qualidade da água na lagoa da Pampulha através de medições obtidas com a sonda multiparâmetros.....	77
5.2.1	Clorofila-a.....	79
5.2.2	Ficocianina.....	81
5.2.3	pH.....	83
5.2.4	Oxigênio dissolvido.....	84
5.2.5	Turbidez.....	86
5.2.6	Profundidade do disco de Secchi.....	87
5.2.7	Análise de cluster.....	89
5.3	Desenvolvimento da metodologia de calibração e utilização do sensor hiperespectral em laboratório.....	91
5.3.1	Reflectância hiperespectral das imagens de laboratório.....	91
5.3.2	Estimativa da concentração de algas a partir da reflectância hiperespectral em laboratório.....	94
5.4	Estimativa da concentração de clorofila-a através do sensor hiperespectral em campo através do desenvolvimento de modelos empíricos.....	97
5.4.1	Reflectância hiperespectral das imagens coletadas em campo.....	97
5.4.2	Estimativa da concentração de clorofila-a a partir das imagens de campo do sensor hiperespectral.....	99
6	CONCLUSÕES.....	109
	REFERÊNCIAS.....	111
	APÊNDICE A – Coordenadas dos pontos amostrais monitorados no presente estudo e sua subdivisão em relação aos pontos monitorados pelo IGAM, pela sonda multiparâmetros e pela sonda multiparâmetros/sensor hiperespectral.....	121
	APÊNDICE B – Estatística descritiva dos dados de monitoramento do reservatório da Lagoa da Pampulha, referentes aos pontos PV230, PV235 e PV240. Subdivididos em período seco e chuvoso.....	122
	APÊNDICE C - Resultados do teste de Mann-Whitney para comparação das medianas dos parâmetros analisados entre os períodos seco e chuvoso nos pontos PV230, PV235 e PV240. Valor de $\alpha$ foi igual a 0,05.....	126
	APÊNDICE D - Resultados para o teste de Kruskal-Wallis para verificar diferenças significativamente estatísticas entre as medianas dos 3 pontos analisados para os períodos seco e chuvoso, complementado pelo teste <i>post-hoc</i> de Dunn (Pontos 1, 2 e 3 representam os pontos amostrais PV230, PV235 E PV240, respectivamente) ...	127
	APÊNDICE E - Resultado gráfico do número ótimo de clusters obtido através do método gap statistic.....	128
	APÊNDICE F - Resultados dos modelos para o teste com as bandas preditivas, expondo $R^2$ para o treino, teste e validação e RMSE de validação para cada tipo de entrada dos dados utilizada, com destaque para o melhor resultado observado....	129

## 1 INTRODUÇÃO

Alterações no uso do solo em áreas urbanas e rurais geram inúmeros impactos sobre os recursos hídricos comprometendo sua qualidade e inviabilizando seus usos previstos. O crescente lançamento de águas residuárias domésticas, industriais e o carreamento de sedimentos e resíduos em períodos chuvosos contribuem substancialmente para este processo em regiões urbanizadas (PAIVA et al., 2020). No meio rural a piora da qualidade da água ocorre, principalmente, pelo avanço da agricultura (AGUIAR; NETO; RANGEL, 2011). Além disso, em um cenário de escassez hídrica e crescente demanda por águas de abastecimento, estes problemas se tornam ainda mais relevantes (GIRI; QIU, 2016).

Em particular, estas alterações provocadas por ações antrópicas afetam a ciclagem de fósforo e nitrogênio. Escoamentos com altas concentrações destas substâncias alteram a dinâmica da produção primária em águas superficiais. Este fato ocorre porque estão relacionadas ao crescimento de algas e plantas aquáticas e quando disponíveis em quantidade excessiva podem levar à eutrofização, floração de algas e ao aparecimento de cianobactérias potenciais produtoras de toxinas, que podem afetar a biota e a saúde humana (SAGAN et al., 2020).

Diversos tipos de algas compõem o fitoplâncton e sua ecologia está diretamente relacionada com as características físicas, químicas e biológicas do corpo hídrico, sendo um componente importante para a compreensão destas variáveis e do estado de trofia do ambiente. A análise quantitativa do fitoplâncton pode ser obtida, dentre outras formas, pela estimativa da concentração de clorofila-a (chl-a), pigmento presente no fitoplâncton responsável pela fotossíntese (ESTEVES, 1998).

Entender como a poluição afeta os aspectos espaço-temporal dos recursos hídricos é de suma importância para a fiscalização, gerenciamento, proteção e melhoria dos ecossistemas aquáticos (BEHMEL et al., 2016). O monitoramento da qualidade da água ocorre, tradicionalmente, por coleta de amostras de água em locais e intervalos pré-determinados e padronizados, com posterior análise laboratorial e comparação dos resultados com padrões estabelecidos pelo poder público, através dos órgãos responsáveis (SÃO PAULO, 2012).

Entretanto, este processo operacional de monitoramento, apesar de apresentar elevada precisão, é trabalhoso e demorado, além de demandar altos custos devido aos recursos humanos especializados, insumos necessários para as análises e pela grande extensão dos sistemas a serem monitorados (GHOLIZADEH; MELESSE; REDDI, 2016).

Estes fatores resultam na baixa frequência de campanhas de monitoramento e número de pontos amostrais, os quais são questões críticas para a eficiência do monitoramento, pois os corpos d'água apresentam grande heterogeneidade espacial e temporal dos seus parâmetros de qualidade. A defasagem destes quesitos, devido aos baixos recursos orçamentários, ocasiona baixa eficiência do monitoramento. Com isto, o monitoramento pode não retratar todo o dinamismo apresentado pelos ecossistemas aquáticos, principalmente em bacias hidrográficas densamente povoadas onde a carga de poluição tende a mudar as características dos corpos d'água rapidamente (VILMIN et al., 2018).

Para minimizar tais problemas, o monitoramento ambiental de ambientes aquáticos utiliza diferentes ferramentas para possibilitar o monitoramento de grandes áreas em tempo real com acurácia e sensibilidade relativamente aceitáveis (SAGAN et al., 2020). Ferramentas como o sensoriamento remoto são muito utilizadas nesse tipo de monitoramento, podendo ser associadas a modelagem matemática e técnicas de aprendizado de máquina para melhoria da qualidade dos dados ambientais gerados (LU et al., 2020). As técnicas de sensoriamento remoto estão sendo utilizadas na área ambiental desde a década de 1970 em estudos de monitoramento de áreas desmatadas, do clima, dos oceanos, rios, lagos e reservatórios (NOVO, 2010).

Com os profundos avanços tecnológicos na tecnologia aeroespacial, nos sistemas computacionais e operacionais, tanto para aquisição quanto para processamento, armazenamento e disseminação das imagens da superfície terrestre, o sensoriamento remoto se tornou uma ferramenta frequentemente utilizada para realização de estudos relacionados ao monitoramento ambiental. A junção destes avanços e técnicas podem integrar sistemas de monitoramento inteligentes dos recursos hídricos, capazes de prover informações até mesmo em tempo real para os gestores destes recursos (MONTEIRO, 2020). Na água o Sensoriamento Remoto (SR) busca captar a interação

da luz com os componentes opticamente ativos (COAs), que seriam indicadores de qualidade da água específicos capazes de alterar o espectro de reflexão da água, pois interagem com a radiação eletromagnética incidente (SAGAN et al., 2020).

Dentre os avanços mencionados, a evolução dos sensores vem permitindo aferir com maior precisão a interação entre a luz e os alvos investigados em sensoriamento remoto. A resolução espectral se caracteriza pela largura e quantidade das bandas do espectro eletromagnético que o sensor consegue captar e quanto menor a largura e maior o número de bandas. Ou seja, quanto melhor a resolução espectral, mais específica será a capacidade de o sensor analisar as interações entre a radiação eletromagnética e os COAs, na massa de água (INPE, 2019). Sensores hiperespectrais, que captam centenas de bandas, gerando dados contínuos do espectro dos alvos, permitindo otimização das estimativas dos parâmetros de qualidade da água, principalmente em recursos hídricos com maior complexidade de poluentes presentes.

O Reservatório da Pampulha, cujo bacia hidrográfica se encontra nas cidades de Contagem e Belo Horizonte, cidade na qual se encontra todo o reservatório, apresenta como característica alto dinamismo em relação aos parâmetros de qualidade da água, sendo caracterizado como hipereutrófico (BEZERRA et al., 2022), influenciado pelo alto grau de urbanização e aumento da densidade populacional (SILVA et al., 2019).

O reservatório possui grande importância para a região, seu conjunto arquitetônico possui o título de Patrimônio Cultural da Humanidade pela Organização das Nações Unidas para a Ciência e a Cultura (UNESCO, sigla em inglês) e é um dos principais pontos turísticos de Minas Gerais. Até a década de 80 foi utilizado como fonte para o abastecimento público, no entanto a alta taxa de urbanização sem o adequado planejamento de uso e ocupação do solo e implementação de infraestrutura deterioraram a qualidade de suas águas (FURTADO; MONTE-MOR; COUTO, 2021).

A introdução de altas cargas de efluentes, principalmente o doméstico e o carreamento de sedimentos da bacia de drenagem são os fatores interferentes neste cenário. (BEZERRA et al., 2022). Os córregos Ressaca e Sarandi, que ocupam cerca de 60% de sua área e 90% da população, têm destaque como tributários que mais contribuem para a incidência de poluição (FURTADO; MONTE-MOR; COUTO, 2021).

Os principais problemas observados na Lagoa da Pampulha são o assoreamento, que provocou a perda de 1/3 de seu volume e a eutrofização, que alterou todo o equilíbrio ecológico do corpo d'água, provocando constantes florações de algas, entre elas as cianobactérias. Foram realizadas campanhas para a revitalização da lagoa, como disposição de uma cortina retentora de resíduos, dragagem de sedimentos, introdução de uma estação de tratamento de águas fluviais e tratamentos utilizando ferramentas de geoengenharia e biorremediação para diminuição da concentração de fósforo e nitrogênio, respectivamente. Todavia, sem resultados satisfatórios que mantiveram as condições hipereutróficas do reservatório, este cenário se deve ao contínuo aporte de poluição pelos tributários da Lagoa, principalmente os córregos Ressaca e Sarandi (BARÇANTE et al., 2020; BEZERRA et al., 2022; DE MELLO et al., 2018).

Diante do exposto, este estudo propõe analisar a utilização de sensores hiperespectrais para o monitoramento da qualidade da água, no que diz respeito às concentrações de chl-a, no reservatório da Lagoa da Pampulha. Em parceria com outros estudos, que buscam analisar a qualidade da água através de imagens de satélite e drones, o intuito é verificar a possibilidade de uso desta ferramenta para o monitoramento em tempo real das variações espaço-temporais do parâmetro no local estudado.

Estes estudos estão inseridos em maio a um projeto de Pesquisa e Desenvolvimento, intitulado "Monitoramento inteligente da qualidade da água em reservatórios hidrelétricos através do desenvolvimento de um algoritmo fotogramétrico". O projeto conta com recursos da Companhia Hidrelétrica de Minas Gerais (CEMIG), em parceria com o Departamento de Engenharia Sanitária e Ambiental da Universidade Federal de Minas Gerais (DESA-UFMG), através do grupo de pesquisas: Sistema Inteligente de Monitoramento das Águas (SIMOA).

Destarte, o estudo também almeja uma análise histórica das águas superficiais da Lagoa da Pampulha e a dinâmica do fitoplâncton e cianobactérias, em uma escala espacial e temporal, avaliando a eficiência das inúmeras intervenções de melhoria feitas neste reservatório na última década através do uso de remediadores ambientais e contenção de sedimento. Para isso, foi realizada uma ampla amostragem que abrangeu mensurações mensais, coletados com a sonda multiparâmetros, em 30

pontos no reservatório durante um ano, além de analisar historicamente a qualidade das águas mediante dados secundários, aos quais foram aplicadas ferramentas estatísticas para avaliar sua dinâmica temporal e espacial.

## **2 OBJETIVOS**

### **2.1 Objetivo Geral**

Este trabalho tem por objetivo utilizar técnicas de sensoriamento remoto para estimar a qualidade da água utilizando um sensor hiperespectral para determinação indireta de clorofila-a na Lagoa da Pampulha

### **2.2 Objetivos Específicos**

- Caracterizar a qualidade da água na Lagoa da Pampulha através do levantamento de dados históricos de monitoramento;
- Caracterizar a qualidade da água na Lagoa da Pampulha através de dados primários, coletados com a sonda multiparâmetros;
- Desenvolver a metodologia de calibração e utilização do sensor hiperespectral em laboratório, para avaliação da concentração de Clorofila-a em água;
- Aplicar a metodologia desenvolvida em campo, Lagoa da Pampulha, e desenvolver modelos empíricos para estimar a concentração de clorofila-a usando imagens hiperespectrais;
- Avaliar, a partir dos resultados obtidos, a aplicabilidade e sensibilidade do método para monitoramento de clorofila-a na Lagoa da Pampulha.

### **3 REVISÃO DA LITERATURA**

#### **3.1 Recursos hídricos em bacias hidrográficas urbanas**

Os recursos hídricos sempre apresentaram caráter estratégico para a humanidade, devido à sua essencialidade relacionada aos processos fisiológicos, comum a todos os organismos vivos e pelos diversos usos possíveis que rios, lagos, oceanos e, mais recentemente os reservatórios, podem proporcionar. Dentre os múltiplos usos da água destacam-se os usos na agricultura (irrigação, pecuária e aquicultura), responsável por cerca de 70% da demanda de água, abastecimento público, produção de energia elétrica, usos industriais, pesca, transporte e navegação, turismo, recreação, usos estéticos e no caso dos reservatórios, controle de cheias (BRASIL, 2005; TUNDISI; MATSUMURA-TUNDISI, 2013).

São muitos os impactos gerados pelas atividades humanas sobre os recursos hídricos em áreas urbanas e rurais que comprometem sua qualidade e inviabilizam os usos previstos. Em bacias hidrográficas o intenso crescimento populacional e a rápida urbanização provocam o aumento da impermeabilidade do solo, este processo geralmente ocorre sem a devida instalação de infraestruturas de drenagem (PAIVA et al., 2020). Com isto, há maior velocidade no escoamento provocando enchentes, aumento da erosão, que aumenta os níveis de sólidos carregados aos cursos de água, aumentando a poluição difusa no período chuvoso (SEIDL et al., 2020; SILVA et al., 2019).

A falta de infraestrutura também atinge o sistema de saneamento, problemas como a falta de estações de tratamento de esgoto, de cobertura da rede coletora em todos os domicílios e ligações irregulares nos sistemas pluviais de drenagem provocam o despejo de efluentes domésticos e industriais nos corpos d'água (SILVA et al., 2019). Com isto, há o aumento dos níveis de nutrientes, fósforo e nitrogênio, de produtos químicos e de metais pesados na água, o que altera os níveis de poluição e, conseqüentemente, todo o equilíbrio aquático (BARÇANTE et al., 2020).

Em particular, as alterações provocadas por ações antrópicas, associadas ao lançamento de esgoto sem tratamento, afetam a ciclagem de fósforo e nitrogênio nos ambientes aquáticos, e culminam no aparecimento do fenômeno denominado

eutrofização, que é o aumento do grau de trofia do meio caracterizado pelo aumento da concentração de nutrientes (SILVA et al., 2019). Em um ambiente eutrofizado, os altos níveis de nutrientes provocam aumento da produtividade, principalmente no fitoplâncton ou algas, resultando em acentuadas alterações qualitativas e quantitativas nas comunidades aquáticas e em alterações nos parâmetros físicos e químicos do meio (ESTEVES, 2011).

O aumento da densidade populacional do fitoplâncton é denominado floração ou *Bloom* de algas. Este é um grave problema ambiental, quando provocado por forças antrópicas, pois geralmente está relacionado ao aparecimento de cianobactérias, que possuem vantagens adaptativas em ambientes eutróficos e possuem como característica a potencial produção de toxinas que podem afetar a biota aquática e a saúde humana (SAGAN et al., 2020). Em ambientes eutróficos o ecossistema passa a produzir quantidade excessiva de matéria orgânica e sua decomposição diminui os níveis de oxigênio dissolvido no meio, podendo provocar a morte de espécimes superiores, como os peixes, por falta de oxigênio (ESTEVES, 2011).

Outros problemas provocados pela eutrofização são a perda de biodiversidade e a interrupção dos usos da água, sendo que o aumento do grau de trofia e o consequente aumento da produção primária ocorre em maior magnitude em lagos e reservatórios em consequência do maior tempo de detenção da água (SILVA et al., 2019).

### **3.2 Monitoramento de recursos hídricos**

Entender como a poluição afeta os aspectos espaço-temporais dos parâmetros de qualidade da água é de suma importância para se prover o gerenciamento adequado dos recursos hídricos, englobando a fiscalização, proteção e melhoria dos ecossistemas aquáticos (BEHMEL et al., 2016).

O monitoramento ambiental sistemático é instrumento fundamental para se obter dados contínuos dos aspectos quantitativos e qualitativos dos corpos d'água. O acompanhamento dos indicadores de qualidade da água ocorre, tradicionalmente, por meio da coleta de amostras de água em locais e intervalos de tempo pré-determinados e padronizados, com posterior análise laboratorial. A aquisição dos dados pode ocorrer ainda com a utilização de equipamentos de coletas de dados automáticos

posicionados nos pontos selecionados para amostragem e por meio de sondas multiparâmetros que medem a concentração dos parâmetros *in situ*. Em todos os casos os resultados obtidos são posteriormente comparados com padrões estabelecidos pelo poder público através dos órgãos responsáveis. A seleção dos parâmetros e pontos amostrais é realizada em função do tipo de ambiente aquático, dos objetivos do monitoramento, do uso das águas, da localização das atividades que podem influenciar na qualidade das águas e no tipo de poluição incidente (SÃO PAULO, 2012).

A análise dos dados obtidos no monitoramento ambiental torna possível a caracterização, avaliação temporal e tendências da qualidade das águas (JIANG et al., 2020), a identificação de áreas prioritárias para o controle da poluição, o controle e manutenção da qualidade da água para abastecimento, subsidiando as tomadas de decisões dos órgãos responsáveis por manter a disponibilidade e capacidade de atender aos usos múltiplos dos recursos hídricos (SÃO PAULO, 2012).

O monitoramento deve subsidiar o gerenciamento, de forma a atender os usos múltiplos previstos para cada corpo d'água. No Brasil a resolução 357 do Conselho Nacional do Meio Ambiente (CONAMA) de 7 de março de 2005, estabelece as diretrizes ambientais para o enquadramento dos recursos hídricos superficiais, fundamentado em princípios do desenvolvimento sustentável e visando o bem-estar humano, bem como o equilíbrio ecológico aquático e as condições e padrões de lançamento de efluentes. No âmbito do Estado de Minas Gerais a Deliberação Normativa Conjunta N.º 8, de 21 de novembro de 2022, do Conselho Estadual de Política Ambiental (COPAM-MG) e do Conselho Estadual de Recursos Hídricos do Estado de Minas Gerais (CERH-MG) estabelece as diretrizes ambientais relacionadas ao enquadramento e condições a serem mantidas nos corpos d'água.

### 3.2.1 Parâmetros de qualidade da água

Para o monitoramento da qualidade das águas são utilizados indicadores que servem de subsídio para avaliar as características dos cursos d'água. Uma série de parâmetros físicos, químicos e biológicos são estabelecidos pelos órgãos responsáveis pela gestão das águas, assim como seus limites. Leva-se em consideração a qualidade requerida para os diversos usos preponderantes dos

recursos hídricos (SÃO PAULO, 2012). A Tabela 2 apresenta alguns destes parâmetros que serão abordados no presente trabalho.

**Tabela 1 - Relação dos parâmetros avaliados no presente estudo e seus limites para as classes 1, 2 e 3, segundo a resolução CONAMA 357/2005 e portaria 05/2017 do Ministério da Saúde (microcistina e saxitoxina)**

Parâmetro	Limite Classe 1	Limite Classe 2	Limite Classe 3	Unidade de medida
Chl-a	10	30	60	µg/L
<p><b>Descrição:</b> O fitoplâncton, importante componente da biota aquática, é composto por algas fotoautotróficas que fazem parte da comunidade de organismos que não possuem movimentos próprios e vivem suspensos em ambientes aquáticos, denominada plâncton. Além do fitoplâncton, esta comunidade é composta pelo zooplâncton, que são pequenos animais, e pelo bacterioplâncton, que também apresentam organismos fotossintetizantes, as cianobactérias (ESTEVES, 2011). Desta forma, o termo algas é usado para discriminar, principalmente, um grupo diverso de organismos fotossintéticos eucariotos e procariotos (cianobactérias), que estão presentes em águas continentais e marinhas por todo o mundo (STENGEL; CONNAN; POPPER, 2011). As algas são espécies microscópicas e unicelulares, estão presentes na coluna d'água de forma isolada, em filamentos ou em colônias (ESTEVES, 2011).</p> <p><b>Importância:</b> Estes indivíduos são caracterizados por utilizarem a luz, do espectro do visível entre 360 nm e 700 nm, e gás carbônico como fonte de energia para sintetizar compostos orgânicos (PATEL, 2022) e são responsáveis por grande parte da produção primária no ambiente aquático, ou seja, são a base da cadeia trófica nestes ambientes HUISMAN, 2018). A fotossíntese é realizada por pigmentos, principalmente a clorofila-a e os carotenóides, sendo que outros pigmentos que compõem estes organismos como as clorofilas b e c, ficocianinas, xantofilas, ficoeritrinas e xicobilinas são consideradas acessórios (TUNDISI; MATSUMURA-TUNDISI, 2013).</p> <p><b>Fatores intervenientes:</b> São vários os fatores que interferem na fisiologia, crescimento e reprodução das comunidades algais. As adaptações adquiridas por estes organismos para permanecer na zona eufótica do lago ou reservatório, ou seja, zona com incidência da luz, tem grande relevância, em razão da necessidade de luz para a produção de energia (ESTEVES, 2011). Os índices de radiação incidente na coluna d'água, fundamental para o crescimento do fitoplâncton, a temperatura da água, que influencia diretamente nos processos metabólicos dos organismos, a concentração e distribuição espacial dos principais nutrientes como carbono, hidrogênio, oxigênio, enxofre, nitrogênio e fósforo, os dois últimos em especial, já que podem ser fator limitante na produção primária, têm caráter fundamental no desenvolvimento das comunidades (CAMPOS, 2010). Além destes, oscilações no clima, na concentração de matéria orgânica, no parasitismo e predação por herbívoros, na presença de sólidos em suspensão, que pode inibir a penetração da luz, na incidência de ventos, no tempo de residência da água e na estratificação térmica podem atuar como modificadores nos processos que determinam as condições sobrevivência do fitoplâncton (TUNDISI; MATSUMURA-TUNDISI, 2013).</p>				
Parâmetro	Limite Classe 1	Limite Classe 2	Limite Classe 3	Unidade de medida
Densidade de cianobactérias	2	5	10	mm <sup>3</sup> /L
<p><b>Descrição:</b> As cianobactérias são constituintes procariontes do fitoplâncton, possuem origem remota, há 3,5 bilhões de anos, compondo o grupo dos mais antigos a habitarem o planeta (PATEL, 2022). Conhecidas como algas verde-azuladas, podem na verdade apresentar variadas cores que dependem dos principais pigmentos associados a cada espécie. A ficocianina, por exemplo, concede às cianobactérias o tom ciano (azul-verde), porém, o tom verde é mais pronunciado devido à maior presença da chl-a. Além destes, outros acessórios como a ficoeritrina e carotenóides, pigmentos vermelhos e amarelo-laranja, respectivamente, dão às cianobactérias várias tonalidades de cores distintas (HUISMAN, 2018).</p> <p><b>Importância:</b> Em lagos e reservatórios com intensa eutrofização as estratégias de adaptação das cianobactérias as tornam dominantes, provocando densas florações destes organismos (FIGUEIREDO et al., 2016). Altas taxas de biomassa de cianobactérias é considerado um problema ambiental relevante, pois estes organismos têm o potencial de produzir toxinas, que podem ser prejudiciais tanto para a biota quanto para a os seres humanos (BARÇANTE, 2019). Desta forma, o</p>				

**Tabela 1 - Relação dos parâmetros avaliados no presente estudo e seus limites para as classes 1, 2 e 3, segundo a resolução CONAMA 357/2005 e portaria 05/2017 do Ministério da Saúde (microcistina e saxitoxina)**

monitoramento destes organismos é fundamental. A contagem da densidade de cianobactérias é utilizada, principalmente em ambientes lênticos, para mensurar sua ocorrência (IGAM, 2021).				
Parâmetro	Limite Classe 1	Limite Classe 2	Limite Classe 3	Unidade de medida
Fósforo total	0,02	0,03	0,05	mg/L
<p><b>Descrição:</b> O fósforo é uma das principais moléculas para os recursos hídricos, é considerado um nutriente essencial no desenvolvimento dos organismos. O fósforo total corresponde a todas as amostras de fósforo em uma amostra de água, particulado e dissolvido, subdivididos em orgânicos e inorgânicos (FIGUEIREDO et al., 2016).</p> <p><b>Importância:</b> A fração inorgânica dissolvida é a principal forma de fósforo assimilada por microalgas e bactérias. Sua presença no ecossistema aquático é considerada um fator limitante para a produtividade primária, pois as algas dependem de fontes externas deste nutriente para processos metabólicos como o armazenamento de energia, fotossíntese e reprodução. Porém, em excesso desencadeia processos de eutrofização, provocando proliferação de algas (ESTEVES, 2011).</p> <p><b>Fontes:</b> O fósforo não apresenta componentes gasosos na coluna d'água e sua presença em águas naturais depende da dissolução das rochas presentes na bacia de drenagem, neste caso o fósforo está presente em baixas concentrações e regula a concentração algal (ESTEVES, 2011). Entretanto, ações antrópicas como o uso inadequado de fertilizantes no meio rural e efluentes industriais e, principalmente, domésticos em áreas urbanizadas aumentam consideravelmente a concentração deste parâmetro nas águas (AGUIAR; NETO; RANGEL, 2011).</p>				
Parâmetro	Limite Classe 1	Limite Classe 2	Limite Classe 3	Unidade de medida
Turbidez	40	100	100	NTU
<p><b>Descrição:</b> Compreende o grau de atenuação sofrido por um feixe de luz ao atravessar a água, por processos de absorção e espalhamento em razão da presença de partículas de sólidos em suspensão orgânicos e inorgânicos (SAGAN et al., 2020).</p> <p><b>Importância:</b> Além de alterar a atenuação da luz, que pode modificar a dinâmica da produção primária do fitoplâncton, os sólidos suspensos quando se depositam no fundo do corpo hídrico formam bancos de lodo levando à formação de gases como o metano e gás carbônico, associados ao efeito estufa e o gás sulfídrico, que provoca odor (ESTEVES, 2011).</p> <p><b>Fontes:</b> A erosão das margens dos rios devido à retirada da mata ciliar, mau uso do solo em áreas urbanas e rurais, o aporte de efluentes domésticos e industriais e a mineração provocam aumento da turbidez, alterando todo o ecossistema aquático (GHOLIZADEH; MELESSE; REDDI, 2016). As algas, enquanto partículas em suspensão nas águas, também alteram a turbidez (SAGAN et al., 2020).</p>				
Parâmetro	Limite Classe 1	Limite Classe 2	Limite Classe 3	Unidade de medida
pH	6 a 9	6 a 9	6 a 9	-
<p><b>Descrição:</b> Mede a acidez e a basicidade de uma solução. Valores de pH acima de 7, que representa uma solução neutra, são considerados ácidos e abaixo deste valor são considerados básicos (LIBÂNIO, 2010).</p> <p><b>Importância:</b> O pH é um parâmetro de grande relevância por ter influência em diversos equilíbrios químicos que ocorrem nas águas naturais e afeta diretamente a fisiologia das espécies presentes nos sistemas aquáticos, como nos processos de permeabilidade celular (VON SPERLING, 2005).</p> <p><b>Fontes:</b> Origens antropogênicas que interferem na dinâmica do pH estão relacionadas com o despejo de efluentes domésticos e industriais e aumento de sedimentos provocados por processos erosivos na bacia hidrográfica. As comunidades biológicas do meio aquático também interferem na variação do pH, a assimilação do CO<sub>2</sub> durante a atividade fotossintética, por exemplo, pode elevar o pH. Por outro lado, a decomposição da matéria orgânica pelos organismos gera diminuição do pH, devido a produção de CO<sub>2</sub> (VON SPERLING, 2005).</p>				

**Tabela 1 - Relação dos parâmetros avaliados no presente estudo e seus limites para as classes 1, 2 e 3, segundo a resolução CONAMA 357/2005 e portaria 05/2017 do Ministério da Saúde (microcistina e saxitoxina)**

Parâmetro	Limite Classe 1	Limite Classe 2	Limite Classe 3	Unidade de medida
OD	> 6	> 5	> 4	mg/L
<p><b>Descrição:</b> Medida relativa do quantitativo de oxigênio que está dissolvido em um determinado corpo d'água.</p> <p><b>Importância:</b> Responsável pela na manutenção da biota por estar ligado à respira e oxidação da matéria orgânica, o OD aparece como um dos principais parâmetros de qualidade da água. Além disso, participa de inúmeras reações químicas e processos bioquímicos como na produção de energia e na reprodução (LIBÂNIO, 2010).</p> <p><b>Fontes:</b> Em um reservatório, as principais fontes de OD são a atmosfera, processo influenciado pela temperatura e pressão, e pela fotossíntese. A oxidação da matéria orgânica, introduzida por efluentes domésticos e pela eutrofização, a respiração dos organismos e a oxidação de íons metálicos representam as principais vias de perda de OD nos recursos hídricos (ESTEVES, 2011).</p>				
Parâmetro	Limite Classe 1	Limite Classe 2	Limite Classe 3	Unidade de medida
Nitrato	10	10	10	mg/L
Nitrito	1	1	1	mg/L
Nitrogênio Amoniacal (N.Am)	(1)	(1)	(1)	mg/L
<p><b>Descrição:</b> O nitrogênio é um elemento abundante na atmosfera terrestre (aproximadamente 78% do total), na água é encontrado nas formas reduzidas, nitrogênio orgânico e amoniacal (amônia e íon amônio), e oxidadas, nitrito e nitrato.</p> <p><b>Importância:</b> O íon amônio e o nitrato são as principais formas assimiladas pelos produtores primários, sendo de grande importância ecológica. O nitrogênio orgânico em sua forma dissolvida, como a uréia e aminoácidos, também podem ser utilizados como fonte deste nutriente pelos organismos. Por outro lado, a amônia e o nitrito adquirem relevância por serem compostos que apresentam toxicidade ao ambiente aquático (ESTEVES, 2011).</p> <p><b>Fontes:</b> As principais fontes antropogênicas de nitrogênio na água são os esgotos sanitários (orgânico e amoniacal). Além deste, as indústrias, os escoamentos pluviais em áreas urbanas e rurais e a atmosfera também são fontes de nitrogênio. Algumas cianobactérias são capazes de fixar o nitrogênio atmosférico, em virtude disto o nitrogênio pode não ser um fator limitante para a produção primária e as ações para conter o crescimento excessivo de algas devem priorizar o controle do fósforo. O nitrogênio pode ainda ser um indicador da distância da fonte de poluição em um corpo d'água, formas reduzidas são encontradas em locais próximos aos despejos poluidores e as formas oxidadas em locais distantes da fonte inicial de contaminação, sendo que o nitrato é o último composto desta cadeia (VON SPERLING, 2005).</p>				
Parâmetro	Limite Classe 1	Limite Classe 2	Limite Classe 3	Unidade de medida
Microcistina	1	1	1	µg/L
Saxitoxina	3	3	3	µg/L
<p><b>Descrição:</b> As toxinas são divididas em três grupos principais, neurotoxinas, hepatotoxinas e dermatotoxinas. As microcistinas são as cianotoxinas mais encontradas em lagos e reservatórios e tem características hepatotóxicas, a ocorrência de saxitoxina é a segunda mais comum e tem características e são consideradas neurotóxicas (BARÇANTE, 2019).</p> <p><b>Importância:</b> A microcistina tem grande relevância, pois podem provocar a morte em seres humanos. Não existem casos de incidentes mais graves envolvendo humanos relacionados a esta toxina. No entanto, a saxitoxina pode acumular na cadeia trófica e impactar substancialmente o ecossistema aquático (BATISTA; FIGUEIREDO; GIANI, 2018).</p> <p><b>Fontes:</b> São produzidas pelas cianobactérias, sendo que as microcistinas são produzidas, principalmente, pelo gênero <i>Microcystis</i> e as saxitoxinas são produzidas em sua maioria pelos gêneros <i>Anabaena</i>, <i>Cylindrospermopsis</i> e <i>Planktothri</i></p>				

**Tabela 1 - Relação dos parâmetros avaliados no presente estudo e seus limites para as classes 1, 2 e 3, segundo a resolução CONAMA 357/2005 e portaria 05/2017 do Ministério da Saúde (microcistina e saxitoxina)**

Parâmetro	Limite Classe 1	Limite Classe 2	Limite Classe 3	Unidade de medida
Temperatura da água	-	-	-	°C
<p><b>Descrição:</b> Consiste na variação térmica da coluna d'água. A radiação incidente nos comprimentos de onda maiores que 740 nm, principalmente no vermelho e infravermelho, são fortemente absorvidos pela água, influenciando a temperatura da água. As variações diárias e sazonais da incidência luminosa, a localização geográfica e condições climáticas são fatores que provocam alterações neste parâmetro (ESTEVEZ, 2011)</p> <p><b>Importância:</b> A temperatura da água tem significativa importância ambiental, pois interfere em inúmeros processos físicos, químicos e biológicos que ocorrem no corpo hídrico (TUNDISI; MATSUMURA-TUNDISI, 2013). Processos físicos como alterações na densidade, viscosidade e a estratificação térmica são influenciados pela temperatura da água e interferem diretamente na distribuição dos nutrientes e na movimentação da comunidade biológica na água (CAMPOS, 2010). Além disso, o teor de gases dissolvidos, como o oxigênio, a taxa reprodutiva e por consequência a produção primária e o tamanho da população fitoplanctônica sofrem influência da temperatura (PLEC, 2021). A estratificação térmica ocorre por diferenças verticais na densidade ao longo da coluna d'água que são provocados por gradientes de temperatura, onde as águas mais densas e frias tendem a permanecer na parte inferior do lago ou reservatório. Com isto, não há mistura entre os compartimentos do ambiente, formando-se extratos diferenciados física, química e biologicamente (TUNDISI; MATSUMURA-TUNDISI, 2013).</p>				

### 3.3 Princípios do sensoriamento remoto

O sensoriamento remoto consiste na obtenção de informações de determinado objeto, área de superfície terrestre ou fenômeno natural através de sensores que não estão em contato direto com este alvo, através da detecção e mensuração das mudanças impostas pelo alvo na luz incidente. Para tal, os sensores (passivos) captam a interação da luz ou radiação eletromagnética (REM) emitida pelo sol, principalmente, com os componentes da atmosfera ou da superfície (INPE, 2019). A REM, subdividida entre os campos elétrico e magnético, se propaga no vácuo, onde atinge a velocidade da luz, oscilando de forma ondulatória.

A Tabela 2 apresenta os intervalos do espectro eletromagnético e seus respectivos comprimentos de onda, destacam-se o visível, subdividido em azul (entre 400 nm e 500 nm), verde (entre 500 nm e 600 nm) e vermelho (entre 600 nm e 700 nm), e o infravermelho que são consideradas as regiões espectrais mais importantes para o sensoriamento remoto, já que são as faixas que, predominantemente, integram o espectro solar (NOVO, 2010).

**Tabela 2 – Divisão do espectro eletromagnético, com destaque para os comprimentos de ondas da luz visível**

<b>Intervalo espectral</b>	<b>Comprimento de onda</b>
Raios cósmicos	0,01 A <sup>o</sup>
Raios gama	0,01 – 0,1 A <sup>o</sup>
Raios X	0,1 – 10 A <sup>o</sup>
Ultravioleta	100 nm – 0,38 µm
<b>Visível (Azul)</b>	<b>0,38 – 0,50 µm</b>
<b>Visível (Verde)</b>	<b>0,50 – 0,60 µm</b>
<b>Visível (Vermelho)</b>	<b>0,60 – 0,76 µm</b>
Infravermelho próximo	0,76 – 1,2 µm
Infravermelho de ondas curtas	1,2 – 3,0 µm
Infravermelho médio	3,0 – 5,0 µm
Infravermelho termal	5,0 µm – 1 mm
Micro-ondas	1 mm – 1 m
Rádio	1 m – 10 km
Áudio	10 km – 100 km
Corrente alternada	> 100 km

Obs: A<sup>o</sup> = 10<sup>-10</sup> m; nm = 10<sup>-9</sup> m; µm = 10<sup>-6</sup> m  
 Fonte: Novo (2010)

A REM ao incidir sobre uma superfície pode ser refletida, transmitida ou absorvida pelos elementos que a compõem, estas interações podem ocorrer em dois níveis: macroscópico e microscópico. No nível microscópico observa-se as interações de troca de energia dos diferentes comprimentos de onda em função das características químicas do objeto. Quando ocorre, o material absorve o comprimento de onda específico no qual houve interação com a REM, sendo que o restante do espectro onde não existe esta interação será refletido ou disperso (INPE, 2019). A interação macroscópica se relaciona com o tamanho do comprimento de onda e dos objetos, caso o comprimento de onda seja menor que o material com o qual interage a REM incidente será barrada por este material e será refletida, caso contrário a interação entre a REM e o material será mínima (MENEZES e ALMEIDA, 2012).

Desta forma, cada material interage especificamente com a REM a depender de sua composição e tamanho, absorvendo comprimentos de onda espectrais específicos e refletindo o restante. Analisando os espectros refletidos captados pelos sensores depois da interação da REM com a superfície é possível distinguir a composição dos alvos através da assinatura espectral, que é definida pelos processos de interação (GHOLIZADEH; MELESSE; REDDI, 2016).

### 3.3.1 Sensoriamento remoto em ambientes aquáticos

Na água, a interação com a luz pode ser conduzida por dois processos: absorção e espalhamento. A reflexão observada após estes processos seria o sinal captado pelo sensor que é a resultante do fluxo de energia eletromagnética que incide sobre a superfície da água e o que é refletido de volta à atmosfera depois das interações de absorção e espalhamento (SAGAN et al., 2020; INPE, 2019).

Na absorção ocorre a interação microscópica, onde a energia contida no fóton é transferida para a molécula de água ou outras substâncias presentes na mesma. Este processo culmina no desaparecimento do fóton, alterando o espectro incidente em comprimentos de onda específicos. A própria água absorve cerca de 80% da radiação incidente no espectro do vermelho do visível (ESTEVES, 2011). Outra fonte de alteração do espectro são os organismos clorofilados, que possuem pigmentos eficientes na absorção da radiação nas faixas do azul e do vermelho para realização da fotossíntese (PATEL, 2022). Substâncias húmicas também absorvem na faixa do azul e sua presença pode interferir na fotossíntese, a absorção por detritos orgânicos e inorgânicos particulados ou dissolvidos é variada em função da sua composição heterogênea, em geral, apresenta semelhança com os compostos húmicos, pois normalmente estão adsorvidos nestes (SAGAN et al., 2020).

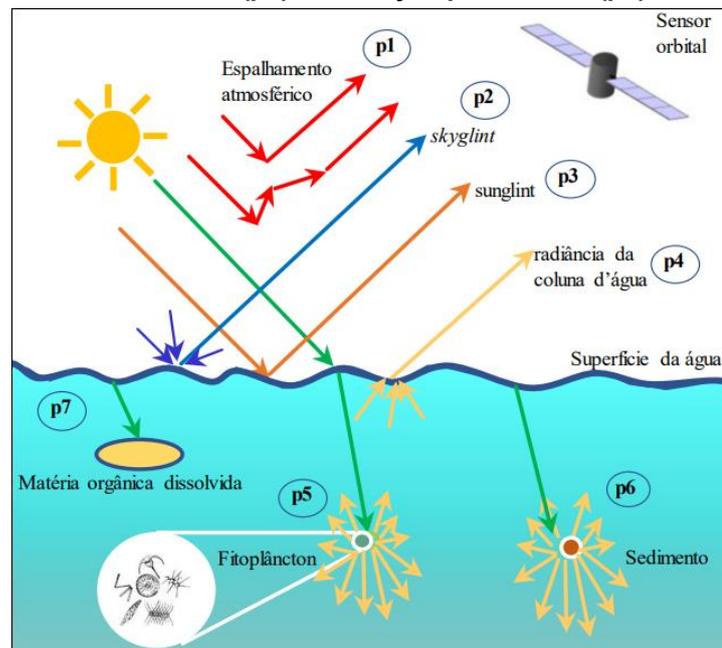
O espalhamento consiste na alteração da direção da REM ao interagir com as partículas orgânicas e inorgânicas em suspensão presentes na água. Ao ser espalhada no meio aquático, a luz apresenta mais chances de ser refletida de volta à superfície ou absorvida, o que reduz o alcance de profundidade da luz (INPE, 2019).

Desta forma, água na presença de baixas concentrações de partículas apresenta comportamento espectral diferente de quando apresenta concentrações elevadas de fragmentos vivos e não vivos em suspensão. Material particulado total como fitoplâncton, sedimentos minerais e orgânicos, matéria orgânica dissolvida colorida (CDOM) e a própria água são considerados constituintes opticamente ativos (COAs) e se caracterizam por interagir com a radiação solar (MENON; ADHIKARI, 2018). Os COAs são responsáveis pelos processos de absorção e de espalhamento da luz em um corpo d'água e podem coexistir em diferentes proporções, sendo que a variação

de suas concentrações e a interação entre elas determinam o comportamento espectral da massa de água (INPE, 2019).

A Figura 3.1 retrata um resumo dos possíveis caminhos que podem ser percorridos pela luz que incide em um corpo hídrico e sua reflexão até ser captada pelo sensor. O aumento do percentual de componentes na coluna de água pode aumentar os processos de absorção e espalhamento, interferindo no percentual de radiação incidente que é refletida de volta para a atmosfera. Na imagem observa-se ainda, os processos de *skylint* (irradiância atmosférica difusa refletida pela água) e *sunglint* (reflexão da luz solar refletida diretamente pela água). Estes fenômenos provocam distorções no espectro captado pelo sensor, inserindo incertezas nos resultados (GEGE; GROESTSCH, 2016).

**Figura 1 – Trajetória da luz incidente (solar) em um corpo hídrico até atingir o sensor. Demonstrando o espalhamento pela atmosfera (p1), skylint (p2), sunglint (p3), radiação da coluna de água (p4), absorção e espalhamento pelo fitoplâncton (p5), espalhamento pelo sedimento (p6) e absorção pelo CDOM (p7)**



Fonte: INPE (2019)

Parâmetros de qualidade da água opticamente ativos são indicadores específicos que podem ser detectados através das técnicas de sensoriamento remoto. Cita-se como exemplo a CDOM, turbidez, ficocianina e chl-a (PIZANI; MAILLARD e AMORIM, 2020).

Devido às diferenças observadas entre as águas oceânicas e continentais foram definidas duas classes para diferenciação das análises no sensoriamento remoto. Águas oceânicas caracterizam o Caso 1 e possuem um gradiente de transição suave influenciado principalmente pela concentração do fitoplâncton e CDOM. Águas interiores como rios, lagos, reservatórios, estuários e águas costeiras influenciadas por rios, possuem características ópticas complexas com concentrações de COA variadas e dinâmicas devido à sinergia entre os componentes físicos, químicos e biológicos do corpo hídrico. Estes, são influenciados pela interação com as respectivas bacias hidrográficas, esta categoria é denominada Caso 2 e a criação de modelos e algoritmos para este caso apresenta maior dificuldade e a replicação dos resultados em locais diferentes geralmente não é possível (HOOKER et al., 2021).

### **3.4 Sensoriamento remoto aplicado ao monitoramento de recursos hídricos**

O monitoramento ambiental tem possibilitado avanços na investigação da presença de diferentes contaminantes na água, ganhando inclusive celeridade com o desenvolvimento de tecnologia de sensores e equipamentos analíticos de maior acurácia e robustez. No entanto, o processo operacional de monitoramento dito tradicional, apesar de apresentar elevada precisão, é trabalhoso e demorado, além de demandar altos custos devido aos recursos humanos especializados, insumos necessários para as análises e pela grande extensão dos sistemas a serem monitorados (GHOLIZADEH; MELESSE; REDDI, 2016).

Estes fatores resultam na baixa frequência de campanhas de monitoramento e número reduzido de pontos amostrais, os quais são questões críticas para a eficiência do monitoramento, pois os corpos d'água apresentam grande heterogeneidade espacial e temporal dos seus parâmetros de qualidade (VILMIN et al., 2018).

Para minimizar os problemas relacionados às escalas temporais e espaciais do monitoramento, ferramentas como o sensoriamento remoto podem ser incorporadas com o intuito de auxiliá-lo (SAGAN et al., 2020). As técnicas de sensoriamento remoto têm sido amplamente utilizadas na área ambiental desde a década de 1970 impulsionadas pelo acesso a imagens de satélites que cobrem vastas áreas em todo o planeta, o que permitiu a realização de estudos em escalas regionais, ecossistêmicas, continentais e globais (AMANI et al., 2020). Com os profundos

avanços tecnológicos na tecnologia aeroespacial, nos sistemas computacionais e operacionais, tanto para aquisição quanto para processamento, armazenamento e disseminação das imagens da superfície terrestre, o sensoriamento remoto se tornou uma ferramenta frequentemente utilizada para realização de estudos relacionados às ciências ambientais (GHOLIZADEH; MELESSE; REDDI, 2016).

Várias plataformas foram desenvolvidas e os dados disponibilizados para tais estudos de forma gratuita por agências governamentais por todo o mundo, destaque para a série de satélites LANDSAT (*Land Remote Sensing Satellite*), desenvolvido pela Agência Espacial Americana. O programa está ativo desde a década de 70 e já lançou 9 satélites que disponibilizam imagens para observação de recursos naturais terrestres (GROSS et al., 2022).

Além do uso dos satélites, o aumento da sensibilidade dos sensores na captação de imagens com melhores resoluções, principalmente espacial e temporal, e a diversidade de plataformas utilizadas para a aquisição de dados como aviões, drones e mensurações in situ, permitiram que estudos ambientais em diversas áreas fossem realizados (GAI et al., 2020).

O uso de todo o avanço tecnológico observado relacionado ao sensoriamento remoto nas últimas décadas, principalmente a otimização dos sensores, integrado ao uso de técnicas da tecnologia da informação como processamento de dados e de imagens, o uso de computação e do aprendizado de máquina vêm recebendo o nome de sistemas de monitoramento inteligente da qualidade da água (MONTEIRO, 2020).

Segundo Dong et al. (2015), o monitoramento inteligente da qualidade da água interliga todo o progresso na capacidade de coleta de dados, comunicação, análise de dados e sistemas de alerta, sendo considerado como futuro no monitoramento da água. Sistemas propondo o uso da tecnologia da Internet das Coisas para coleta de dados e monitoramento em tempo real da qualidade da água (SARAVANAN et al., 2018; UDANOR et al., 2022; CAO et al., 2022) vem convergindo diferentes tecnologias para implementar novos métodos para monitorar os recursos hídricos.

Pesquisas para monitoramento da qualidade e quantidade da água, mapeamento da vegetação aquática e sedimentos, comportamento climático, processos de uso e

ocupação do solo em áreas urbanas e rurais, na agricultura, monitoramento de florestas e queimadas são exemplos do potencial de uso das combinações das técnicas de sensoriamento remoto em aplicações ambientais (MONTEIRO, 2020). Além disso, o monitoramento inteligente dos recursos hídricos pode auxiliar na gestão dos recursos hídricos, gerando informações que auxiliem na tomada de decisão relacionadas ao monitoramento e na resposta a eventos climáticos extremos ou desastres ambientais (DONG et al., 2018).

### **3.5 Métodos de aquisição de dados no sensoriamento remoto**

Os primeiros estudos relacionados à detecção de COAs através do sensoriamento remoto abordaram águas do Caso 1 com a utilização de sensores embarcados em satélites orbitais para estimativa do fitoplâncton, impulsionados pelo lançamento de várias missões de observação da Terra a partir da década de 70. Desde então missões como MODIS, LANDSAT e a série de satélites Sentinel fornecem imagens com informações que permitem, entre outros usos, caracterizar as concentrações de COAs nos oceanos e são utilizados na criação de diversos algoritmos e índices para mensuração destes parâmetros nos recursos hídricos (HE et al., 2020).

O monitoramento dos recursos hídricos com imagens de satélites tem a vantagem de abranger vastas áreas a um baixo custo operacional, uma vez que muitas missões disponibilizam os dados de forma gratuita. Contudo, podem apresentar limitações devido às suas menores resoluções, principalmente, a espacial e espectral, motivo pelo qual o sensoriamento por satélite apresenta resultados com menor precisão para águas interiores (WANG et al., 2017).

Outro fator que interfere na qualidade das imagens obtidas por satélites é a atenuação atmosférica. A REM interage com os gases e partículas atmosféricos tanto ao incidir sobre a superfície quanto ao retornar pela atmosfera até o sensor, interferindo no resultado do sinal detectado e gerando incertezas na qualidade dos resultados observados (INPE, 2019). Vários algoritmos foram desenvolvidos para a correção desta perturbação na captação do sinal por meio dos sensores, porém, mesmo atenuado, esta interferência provoca distorções nas análises de sensoriamento remoto (SAGAN et al., 2020).

Segundo Novo (2010) a altura do sensor em relação ao alvo define o nível de aquisição dos dados e, atualmente, ocorre a subdivisão nos níveis de laboratório/campo, de aeronave e orbital. Sendo que os níveis de obtenção interferem na proporção de interferência e intensidade do sinal registrado aumentando ou diminuindo a complexidade da caracterização espectral e por consequência na qualidade final dos dados e no esforço requerido para processamento e análise.

Recentemente, sensores instalados em plataformas não orbitais como aviões e veículos aéreos não tripulados (VANTs), popularmente conhecidos como drones, melhoraram as resoluções das imagens obtidas e atenuaram os problemas com a interferência atmosférica (GHOLIZADEH; MELESSE; REDDI, 2016). Na última década trabalhos utilizando tais plataformas vêm apresentando bons resultados para o monitoramento de COAs em águas do Caso 2, que em razão de sua alta complexidade necessitam de sensores com depurada qualidade em suas resoluções para detecção dos seus constituintes, principalmente a clorofila-a (HOOKER et al., 2021).

Outro método de aquisição de imagens consiste na utilização de sensores para medições de reflectância ao nível do solo ou de campo (in situ), esta metodologia minimiza a interferência atmosférica a níveis quase nulos, resultando em maior qualidade das imagens em relação aos dados obtidos por satélites, aviões e drones. Uma gama de dados de alta resolução são adquiridos e têm demonstrado capacidade de retorno de resultados substancialmente robustos para o monitoramento dos COAs. Estes dados são coletados, por exemplo, em embarcações ou em pontos fixos nas margens de rios a poucos metros da coluna d'água e podem ser utilizados para monitoramento de sistemas complexos com alta carga de poluição, como ocorre em águas interiores (KELLER et al., 2018). Neste nível de aquisição os fatores que podem interferir na qualidade da imagem capturada são a reflexão direta da luz na superfície que é dependente de aspectos como geometria de observação (azimute entre o sol e os ângulos do sensor), ondulações na água, condições atmosféricas (carga de nuvens) e horário da captura das imagens (ZIBORDI, 2016).

Ao nível de laboratório a análise ocorre com amostras dos ambientes ou material a ser estudado, ou ainda experimentos que simulam as características destes. Há o

controle ou exclusão dos fatores ambientais como reflectância difusa, interferência atmosférica, ondulações de superfície que podem interferir na reflexão espectral das amostras e, com isto, a aquisição de imagens com as assinaturas espectrais específicas para os alvos analisados (NOVO, 2010).

### **3.6 Sensoriamento remoto hiperespectral**

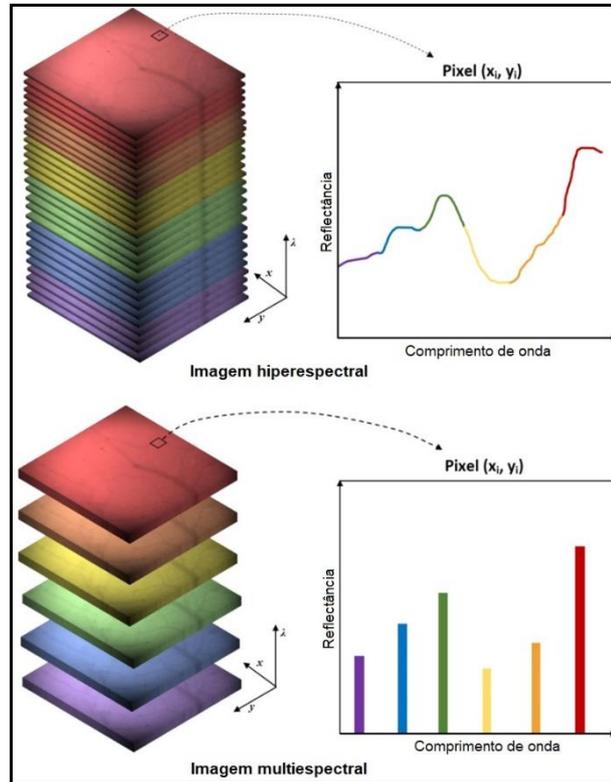
A resolução espectral se caracteriza pela largura e quantidade das bandas do espectro eletromagnético que o sensor consegue captar. Sensores com faixas espectrais largas não conseguem aferir as características individuais de cada interação entre a REM e a superfície, e quanto menos bandas do espectro forem captadas menor será a individualização dos processos de absorção e espelhamento específicos para cada elemento detectado pelo sensor. Desta forma, quanto menor a largura e maior o número de bandas, ou seja, quanto melhor a resolução espectral, mais específica será a capacidade de análise das interações entre a radiação eletromagnética e os COAs na massa de água (INPE, 2019).

Os sensores utilizados no sensoriamento remoto podem ser subdivididos em três classes segundo a sua resolução espectral: (i) câmeras que captam apenas três bandas na faixa do visível com espessas larguras no comprimento de onda do vermelho, verde e azul (RGB, sigla do inglês); (ii) multiespectrais que possuem um número maior de áreas imageadas e podem detectar faixas do espectro além do visível (GHOLIZADEH; MELESSE; REDDI, 2016); (iii) sensores hiperespectrais que detêm resolução superior aos demais, podendo chegar a centenas de bandas estreitas e adjacentes captadas (KELLER et al., 2018). Os avanços com os sensores hiperespectrais na última década vêm permitindo a ampliação do seu uso para o monitoramento dos parâmetros de qualidade da água, resultando em melhores estimativas, principalmente para as águas com maior grau de poluição e de sólidos suspensos (CUI et al., 2021; LU et al., 2020).

Estes resultados ocorrem em função de sua resolução espectral superior, que constrói um espectro contínuo para cada pixel através das inúmeras imagens captadas de cada banda que o sensor é capaz de captar. A partir destas imagens é formado um cubo hiperespectral, demonstrado na Figura 2, uma estrutura tridimensional ( $x$ ,  $y$  e  $\lambda$ ) com informações espaciais e espectrais do alvo. As dimensões  $x$  e  $y$  comportam a

informação espacial, enquanto a dimensão  $\lambda$  representa a informação espectral formada pela junção das imagens de todas as bandas adquiridas pelo sensor (GIANNONI; LANGE; TACHTSIDIS, 2018; SOARES, 2016).

**Figura 2 – Comparação entre um espectro contínuo retirado de uma imagem hiperespectral (a) e o espectro de uma imagem multiespectral (b)**



Fonte: Giannoni; Lange e Tachtsidis (2018)

Os sensores hiperespectrais utilizam matrizes de detectores, como Charge-Coupled Device (CCD) ou Complementary Metal-Oxide-Semiconductor (CMOS), para medir a intensidade da radiação em cada banda espectral. Cada pixel na matriz corresponde a uma banda espectral específica. Além disso, esses sensores incorporam filtros dispersivos, como prismas ou grades de difração, que permitem a separação da luz incidente em diferentes comprimentos de onda. Esses filtros direcionam cada banda espectral para conjuntos específicos de detectores. A unidade de aquisição de dados é responsável por converter os sinais detectados pelos detectores em valores digitais, por meio de processos como digitalização, amplificação e conversão analógico-digital dos dados espectrais (ALLEN et al., 2020).

Utilizando-se os dados hiperespectrais contínuos é possível estimar as concentrações dos COAs com maior precisão, através de comparações com dados das

concentrações dos parâmetros coletados simultaneamente às imagens (CUI et al., 2021). Na Figura 2, observa-se ainda a comparação entre uma imagem hiperespectral, com o espectro contínuo que representa a resultante da interação entre a luz e o alvo monitorado de forma completa e uma imagem multiespectral, como a obtidas em drones e satélites, que apresenta espectro discreto e representa apenas algumas faixas de reflexão do alvo.

A Tabela 3 apresenta estudos que utilizaram dados de sensores hiperespectrais cujo imagens foram coletadas *in situ* e em laboratório, formas de obtenção das imagens que são objetos do presente estudo. Nesta tabela observa-se que os sensores utilizados apresentam altas resoluções espectrais, com destaque para o ASD portátil da FieldSpec que possui tecnologia para captar 512 bandas com 1,3 nm de comprimento entre as faixas de 350 a 1050 nm.

Dentre estes estudo listados na tabela 3, nota-se que Lu et al. (2020) e Cui et al. (2021) apresentaram resultados robustos com  $R^2 > 0,9$  para a estimativa de chl-a. Cao et al. (2022) posicionaram nas margens da calha de um rio um sensor experimental com resolução espectral de 2,4 nm, que capta 354 bandas entre os comprimentos de onda de 280 e 1130 nm. Os autores analisaram uma ampla quantidade de parâmetros, onde todos os resultados para o coeficiente de determinação foram superiores a 0,7. Parâmetros como o fósforo e amônio, que não são opticamente ativos, apresentaram  $R^2 >$  que 0,9, sinalizando que a metodologia, que pode monitorar o corpo hídrico em tempo real, tem grande poder de aplicação na inferência destes parâmetros de qualidade da água.

**Tabela 3 – Estudos utilizando sensores hiperespectrais, detalhando o sensor, a faixa do espectro imageada, número de bandas, parâmetros utilizados, modo de aquisição e resultados obtidos**

Sensor	Faixa do espectro imageada (nm)	Nº de bandas/ Resolução espectral (nm)	Parâmetros	Desempenho	Modo de captura	Autor
ASD portátil FieldSpec	350 a 1050	512 / 1,3	Ficocianina	$R^2 = 0,9989$ e $RMSE = 5,9$ (mg/m <sup>3</sup> )	<i>In situ</i>	SUN et al. (2013)
ASD FieldSpec HandHeld – 2	350 a 1050	233 / 3,0	Sólidos suspensos	$R^2 > 0,99$ para as diferentes granulometrias de sedimentos utilizadas	Laboratório	QU et al. (2016)

**Tabela 3 – Estudos utilizando sensores hiperespectrais, detalhando o sensor, a faixa do espectro imageada, número de bandas, parâmetros utilizados, modo de aquisição e resultados obtidos**

Sensor	Faixa do espectro imageada (nm)	Nº de bandas/ Resolução espectral (nm)	Parâmetros	Desempenho	Modo de captura	Autor
ASD FieldSpec HandHeld – 2	350 a 1050	233 / 3,0	chl-a e sólidos suspensos	Chl-a, $R^2 = 0,98$ e RMSE = 6,15 ( $\mu\text{g/L}$ ) e SS, $R^2 = 0,97$ e RMSE = 1,91 (mg/L) 8 diferentes faixas de turbidez analisadas, todas com $R^2 > 0,8000$	<i>In situ</i>	Wang et al. (2017)
SVC HR 1024	350 a 2500	**	Turbidez		<i>In situ</i>	GARG et al. (2017)
FieldSpec 4 Hi-Res/ FieldSpec 3	350 a 2500	2151 / 3 nm até 700 nm e 8 entre 1400 e 2500	Chl-a	$R^2 = 0,92$	<i>In situ</i>	SANTOS et al. (2018)
ASD portátil FieldSpec	350 a 1050	512 / 1,3	chl-a	$R^2 = 0,99$ e MRE = 5%	Laboratório	ZHOU et al. (2019)
ASD FieldSpec HandHeld	235 a 1075	**/ 1,5 (visível)	chl-a Turbidez	Chl-a, $R^2 = 0,67$ e RMSE = 6,38 e turbidez $R^2 = 0,74$ e RMSE = 0,03 $R^2 = 0,75$ e RMSE = 11,54, águas hipertróficas $R^2 = 0,9485$ , MAPE = 54,43% e RMSE = 27,11	<i>In situ</i>	SOUZA (2019)
ASD portátil FieldSpec	350 a 1050	512 / 1,3	chl-a	$R^2 = 0,9024$ e RMSE = 1,1312 ( $\mu\text{g/L}$ )	<i>In situ</i>	CHENG et al. (2019)
ASD portátil FieldSpec	350 a 1050	512 / 1,3	chl-a	MAPE = 54,43% e RMSE = 27,11	<i>In situ</i>	LU et al. (2020)
ASD portátil FieldSpec	350 a 1050	512 / 1,3	chl-a	RMSE = 1,1312 ( $\mu\text{g/L}$ )	<i>In situ</i>	CUI et al. (2021)
ATP200P Optosky Photonics	280 a 1130	354 / 2,4	*	Todos os $r^2$ ultrapassaram 0,7, destaque para pH e $\text{NH}_4\text{-N}$ , $R^2 = 0,97$ e 0,90, respectivamente	<i>In situ</i>	CAO et al. (2022)
ASD FieldSpec 4	350 a 2500	**	chl-a	$R^2 = 0,78$ e MAPE = 35,21%	<i>In situ</i>	LI et al. (2022)

\* =  $\text{NH}_4\text{-N}$ , nitrogênio total, Fósforo total, sólidos suspensos totais, chl-a, ficocianina,  $\text{NO}_3\text{-N}$  e pH;

\*\* = Não mencionado.

Fonte: Autor (2023)

No Brasil, Santos et al. (2018) utilizaram 2 sensores hiperespectrais capazes de obter imagens de até 3 nm de resolução espectral. O rio Subaé na Bahia foi monitorado e

foram encontrados  $R^2$  de 0,92, demonstrando que a tecnologia de SR hiperespectral tem capacidade de detectar com boa acurácia nos resultados das concentrações de chl-a no corpo hídrico. Souza (2019) monitorou chl-a e turbidez com imagens coletadas *in situ* na Lagoa da Pampulha, foi utilizado um sensor com resolução espectral de 1,5 nm na zona do visível. Concluiu-se que modelos utilizando dados subdivididos em razão dos compartimentos do reservatório apresentaram melhores resultados para os modelos de chl-a.

### 3.6.1 Propriedades ópticas da clorofila-a

A chl-a é um pigmento presente na maioria dos organismos fitoplanctônicos, pode ser um indicador da presença de algas no ecossistema e tem um papel importante na absorção de energia solar em uma coluna de água (ESTEVES, 2011). Além da chl-a, as algas podem apresentar outros pigmentos, que são caracterizados como acessórios ou estão presentes em uma gama restrita de espécies, sendo que a chl-a está presente na maioria dos organismos do fitoplâncton. Este pigmento apresenta eficiente absorção da luz na região do visível, especificamente, absorve a REM em comprimentos de ondas específicos nas regiões azul e vermelha e tem picos de reflectância na região do verde da luz visível (PATEL, 2022; SAGAN et al., 2020). A Tabela 4 apresenta alguns estudos em que são relatadas as faixas de comprimento de onda em que ocorre maior absorção e os picos de reflexão para a chl-a.

**Tabela 4 - Estudos relatando as faixas espectrais de absorção e reflectância da clorofila-a em ambientes aquáticos e as formas de aquisição das imagens**

Pigmento	Faixa de absorção (nm)	Faixa de reflexão (nm)	Método de aquisição	Autor (ano)
chl-a	412, 443, 678	488 e 531	Satélite MODIS, multiespectral	HE et al. (2020)
	Entre 400 e 500 e em 670	Entre 550 e 580	In situ, Hiperespectral	CUI et al. (2021)
	Azul e verde do visível e 670	662 e 726	In situ, Hiperespectral	KELLER et al. (2018)
	420 e 670	Entre 520 e 560	Drone, Hiperespectral	GAI et al. (2020)

Fonte: Autor (2023)

O comprimento de onda de 670 nm representa uma faixa importante na detecção remota da chl-a, vários estudos encontraram correlação desta faixa do espectro com o parâmetro (SAGAN, 2020). Os diferentes resultados nos comprimentos de onda encontradas na Tabela 4 podem ser explicadas por diferenças na concentração,

tamanho, forma e estágio das células e colônias dos organismos clorofilados (ESTEVES, 2011), bem como na forma de aquisição dos dados de sensoriamento remoto, interação com outros COAs e métodos de processamento e análise de dados utilizados (GHOLIZADEH; MELESSE; REDDI, 2016; SAGAN et al., 2020). As condições da luz incidente, como sua intensidade, são fatores determinantes para o crescimento das algas, que apresentam comportamento distinto de interação com a radiação em função de suas diferentes fases de crescimento (PATEL, 2022). Além disso, a presença de diferentes concentrações de outros parâmetros, como sólidos suspensos e CDOM, por exemplo, pode interferir nas análises de reflectância. Isto ocorre porque a matéria orgânica pode interagir fortemente com a luz na região do azul e do verde e os sólidos suspensos têm banda de reflexão ativa na faixa do vermelho (HE et al., 2020; GAI et al., 2020).

### **3.7 Processamento de dados**

Muitas abordagens podem ser adotadas para estimar as concentrações dos COAs na água. Os métodos empíricos, semiempíricos e analíticos podem gerar modelos que buscam encontrar correlações entre as reflectâncias de sensoriamento remoto inseridas nas imagens captadas pelos sensores e a concentração dos parâmetros que estão presentes no corpo d'água monitorado (DÖRNHÖFER; OPPELT, 2016).

Nos modelos analíticos ou bio-ópticos os dados são manipulados com base na transferência radiativa, que leva em conta os processos físicos da propagação da radiação eletromagnética e os fatores que a influenciam como processos de absorção, emissão e espelhamento, para simular a relação entre os COAs e o espectro de reflexão captado da água (SAGAN et al., 2020). Esta abordagem é bastante complexa, pois a construção do modelo exige conhecimento das características ópticas inerentes da água, movimentação da superfície e da variação dos parâmetros de qualidade da água, se tornando uma abordagem desafiadora para águas continentais (YANG et al., 2022).

Abordagens empíricas são métodos que utilizam regressões lineares ou não lineares que relacionam a reflectância em comprimentos de ondas específicos com valores de concentrações dos parâmetros a serem analisados, coletados de forma simultânea (KHAN et al., 2021). Possuem a vantagem de fácil operação e de ter a capacidade de

refinar a seleção dos canais espectrais mais precisos para o conjunto de dados em análise, o que pode aumentar a precisão do algoritmo em relação à predição dos parâmetros de qualidade da água. No entanto, é afetado pela complexidade óptica de águas interiores, necessitam de grande quantidade de dados para geração do modelo e não podem ser generalizados para outros corpos d'água (YANG et al., 2022).

Os modelos semiempíricos são uma combinação das abordagens citadas anteriormente. Este método utiliza informações dos processos físicos e espectrais da interação da radiação eletromagnética com as mudanças e variações dos componentes presentes na água para desenvolver os algoritmos, que posteriormente são correlacionados com os constituintes medidos (GHOLIZADEH; MELESSE; REDDI, 2016).

Os modelos empíricos e semiempíricos utilizam análises de regressão bi e multivariada, sendo que para o desenvolvimento de algoritmos em águas do caso 2, devido à sua complexidade óptica, as análises multivariadas são recomendadas (SAGAN et al., 2020).

Modelos que utilizam inteligência artificial vêm sendo utilizado de forma ampla para inversão de parâmetros de qualidade da água recentemente (LU et al., 2021), este processo utiliza o aprendizado de máquina para melhor explorar a relação complexa entre as variáveis, dependentes e independentes, analisadas (CAO et al., 2022). O melhor desempenho destes modelos se deve à capacidade de captar relações lineares e não lineares, características em águas interiores (YANG et al., 2022). Modelos utilizando redes neurais (LU et al., 2021), máquinas de vetores de suporte (KELLER et al., 2018) e XGBooster, um tipo de árvore de decisão (CUI et al., 2021) obtiveram êxito ao estimar parâmetros de qualidade da água através de dados de sensores hiperespectrais. A tabela 5 descreve os modelos que serão empregados no presente trabalho.

**Tabela 5 - Relação e descrição dos modelos e suas descrições utilizados na etapa de estimativa da concentração de clorofila-a com dados hiperespectrais de campo**

Modelo	Requisitos	Descrição
Regressão Linear (RL)	Necessita de relações lineares entre as variáveis	Estima uma variável dependente a partir de valores de uma variável independente, estabelecendo relação entre elas. Pode ser simples (uma variável independente) ou múltipla (mais de uma variável independente) (SPERLING; VERBYLA; OLIVEIRA, 2020).
<i>Support Vector Regressor</i> (SVR)	Recomendado para dados não lineares e complexos, porém não é adequado para grande quantidade de dados	Derivada da <i>Support Vector Machine</i> , este tipo de aprendizado de máquina molda regressões não lineares criando um hiperplano capaz de minimizar os erros inseridos no processo de modelagem (QUAN et al., 2022).
<i>Random Forest Regressor</i> (RFR)	Funciona bem para variáveis contínuas e categóricas, além de ser eficiente para dados não lineares, com valores ausentes e com <i>outliers</i> . Requer maior esforço computacional e longo período de treinamento	Também é um tipo de aprendizado de máquina que utiliza no processo de predição a criação de várias árvores de decisões, de maneira aleatória e simultânea, através dos dados originais. A partir daí, são selecionados atributos com maior ganho de informação (“notas”), gerando o resultado da predição (LU et al., 2021).
<i>XGBooster</i> (XGB)	Eficiente para dados não lineares e com dados ausentes. No entanto, não funciona bem em dados esparsos e apresenta alta sensibilidade a <i>outliers</i> .	O <i>Extreme Gradient Boosting</i> é considerado uma evolução do RFR. Utiliza árvores de decisões, porém em sequência, onde as árvores são analisadas uma a uma. Os erros da árvore anterior são minimizados a cada processo de decisão (CUI et al., 2021).

Fonte: Autor (2023)

### 3.8 Estudo de caso: Lagoa da Pampulha

O Reservatório da Lagoa da Pampulha, popularmente conhecido como Lagoa da Pampulha, situado na região metropolitana de Belo Horizonte/MG, cuja bacia hidrográfica de 98 km<sup>2</sup> abrange as cidades de Contagem (55%) e Belo Horizonte (45%), cidade na qual se encontra todo o reservatório. A Lagoa da Pampulha, que faz parte da bacia do ribeirão do Onça e deságua no rio das Velhas, apresenta características de alta dinamicidade em suas concentrações de sedimentos e compostos dissolvidos devido à interação com a bacia hidrográfica e processos limnológicos internos (BEZERRA et al., 2022), caracterizando-se como águas do Caso 2.

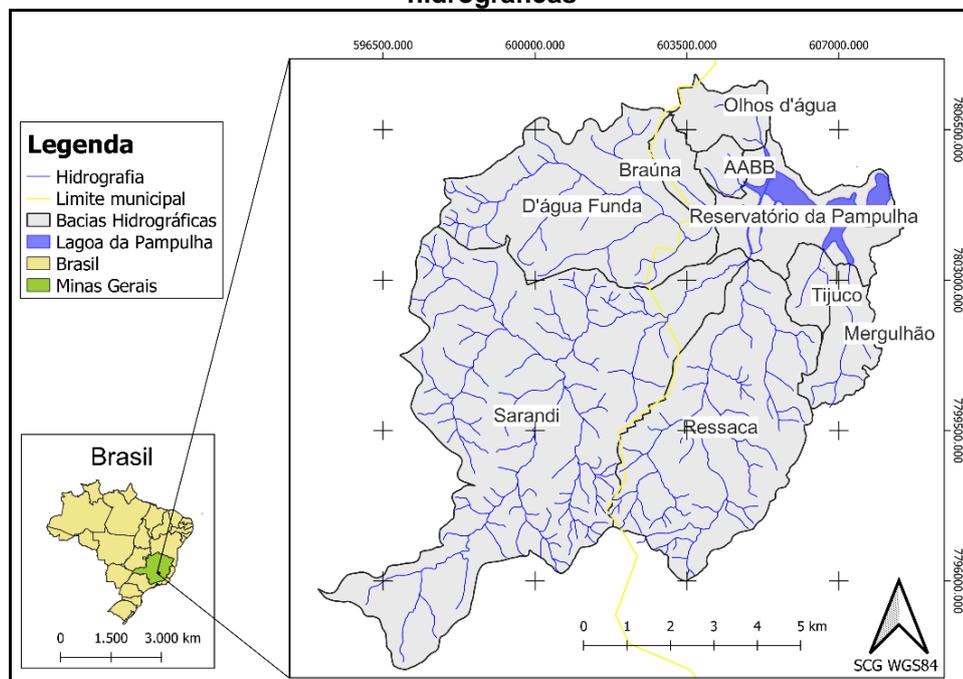
O alto grau de urbanização, promovido sem a devida infraestrutura, o aumento da densidade populacional e a industrialização, que ocorrem desde a década de 70, são fatores que vêm provocando mudanças nas características de uso e ocupação do solo da bacia e por consequência, provocam efeitos adversos na composição das águas

carreadas em toda a rede de drenagem, alterando as características físico, químicas e biológica das águas do reservatório (BARÇANTE et al., 2020).

A Lagoa possui área superficial de aproximadamente 2 km<sup>2</sup>, tempo de retenção hidráulica de 89 dias (média anual) e profundidade variando entre 0,7 m e 14 m, caracterizando-se como um reservatório pequeno e raso (PLEC et al., 2021). A Figura 3 expõe a localização geográfica da bacia hidrográfica subdividida em 8 sub-bacias, com destaque para as sub-bacias dos córregos Ressaca e Sarandi que ocupam cerca de 60% da área e 90% da população, que tem mais 500.000 habitantes (FURTADO; MONTE-MOR; COUTO, 2021).

O clima da região, segundo a Classificação Climática de *Köppen*, é do tipo Cwa (Tropical de altitude com inverno seco e verão chuvoso), as chuvas ocorrem entre os meses de outubro e março (média anual de 1.524,5 mm), período caracterizado também por temperaturas mais elevadas. O período seco ocorre entre abril e setembro, período este que apresenta menores temperaturas e baixa pluviosidade (CBHSF, 2016).

**Figura 3 - Localização geográfica da Lagoa da Pampulha e subdivisão de suas sub-bacias hidrográficas**



Fonte: Autor (2023)

O reservatório, construído na década de 1930, com o propósito de minimizar impactos de enchentes, auxiliar no abastecimento de água e como área de lazer, possui grande importância para a região. Seu conjunto arquitetônico (projetado por Oscar Niemeyer) obteve o título de Patrimônio Cultural da Humanidade pela UNESCO em 2016 e é um dos principais pontos turísticos de Minas Gerais (LOPES et al., 2020). O uso de suas águas como fonte para o abastecimento público prosseguiu até a década de 70. No entanto, a alta taxa de urbanização sem o adequado planejamento de uso e ocupação do solo e implementação de infraestrutura deterioraram a qualidade de suas águas (FURTADO; MONTE-MOR; COUTO, 2021).

A introdução de altas cargas de efluentes industriais e, principalmente, domésticos, oriundos de fontes clandestinas lançados diretamente nos cursos d'água e o carreamento de sedimentos da bacia de drenagem, são fatores que interferem neste cenário (BARÇANTE et al., 2020; BEZERRA et al., 2022). Segundo a Companhia de Saneamento de Minas Gerais (COPASA, 2022), existem cerca de 10 mil imóveis que lançam esgotos domésticos de forma irregular na bacia hidrográfica da Lagoa da Pampulha.

Os principais problemas observados no reservatório são o assoreamento, que provocou a perda de cerca de 1/3 do volume (MELLO et al., 2018), lançamento de grandes volumes de lixo (SOUZA, 2019) e a eutrofização, que altera todo o equilíbrio ecológico do corpo d'água e provocam problemas como a floração de algas no reservatório (RIETZLER et al., 2018).

Estas florações ou *Blooms* de algas na Lagoa da Pampulha são caracterizados principalmente pela dominância de cianobactérias, que apresentam alta adaptabilidade às condições de hipertrofia encontradas no reservatório, podendo representar a quase totalidade das espécies presentes no meio (CAMPOS, 2010; BATISTA; FIGUEIREDO; GIANI, 2018; SILVA et al., 2019; BARÇANTE et al., 2020). Estas espécies são potenciais produtoras de cianotoxinas e sua presença é fator de risco para o meio. Barçante (2019) analisou a presença de toxinas na Lagoa da Pampulha (microcistina e saxitoxina), como resultado foi possível constatar a presença destes compostos em todo período estudado (abril/16 a maio/18).

Nas últimas décadas foram realizadas diversas ações para revitalização da lagoa, em 2003 iniciou-se a operação de uma Estação de Tratamento de Águas Fluviais (ETAF) dos córregos Ressaca e Sarandi (SOUZA, 2019). Segundo Coutinho (2007), a ETAF, que opera com o processo de flotação a ar dissolvido e tem capacidade de tratamento de aproximadamente 750 l/s, apresentou eficiência satisfatória na remoção média de turbidez, sólidos em suspensão, DBO e DQO.

Entretanto, por se tratar de um tratamento primário, apresenta ineficiência em tratar parâmetros associados à eutrofização (BEZERRA et al., 2022; SILVA et al., 2019). Além disso, a estação, que não é capaz de tratar todo o volume incidente dos córregos Ressaca e Sarandi, principalmente no período de chuvas, quando a vazão é aumentada. Neste período a ETAF opera tratando parcialmente a vazão que chega ao reservatório ou fechando suas comportas no período em que ocorrem as precipitações, com o intuito de não prejudicar todo o processo de tratamento com vazões acima das previstas no projeto (COUTINHO, 2007)

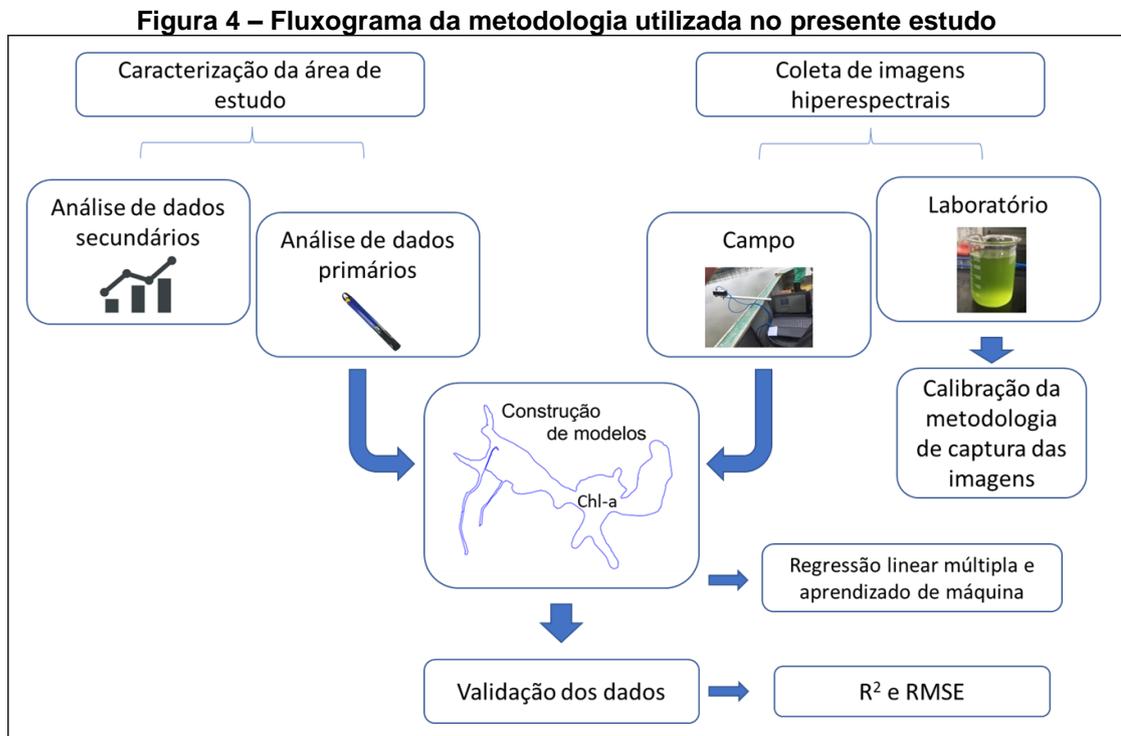
Após receber o título de Patrimônio Cultural da Humanidade, o reservatório passou a receber tratamentos para melhoria dos parâmetros de qualidade da água, a fim de alcançar os padrões estabelecidos pela resolução CONAMA 375/2005 de classe 3 (BEZERRA et al., 2022). Uma das ações consiste no emprego da ferramenta de geoengenharia para sequestro do fósforo da coluna d'água, denominado Phoslock®. Este produto consiste em uma argila, bentonita modificada com lantânio, que quando inserida no meio aquático é capaz de adsorver os íons fosfato, formando o mineral rabdofano ( $\text{LaPO}_4$ ). Este mineral possui granulometria suficiente para se sedimentar, ficando retido permanentemente no fundo do reservatório (BARÇANTE et al., 2019). Outro tratamento utilizado foi a biorremediação microbiana Enzilimp®, que acelera o ciclo do nitrogênio e a degradação da matéria orgânica através de bactérias anaeróbicas, aeróbicas e facultativas (BEZERRA et al., 2022). Além disso, outros programas como a disposição de uma cortina retentora de resíduos (SOUZA, 2019), dragagem de sedimentos (MELLO et al., 2018), ações para conter a erosão nas bacias dos principais tributários e melhorias no sistema de tratamento e captação dos esgotos foram implementados pelos órgãos responsáveis (BARÇANTE, 2019)

Todavia, as tentativas de revitalização da Lagoa da Pampulha apesar de suavizar as concentrações dos principais poluentes que provocam a eutrofização, como matéria orgânica, nitrogênio e fósforo, não foram satisfatórias a ponto de interferir nas condições de hipertrofia do reservatório. Este cenário se deve, principalmente, ao contínuo aporte de altas cargas poluidoras que desaguam na Lagoa oriundas dos principais tributários da Lagoa, os córregos Ressaca e Sarandi, que representam 70% do fluxo de água que chega ao reservatório (BARÇANTE et al., 2020; BEZERRA et al., 2022).

Recentemente, uma ação conjunta entre a COPASA e as prefeituras de Belo Horizonte e Contagem anunciou um plano para acabar com os lançamentos de esgoto na Lagoa da Pampulha. Estão previstas obras de pequeno e grande porte, plano de comunicação, mobilização social e monitoramento da qualidade das águas com o intuito de zerar os lançamentos de esgoto nos cursos d'água da bacia da Lagoa (COPASA, 2022).

## 4 METODOLOGIA

Para o desenvolvimento desse trabalho as seguintes etapas metodológicas, dispostas na Figura 4, foram seguidas:



Fonte: Autor (2023)

Primeiramente foram adquiridas junto ao IGAM os dados de monitoramento para análise dos dados secundários, primeiro objetivo específico. A coleta dos dados primários foi realizada entre os meses de março de 2022 e fevereiro de 2023, logo após os dados foram compilados e analisados para cumprimento do segundo objetivo específico. Ao mesmo tempo foram realizadas as etapas de coleta das imagens em laboratório e construção do modelo, utilizando o método de mínimos quadrados parciais, para o terceiro objetivo específico. Os dados coletados pela sonda foram então utilizados para construção dos modelos no quarto objetivo específico, através da relação com as imagens hiperespectrais que foram coletadas simultaneamente.

### 4.1 Materiais

#### 4.1.1 Sensor hiperespectral

O sensor utilizado para coleta das imagens se trata do OCI™-F *Ultra-Compact Hyperspectral Imager* da BaySpec (EUA). O sensor utiliza a tecnologia CMOS, que é

composto por uma matriz de pixels sensíveis à luz, onde cada pixel contém um fotodetector e circuitos eletrônicos associados. Eles são capazes de medir a intensidade da radiação

em diferentes bandas espectrais e converter essas informações em sinais elétricos. Comparado aos detectores CCD, os sensores CMOS são conhecidos por seu menor consumo de energia, menor tamanho, maior velocidade de leitura e melhor capacidade de integração com outros circuitos eletrônicos (RIBA, 2022).

As imagens geradas fornecem resolução de 1024 x 1280 pixels, cada coluna desta imagem corresponde a uma banda espectral, ou seja, são captadas 1280 bandas espectrais diferentes. Entretanto, para reduzir a relação sinal-ruído, a resolução espectral é reduzida, para os comprimentos de onda entre 413 nm e 955 nm (visível e infravermelho próximo), com resolução espectral de 5,1 nm e 109 bandas. Esta redução de dimensionalidade é calculada estimando em cada banda a média de uma região de aproximadamente 11 nm da imagem original (KHEZRABAD; ZOEJ; SAFDARINEZHAD, 2022).

O sensor hiperespectral foi adquirido com recursos cedidos pela Companhia Hidrelétrica de Minas Gerais (CEMIG), em meio a um projeto de pesquisa e desenvolvimento em parceria com o Departamento de Engenharia Sanitária e Ambiental da Universidade Federal de Minas Gerais (DESA-UFMG), intitulado “Monitoramento inteligente da qualidade da água em reservatórios hidrelétricos através do desenvolvimento de um algoritmo fotogramétrico”.

Para obter a calibração radiométrica, são registradas as imagens de padrões de referência branco com reflectância de 95% (painel fornecido pela *BaySpec*) e escuro (capturado com o protetor da lente impedindo a passagem da luz), com o objetivo de corrigir variações na distribuição espectral das condições da luz do ambiente e converter os valores dos números digitais (ND) captados pelo sensor em reflectância (SHAIKH et al., 2021).

A fabricante do sensor disponibiliza o software *CubeCreator* (*BaySpec* Inc. San Jose, EUA), que processa as imagens brutas dos alvos monitorados. Este software utiliza as imagens de referência citadas anteriormente para calibrar a imagem bruta com os

valores de reflectância corrigidos. Após a calibração, o CubeCreator transforma as imagens em cubos hiperespectrais, como citado na seção 3.6, combinando automaticamente todas as faixas do comprimento de onda em um único arquivo. Desta forma, cada pixel presente na imagem calibrada obtém informações espectrais de todas as bandas capturadas pelo sensor (ROSSITER et al. 2020).

#### 4.1.2 Sonda multiparâmetros

A coleta dos dados de concentração de chl-a foi mensurada por meio da sonda multiparâmetros EXO2 (YSI). Além deste parâmetro, a EXO2 realiza medições de temperatura da água, pH, OD, turbidez e ficocianina que serão utilizados para caracterizar a qualidade da água na Lagoa da Pampulha (segundo objetivo específico).

### **4.2 Caracterização da área de estudo com dados históricos de monitoramento**

Para efetivação do primeiro objetivo específico foram coletados dados secundários de parâmetros limnológicos da Lagoa da Pampulha que influenciam e são influenciados na dinâmica da chl-a (seção 3.2.1). Desta forma, os parâmetros chl-a, Densidade de Cianobactérias (DC), fósforo total (PT), turbidez, pH, oxigênio dissolvido (OD), nitrato, nitrito, nitrogênio amoniacal (N. Am.), temperatura da água (TA), microcistina e saxitoxina, coletados entre os anos de 2013 e 2021 foram utilizadas neste estudo.

Os dados foram disponibilizados pelo Instituto Mineiro de Gestão das Águas (IGAM) e fazem parte do programa de monitoramento da qualidade das águas no estado de Minas Gerais que conta com amostragem trimestral (março, junho, setembro e dezembro) e nas análises laboratoriais são empregadas as metodologias em conformidade com o *Standard methods for the examination of water and wastewater* (IGAM, 2022).

A caracterização da qualidade da água na Lagoa da Pampulha através do levantamento de dados secundário foi realizada através das seguintes etapas: i) análise preliminar dos dados; ii) a estatística descritiva (Mediana, mínimo, máximo, desvio padrão, percentis e *outliers*) e análise temporal; e iii) aplicação de testes não-paramétricos.

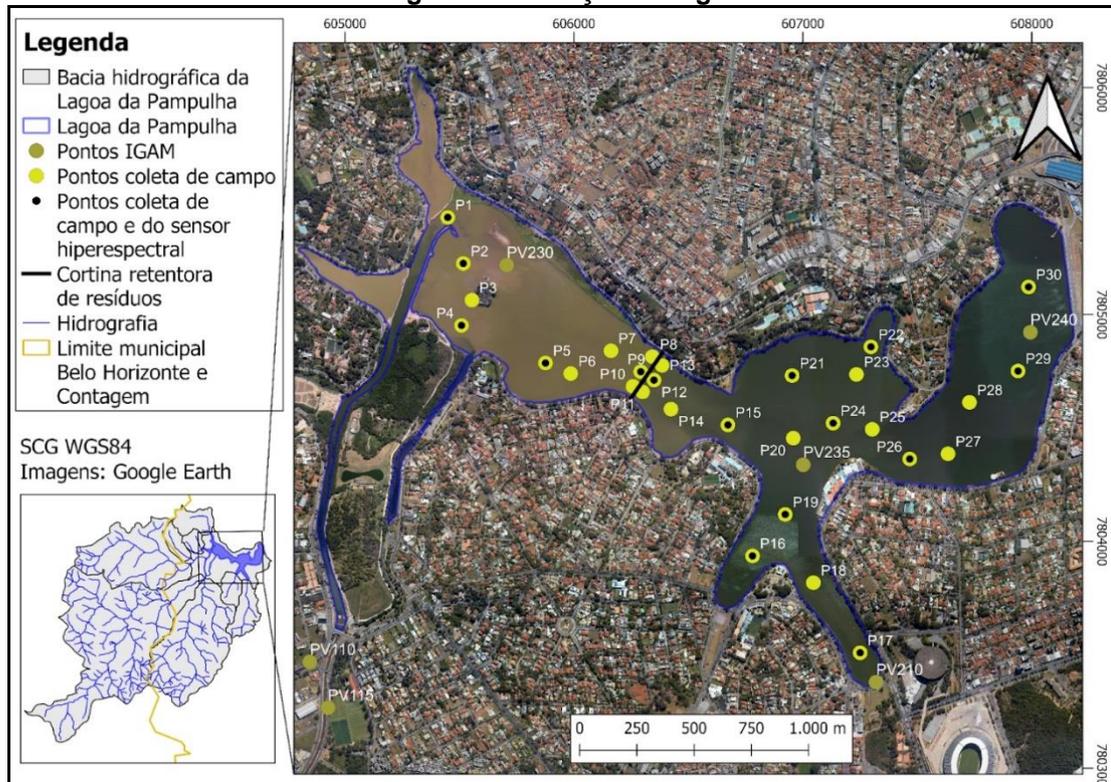
A primeira etapa foi subdividida em seleção e organização do banco de dados e a análise preliminar. Os dados foram organizados em planilhas do *Excel* para verificação da frequência de coleta dos parâmetros analisados. A princípio seriam analisados 10 anos de monitoramento. Entretanto, ao analisar o conjunto de dados, percebeu-se que o repositório para o ano de 2012 apresentava apenas uma amostragem, referente ao mês de dezembro, e optou-se por retirar este ano do escopo do estudo. Nesta etapa também foram analisados os percentuais de dados faltantes. Onde parâmetros com mais de 60% dos dados faltantes devem ser excluídos da análise.

Posteriormente, foi realizada a análise de estatística descritiva para cada parâmetro monitorado, foram calculados ainda o percentual de *outliers*, cada outlier foi analisado individualmente com a finalidade de verificar se deveriam ser mantidos ou excluídos do conjunto de dados antes das análises.

O teste de hipótese de Mann-Whitney foi aplicado para verificar se existem diferenças significativas entre os períodos seco e de chuva, a um nível de significância igual a 5%. Já o teste de Kruskal-Wallis foi conduzido para verificar se há diferenças significativas entre as medianas dos pontos monitorados, complementado pelo teste *post-hoc* de Dunn para apontar em quais pontos se encontram tais diferenças, se existirem. Todas as análises estatísticas foram realizadas utilizando o software RStudio (Versão 4.1.3).

Na Figura 5 podemos observar a disposição dos pontos amostrais de coleta do IGAM no reservatório. O ponto PV230 fica próximo à Ilha dos Amores e ao local de afluência dos principais tributários da Pampulha, os Córregos Ressaca e Sarandi. O ponto PV235 se localiza nas imediações da Igreja de São Francisco de Assis, região intermediária e de transição do reservatório e o ponto PV240 se encontra próximo à barragem do reservatório. Além destes, foram avaliados 3 pontos representando os tributários da Lagoa, os córregos Ressaca (PV110), Sarandi (PV115) e Mergulhão (PV210), que foram selecionados em função de sua influência no reservatório. As coordenadas dos pontos monitorados pelo IGAM e os pontos monitorados pela sonda no presente estudo estão dispostos no Apêndice A.

**Figura 5 - Pontos amostrais monitorados pelo IGAM, pontos utilizados para o monitoramento com a sonda multiparâmetros e os pontos selecionados para o monitoramento com o sensor hiperespectral. Além destes, observa-se a localização da cortina retentora de resíduos, na região de transição da lagoa**



Fonte: Autor (2023)

### 4.3 Caracterização da qualidade da água na Lagoa da Pampulha através de medições obtidas pela sonda multiparâmetros

Nesta etapa, foi utilizada a sonda multiparâmetros EXO2 (YSI) para medição de chl-a, ficocianina, pH, OD e turbidez, realizada pela equipe do grupo de estudos SIMOA, na profundidade de subsuperfície (metade da zona fótica). Além dos dados obtidos pela sonda, foi medida a profundidade do disco de Secchi.

Os pontos amostrais (Figura 5) foram definidos buscando representar a compartimentação do reservatório: i) com condições próximas às características fluviais à montante da Lagoa, devido à influência de seus principais tributários, córregos Ressaca e Sarandi (região onde se encontra o ponto PV230); ii) com características lacustres na área próxima à barragem, pontos 26, 27, 28, 29 e 30 (nas proximidades do ponto PV240); e iii) na região de transição entre estes dois ambientes (entre a região a jusante da cortina retentora de resíduos e o ponto 25), local onde o IGAM monitora o ponto PV235. Estes diferentes pontos representam compartimentos que apresentam características físicas, químicas e biológicas distintas, devido,

principalmente, a fatores como tempo de residência e eventos climáticos (CAMPOS, 2010). Além destes, foram avaliados 3 pontos representando os tributários da Lagoa, os córregos Ressaca (PV110), Sarandi (PV115) e Mergulhão (PV210), que foram selecionados em função de sua influência no reservatório.

Ao todo, 30 pontos foram monitorados, sendo 10 destes na região próxima à montante do reservatório, antes da cortina retentora e 20 após a cortina, a jusante da Lagoa. As coletas foram realizadas entre os meses de março de 2022 e fevereiro de 2023, com o intuito de se observar as variações dos parâmetros em ambos os períodos climáticos encontrados na região da Lagoa da Pampulha. Para realização das coletas dos dados a prefeitura de Belo Horizonte permitiu o acesso à Lagoa da Pampulha e disponibilizou a embarcação e os barqueiros para condução pelo trajeto de forma gratuita.

Os dados foram analisados por meio de gráficos *Box-plot*, teste de Mann-Whitney para análise de sazonalidade, para análise comparativa dos parâmetros monitorados antes e depois da cortina retentora de resíduos, e para comparação entre os dados históricos do IGAM analisados e os dados coletados pela sonda.

A fim de simular as variações dos parâmetros de qualidade da água observados em todo o reservatório foi determinada por Krigagem Ordinária, utilizando o software QGIS na versão 3.22.9. Este método utiliza a interpolação para estimar as concentrações de uma variável em locais não mensurados a partir de dados de locais mensurados. Introduzida no campo da geoestatística, foi desenvolvida por Georges Matheron em 1963, aperfeiçoando estudos do engenheiro de mineração sul-africano Daniel Krige, que começou a considerar a distância entre amostras de concentração de ouro para estudos de variâncias (GUALBERTO, 2020).

A técnica de interpolação utilizada é o *Best Linear Unbiased Estimator* (BLUE), que estima a autocorrelação espacial de pares de pontos no espaço, considerando uma variância mínima do erro, por meio de um variograma empírico (ROSTAMI et al., 2020). Desta forma, o processo de Krigagem pode ser dividido em duas etapas: i) determinação da estrutura de covariância espacial dos pontos amostrados pelo variograma; e ii) os pesos derivados do variograma são utilizados para interpolar valores para os locais não amostrados no espaço em análise (GUALBERTO, 2020).

As equações 1 e 2 representa o cálculo da autocorrelação espacial em que  $Z^*(X_0)$  é o valor estimado em  $X_0$  e  $\sigma_k^2(X_0)$  é a variância do erro em cada ponto  $X_0$ .

$$Z^*(X_0) = \sum_{i=1}^n \phi_i Z^*(X_i) + [1 - \sum_{i=1}^n \phi_i] m(X) \quad (1)$$

$$\sigma_k^2(X_0) = \mu + \sum_{i=1}^n \phi_i \gamma(X_0 - X_i) \quad (2)$$

$Z^*(X_i)$  representa uma variável aleatória na localização  $X_i$ ;  $\phi_i$  os pesos;  $\mu$  é o coeficiente de Lagrange;  $\gamma(X_0 - X_i)$  é o valor do semivariograma com base na diferença entre  $X_0$  e  $X_i$ ; e  $m(X)$  é o valor médio, que por suposição básica da Krigagem Ordinária é estacionário e desconhecido (ROSTAMI et al., 2020). Foram utilizados nas análises os valores medianos de cada ponto, considerando a subdivisão entre o período chuvoso e o de estiagem.

Análise de Cluster (AC) foi utilizada com o intuito de avaliar a semelhança da qualidade da água entre os 30 pontos amostrais, por meio de agrupamentos de pontos semelhantes, e posteriormente compará-los com as regiões onde os pontos monitorados pelo IGAM se localizam. Dessa forma, foi possível analisar a eficácia do monitoramento realizado pelo órgão. Foram utilizados os dados de todos os parâmetros avaliados no monitoramento com a sonda multiparâmetros, além das profundidades do disco de Secchi, em todas as campanhas realizadas (mar/2022 a fev/2023).

A metodologia proposta por Costa (2021) foi utilizada, empregando o método de agrupamento *ward*, distância euclidiana como medida de similaridade e a validação dos resultados obtidos foi realizada com o cálculo do coeficiente de correlação cofenética (CCC). Este método de validação mensura o grau de ajuste dos agrupamentos calculados com a matriz de dissimilaridade (distância euclidiana), CCC acima de 0,7 representa baixas distorções e bom grau de ajuste nos agrupamentos obtidos.

A AC é uma ferramenta multivariada que classifica determinado parâmetro ou conjunto de parâmetros em grupos através de similaridades presentes nos dados. O método utilizado no presente estudo (*Ward*), aglomera hierarquicamente os dados em grupos que apresentam maior similaridade entre si, utilizando a distância euclidiana

como medida de dissimilaridade (KUMAR et al., 2018). Anteriormente à aplicação da AC foi aplicada a padronização dos dados para converter os dados em escalas similares, evitando viés nos resultados pela presença de variáveis com ordens de grandeza discrepantes. Foi utilizado como forma de padronização a escala Z ( $Z = (X - \mu) / \sigma$ ), onde há subtração pela média e divisão pelo desvio padrão, para cada variável.

O método *gap statistic* determina a quantidade ideal de agrupamentos realizando simulações alterando o número de *clusters* e comparando com conjuntos de dados de referência gerados a partir dos dados de entrada. Efetua-se cálculo da dispersão dentro de cada comparação, e a identificação do número ótimo de grupos ocorre quando há grandes distorções entre os agrupamentos (TIBSHIRANI; WALTHER; HASTIE, 2001). Assim como nos dados monitorados pelo IGAM, as análises estatísticas e a AC foram conduzidas por meio do *software* RStudio (Versão 4.1.3).

#### 4.4 Desenvolvimento da metodologia de calibração e utilização do sensor hiperespectral em laboratório

Foram realizados experimentos em ambiente com luminosidade controlada para obtenção das curvas espectrais para estimar as concentrações de chl-a em amostras mono específicas da alga *Raphidocelis subcapitata* (Figura 6), com o sensor OCI™-F *Ultra-Compact Hyperspectral Imager* da *BaySpec* (EUA). O experimento foi conduzido seguindo a metodologia descrita por Qu et al (2016) e objetivou a validação do método de coleta das imagens.

**Figura 6 - Algas *Raphidocelis subcapitata* utilizadas no experimento de laboratório para determinação de clorofila-a**



Fonte: Autor (2022)

As algas utilizadas no experimento foram cultivadas no laboratório de ecotoxicologia do Departamento de Engenharia Sanitária e Ambiental da Universidade Federal de Minas Gerais. Os cultivos foram conduzidos conforme as regras da Associação Brasileira de Normas Técnicas (ABNT) em sua norma ABNT NBR 12.648/2018, que especifica o método para avaliação de toxicidade crônica de amostras líquidas para as microalgas verdes unicelulares.

As imagens hiperespectrais foram extraídas de diferentes concentrações de chl-a obtidas a partir da diluição do cultivo de *Raphidocelis subcapitata* (Tabela 6). Para tal, 2 litros de cada concentração de chl-a foram dispostos em um recipiente retangular de 37 cm x 27 cm, onde a coluna d'água apresentou 3,5 cm (Figura 7a).

**Tabela 6 - Concentrações de chl-a utilizadas na captura das imagens hiperespectrais e posterior construção do modelo de regressão**

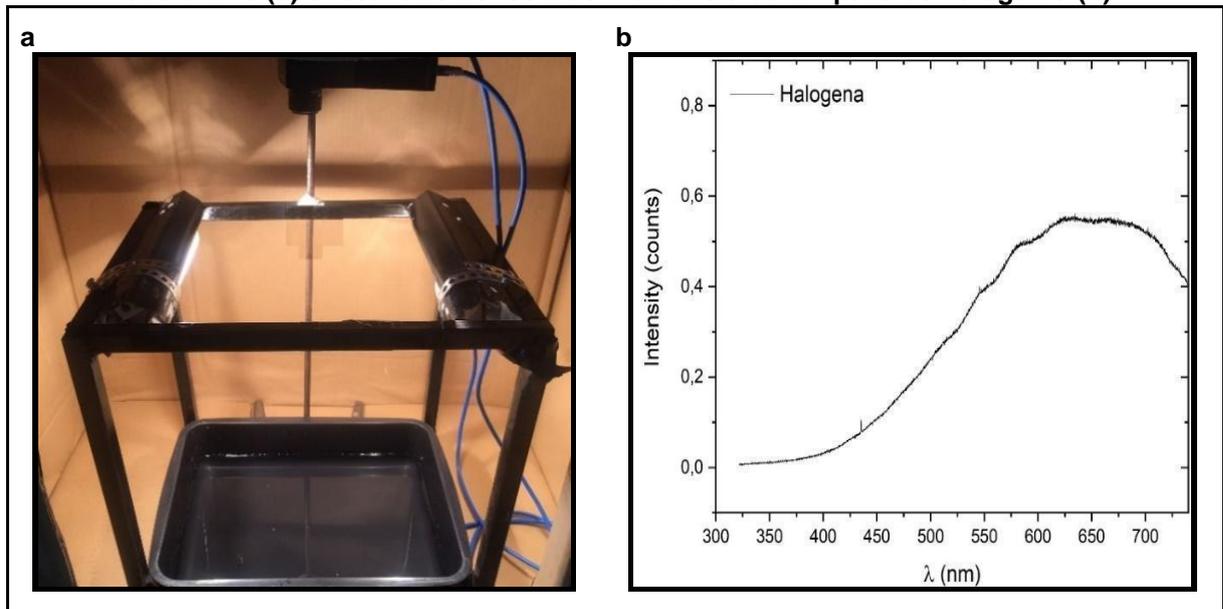
Nº da amostra	Concentração de chl-a	Nº da amostra	Concentração de chl-a
1	0	7	47
2	10	8	58
3	15	9	122
4	19	10	239
5	28	11	407
6	35	12	686

Fonte: Autor (2023)

Qu et al. (2016) indicam o uso de um ambiente escuro e com as amostras dispostas em um recipiente preto para evitar a influência da iluminação do ambiente e a reflectância de fundo, respectivamente. Desta forma, busca-se obter imagens em condições controladas de luz e sem nenhuma interferência ambiental externa ao experimento, para posteriormente extrair a reflectância da superfície da água com um mínimo de interferências externas.

Para a captura das imagens hiperespectrais, o sensor foi posicionado na vertical, focando o centro do recipiente que continha 2 litros do líquido amostrado, a uma distância de 40 cm. Obteve-se, desta forma, campo de visão de 25° e campo de visão instantâneo de 8,8 cm. O ambiente foi iluminado com duas lâmpadas halógenas de 500 W para simular a luz solar. As lâmpadas foram dispostas de forma frontal e a 20 cm da amostra em um ângulo de incidência de 45° do alvo.

**Figura 7 - Aparato experimental para as coletas das curvas espectrais realizadas em laboratório (a) e resultado do teste de irradiância da lâmpada de halogênio (b)**



Fonte: Autor (2022)

Com o intuito de comparar a emissão espectral da lâmpada com a luz solar, realizou-se um teste para aferir a irradiância emitida pelas lâmpadas utilizadas neste experimento. O teste foi conduzido no Laboratório de Nano-Espectroscopia da Universidade Federal de Minas Gerais (LabNS – UFMG) com o equipamento *Spectrometer model Andor ShamRock SR-330i*, que escaneia as emissões da lâmpada entre os comprimentos de onda de 325 nm a 750 nm. Os resultados (Figura 7b) permitiram constatar que a lâmpada tem forte incidência luminosa nas regiões espectrais entre 550 nm e 750 nm, com um platô de pico entre 600 nm e 700 nm, regiões de forte interesse para a interação entre os pigmentos fotossintetizantes e a luz (PATEL, 2022).

#### 4.4.1 Desenvolvimento do algoritmo e validação dos dados para o experimento em laboratório

Após a aquisição dos resultados referentes à quantificação da clorofila-a e das curvas espectrais de reflectância obtidos depois do processamento das imagens hiperespectrais em laboratório, foram realizadas análises de regressão para estimativa das concentrações do parâmetro. A abordagem adotada foi totalmente empírica e a análise dos dados realizada através do *software RStudio* versão 4.2.2.

Como a resolução espectral do sensor é alta, tornou-se necessário a seleção de variáveis preditoras, ou seja, bandas espectrais que mais contribuíram para a

construção do modelo, com o propósito de diminuir efeitos de multicolinearidade provocada pela similaridade das bandas imageadas pelo sensor. A abordagem de seleção de variáveis *forward inclusion stepwise* foi utilizada para esta triagem. O processo de escolha por meio desta abordagem ocorre de forma iterativa, onde as variáveis são inseridas na modelagem, uma a uma, e caso ocorra um aumento na acurácia a variável é mantida no subconjunto de variáveis selecionadas, caso contrário, é eliminada. A modelagem termina quando todas as bandas forem adicionadas ao processo ou quando a acurácia do modelo atingir 100% (FONTES, 2020).

A regressão de Mínimos Quadrados Parciais (Partial Least Squares Regression - PLS) foi utilizada para a construção do modelo. PLS é uma técnica de regressão recomendada quando os dados possuem um elevado número de variáveis independentes e altamente correlacionadas, decompondo os componentes tanto na matriz de reflectância espectral, quanto nos dados de chl-a. Estas matrizes são relacionadas por regressão, estabelecendo a correlação entre as variáveis, verificada por validação cruzada (ZHAO et al., 2021).

A validação dos resultados foi realizada através de duas métricas de avaliação. O coeficiente de determinação ( $R^2$ ), que expressa a proporção do desvio na variável dependente que é previsível a partir da variável independente, utilizando um nível de significância de 5% ( $p < 0,05$ ), quanto maior o  $R^2$  melhor o modelo se ajusta aos dados observados. A outra métrica utilizada será a Raiz do Erro Quadrado Médio (RMSE, sigla em inglês) que é comumente usada para expressar a acurácia dos resultados, expressando seus valores na mesma dimensão da variável dependente (SAGAN et al., 2020).

$$R^2 = \frac{\sum_{i=1}^n (\hat{y}_i - \bar{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y}_i)^2} \quad (3)$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (\hat{y}_i - \bar{y}_i)^2} \quad (4)$$

Onde  $n$  é o número de amostras,  $y_i$  e  $\hat{y}_i$  representam os valores observados e os previstos das amostras de verificação e  $\bar{y}_i$  é o valor médio dos valores observados (HOU et al., 2021).

## 4.5 Estimativa da concentração de clorofila-a através do sensor hiperespectral em campo e do desenvolvimento de modelos empíricos

### 4.5.1 Aquisição das imagens

As campanhas mensais para aquisição das imagens foram realizadas entre os meses de agosto de 2022 e fevereiro de 2023, no total de sete amostragens. A Figura 5 mostra a distribuição dos 16 pontos nos quais foram tomadas as imagens com o sensor. Os pontos foram definidos através de análise de cluster entre os 30 pontos monitorados na seção 4.2 (Figura 5), onde pontos com características similares foram subtraídos, mantendo-se a representatividade de todo o reservatório com os pontos. Na campanha do mês de fevereiro imagens de todos os 30 pontos foram capturadas, com o intuito de recuperar amostras perdidas nas coletas dos meses anteriores.

Para a obtenção dos dados hiperespectrais ao nível de campo, com o intuito de minimizar efeitos de reflexão especular da luz solar, o sensor foi posicionado verticalmente ( $90^\circ$ ) em relação à coluna d'água e a aproximadamente um metro do alvo, conforme CUI et al. (2021) e WANG et al. (2017). As coletas foram realizadas no período da manhã, medida que segundo (KAY; HEDLEY; LAVENDER, 2009; CHONG et al., 2021), minimizam os efeitos da luz solar, pois o ângulo de incidência do sol é diferente do campo de visão do sensor, que não capta a reflexão solar. Além disso, a coleta dos dados amostrais de chl-a e as imagens hiperespectrais ocorreram simultaneamente, para que a fluidez da água provocada pelo vento não sejam uma fonte de incerteza durante a construção dos modelos (SU, 2017). Na Figura 8 é possível observar o método pelo qual o sensor foi posicionado para captura das imagens que foi montado especificamente para alocar o sensor.

**Figura 8 – Disposição do sensor na embarcação para coleta das imagens hiperespectrais**



Fonte: Autor (2022)

O manual do sensor elenca a forma correta de coleta das imagens, que são: (i) conectar o sensor ao notebook via cabos USB 3.0; (ii) Captura das imagens do painel de referência de branco e escura, após ajustar o ganho e tempo de exposição; (iii) coleta das imagens hiperespectrais brutas, após ajuste de ganho e tempo de exposição.

#### 4.5.2 Processamento das imagens coletadas em campo

Na etapa de coleta das imagens em campo, o processamento e criação dos cubos hiperespectrais não foi realizado pelo software disponibilizado pela fabricante do sensor. O erro foi atribuído à velocidade (gerando imagens sem sobreposição) ou estaticidade (gerando imagens totalmente sobrepostas) da embarcação no momento da captura das imagens. No total, cerca de 70% das imagens capturadas apresentaram tal problema.

Para não perder as informações adquiridas, foi necessário criar rotinas computacionais de calibração e extração dos dados das imagens de forma independente para contornar os problemas encontrados no software nativo. Os processos foram realizados nos *softwares* ImageJ e *RStudio* versão 4.2.2.

A calibração das imagens para transformação dos valores de ND em reflectância foi realizada utilizando a Equação 3, que é um processo de normalização amplamente usado para calibração de imagens de sensoriamento remoto (NOVIYANTO; ABDULLA, 2019; SHAIKH et al., 2021).

$$R(x, y) = \frac{R_{ND}(x, y) - D(x, y)}{W(x, y) - D(x, y)} \times 100 \quad (5)$$

Onde  $R(x, y)$  é o dado da imagem calibrada na forma de reflectância relativa (0% - 100%) na posição  $(x, y)$ ,  $R_{ND}(x, y)$  é a imagem original e  $D(x, y)$  e  $W(x, y)$  são as referências de escuro e padrão de reflectância branco, respectivamente.

Nesta etapa não foram criados os cubos hiperespectrais como no *CubeCreator*, sendo que o fluxo de trabalho implementado termina com a aquisição dos dados das curvas hiperespectrais de cada imagem para inserção na etapa de modelagem dos dados.

#### 4.5.3 Desenvolvimento dos algoritmos empíricos e validação dos dados para as imagens coletadas *in situ*.

Como entrada para os modelos foi utilizada a média da reflectância de cada uma das 109 bandas da imagem em cada ponto monitorado durante as campanhas realizadas entre agosto de 2022 e fevereiro de 2023, alinhados com seus respectivos dados de chl-a coletados simultaneamente com a sonda. Nesta etapa também foi utilizada a metodologia de seleção de bandas, *forward inclusion stepwise*, detalhada na seção 4.4.1.

O algoritmo foi desenvolvido em meio ao projeto de pesquisa e desenvolvimento composto por pesquisadores do DESA-UFMG e financiado pela CEMIG, que buscava o monitoramento inteligente da qualidade da água em reservatórios hidrelétricos através do desenvolvimento de um algoritmo fotogramétrico. Utilizou-se metodologia de modelagem empírica, cuja essência é a regressão, técnicas matemáticas foram utilizadas para inferir a concentração de clorofila-a por meio dos pares de dados que combinam as informações espectrais capturadas pelo sensor hiperespectral e dos dados do parâmetro coletado no reservatório, que são coletados simultaneamente.

O algoritmo foi desenvolvido em linguagem de programação Python, onde foram inseridos os modelos de regressão linear, *Support Vector Regressor*, *Random Forrest Regressor* e *Extreme Gradient Boosting*. O algoritmo roda sequencialmente as etapas de separação dos pares de dados, subdividindo tais dados em registro de entrada (80%) para calibração e validação (20%). Nos dados de calibração os hiperparâmetros são testados, selecionando os que resultam em melhores valores de RMSE para compor o modelo. Logo após, os dados de calibração são utilizados novamente, subdivididos em 5 grupos, em que 4 são utilizados para realizarem a calibração do modelo e um para teste de performance. Nesta etapa foram realizadas 100 repetições, obtendo-se 500 valores de métricas calculadas (Validação K-fold), o que possibilita a inferência da generalização ou não do modelo para o conjunto de dados de entrada.

Por fim, segundo Assunção (2023), após a verificação da capacidade de generalização, o modelo calibrado retorna aos 20% dos pares de dados separados anteriormente para o cálculo da performance de dados não calibrados. Desta forma, é possível verificar a robustez do modelo. Ao final do processo, o algoritmo retorna os

resultados de todas as métricas de validação, tanto as verificadas durante a calibração quanto as utilizadas para validação do modelo ( $R^2$  e RMSE), sendo registrados em gráficos e tabelas para a tomada de decisão em relação à escolha do modelo que melhor se ajustou aos dados. No presente estudo, foram selecionados no total 89 pares de dados que foram divididos nos grupos de calibração 62 amostras e, no segundo grupo, foram utilizados 17 pares de dados para a etapa de validação.

## 5 RESULTADOS E DISCUSSÃO

### 5.1 Caracterização da área de estudo com dados históricos de monitoramento

A análise estatística dos dados foi realizada para os três pontos amostrais monitorados pelo IGAM no Reservatório da Pampulha, de forma trimestral, no período entre 2013 e 2021. As coletas ocorrem nos meses de março, junho, setembro e dezembro, no entanto, os dados referentes ao mês de setembro de 2017 não constavam na planilha disponibilizada pelo repositório institucional do IGAM, totalizando 35 campanhas, 18 para o período chuvoso e 17 para o período seco.

Para o parâmetro microcistina 13% dos dados também não estão apresentados no banco de dados, além de 24% de dados censurados (abaixo do limite de detecção). Saxitoxina apresentou 13% e 10% de dados faltantes e censurados, respectivamente. Costa (2021), enfatiza que a porcentagem limite para exclusão de um parâmetro seria a existência de fração menor que 60% dos dados válidos, ou seja, mais de 40% de dados faltantes. Desta forma nenhum parâmetro foi retirado das análises.

Foram retirados ainda, valores discrepantes que estavam bem acima dos valores normalmente observados para os parâmetros em questão e que possivelmente foram inseridos de forma errônea. Ao todo seis dados foram subtraídos, duas para temperatura da água, uma para nitrito e três para turbidez.

As análises estatísticas foram calculadas através do software RStudio. O Apêndice B apresenta os dados de estatística descritiva para cada ponto, subdivididos em período chuvoso e de estiagem, estes resultados serão retratados no presente estudo por meio de gráficos *box-plot*. Foram realizados os testes de comparação não paramétricos de Mann-Whitney, cujo resultados estão dispostos no Apêndice C, para comparações entre o período seco e chuvoso para todos os parâmetros. O teste de Kruskal-Wallis foi conduzido para verificar se existem diferenças significativas entre os pontos amostrais, complementado pelo teste post-hoc de Dunn, que aponta onde se encontram as variações significativas que venham a ser observadas. Os resultados podem ser observados inseridos na Figura 9 e no Apêndice D.

A temperatura da água apresentou variação entre 18,80 °C e 31,10 °C. Foram encontradas, através do teste de Mann-Whitney (Apêndice C), variações

estatisticamente significativas entre os períodos chuvoso e seco em todos os pontos amostrais, seguindo a tendência sazonal do clima na região. O período chuvoso, caracterizado por temperaturas elevadas no local de estudo, apresentou maiores medianas em todos os pontos com valores pouco superiores a 26 °C, já no período seco as medianas dos três pontos giraram em torno de 22 °C. Não foram verificadas diferenças entre os três pontos amostrais na comparação múltipla (Kruskal-Wallis).

A temperatura da água influencia fortemente na dinâmica dos outros parâmetros na Lagoa da Pampulha, assim como nos reservatórios tropicais em geral (GIANI et al., 2020). Seus efeitos estão relacionados com o transporte e variação vertical de nutrientes, fitoplâncton, pH e oxigênio. Segundo Plec et al. (2021), durante o verão quente e chuvoso são observados períodos de forte estratificação térmica no reservatório, que consegue resistir a maioria dos eventos chuvosos, estes precisam ser de grande intensidade para desestabilizar a coluna d'água e o fluxo das precipitações de menor intensidade tendem a afluir pelo fundo do reservatório, diminuindo sua temperatura e a profundidade da zona de mistura. No período seco e frio há maior instabilidade e processos de desestratificação térmica ocorrem com maior constância, provocando a mistura da água e a ressuspensão dos compostos sedimentados.

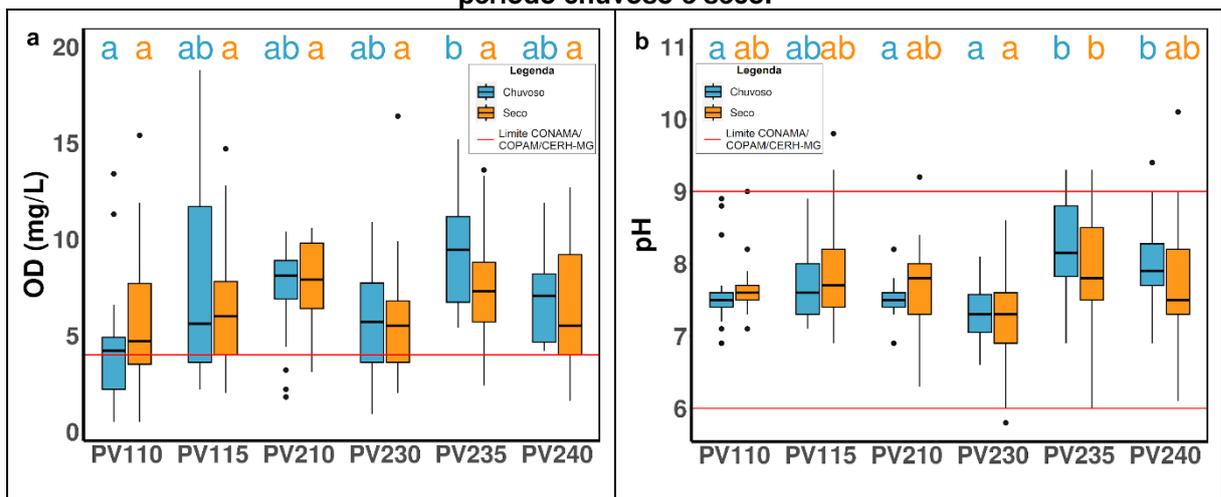
A ocorrência de processos de mistura da coluna d'água e posterior estratificação tem influência direta na concentração do fitoplâncton, principalmente das cianobactérias, pois a maior concentração de nutriente devido à redistribuição e o retorno à estratificação durante este período climático favorecem estes organismos dotados de adaptações a estas variações (BELISCO, 2017; SILVA et al., 2019). No final da estação seca, quando se soma a estes fatores o aumento da concentração de nutrientes, devido à menor diluição ocasionado pelo período sem chuvas, e o aumento gradual da temperatura, ocorrem no reservatório constantes florações de cianobactérias (BATISTA; FIGUEIREDO; GIANI, 2018).

Na Figura 9 são apresentados os gráficos *Box-plot* apontando os valores mínimo, máximo, primeiro e terceiro quartis, mediana, *outliers* e os resultados dos testes de Kruskal-Wallis e Dunn. Com relação aos dados de OD através da Figura 9a, pode-se observar que na lagoa, no período de estiagem, todos os pontos apresentaram amostras inferiores ao limite legal para a classe 3, estabelecido pela resolução

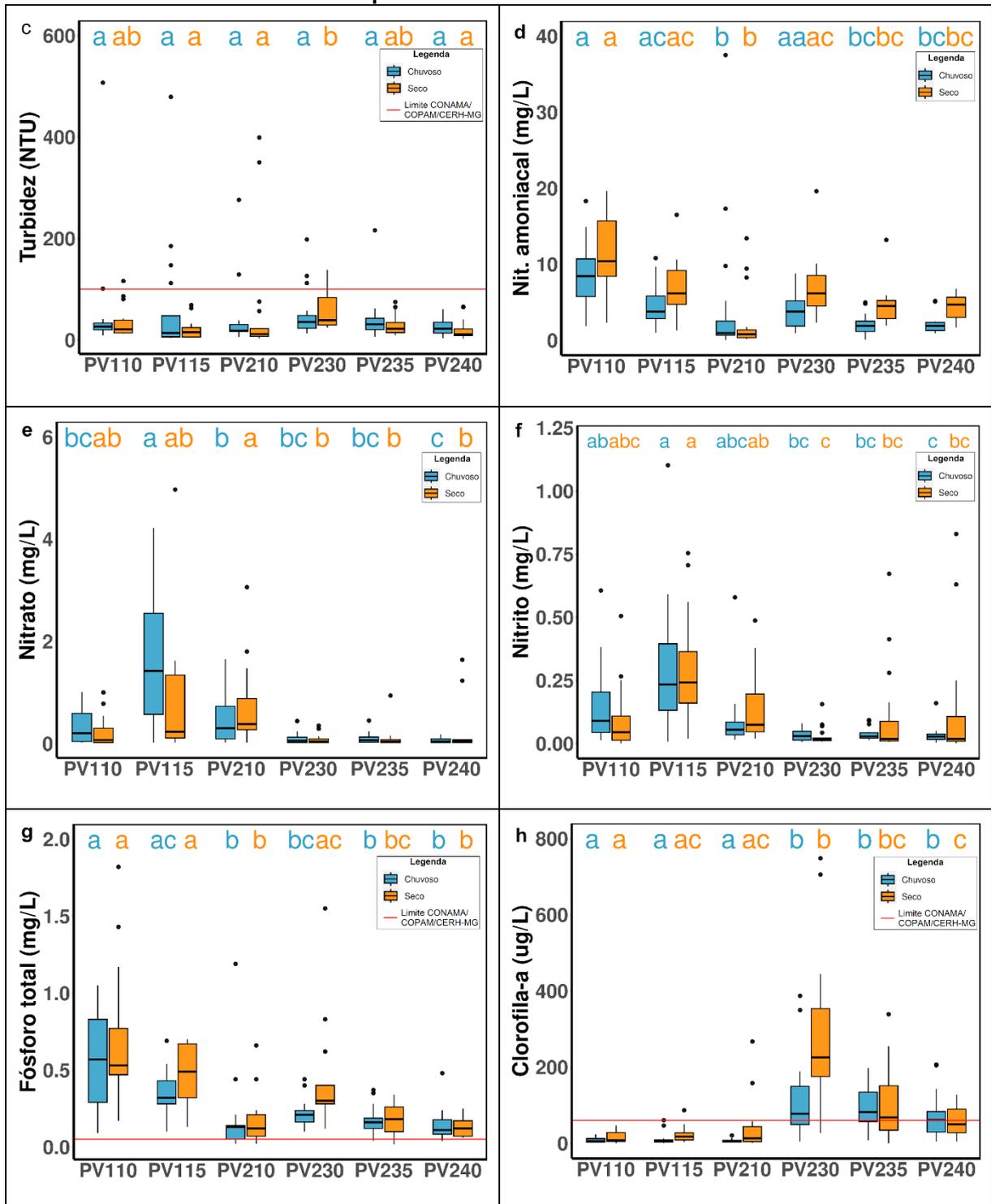
CONAMA 357/05 (4 mg/L) e, na estação chuvosa, somente o ponto PV230 apresentou amostras com concentrações de OD abaixo de 4 mg/L. No entanto, em todos os pontos e em ambos os períodos climáticos observados foi possível constatar níveis medianos de OD acima do limite estabelecido. Fato explicado pela intensa produção primária e constantes florações de algas que produzem oxigênio durante a fotossíntese (BARÇANTE et al., 2020). Porém, não foram identificadas variações sazonais significativas nas concentrações do parâmetro, provavelmente influenciado pela oxigenação provocada pela produção primária, que é elevada durante todo o ano.

Os resultados para o teste de Kruskal-Wallis inseridos na Figura 9a indicaram que não houve diferença significativa entre os pontos dos analisados no período seco. No período chuvoso, para os pontos inseridos no reservatório, o ponto PV230 teve menores índices de OD, porém sem diferença estatisticamente significativa com os outros pontos. Mesmo com altos índices de produção primária, no ponto PV230, as menores concentrações de OD observadas podem se relacionar com a entrada de poluentes orgânicos e com a morte do fitoplâncton, que tendem a diminuir os níveis de oxigênio durante sua decomposição. Considerando os tributários, o ponto PV110, córrego Ressaca, apresentou os piores índices de OD.

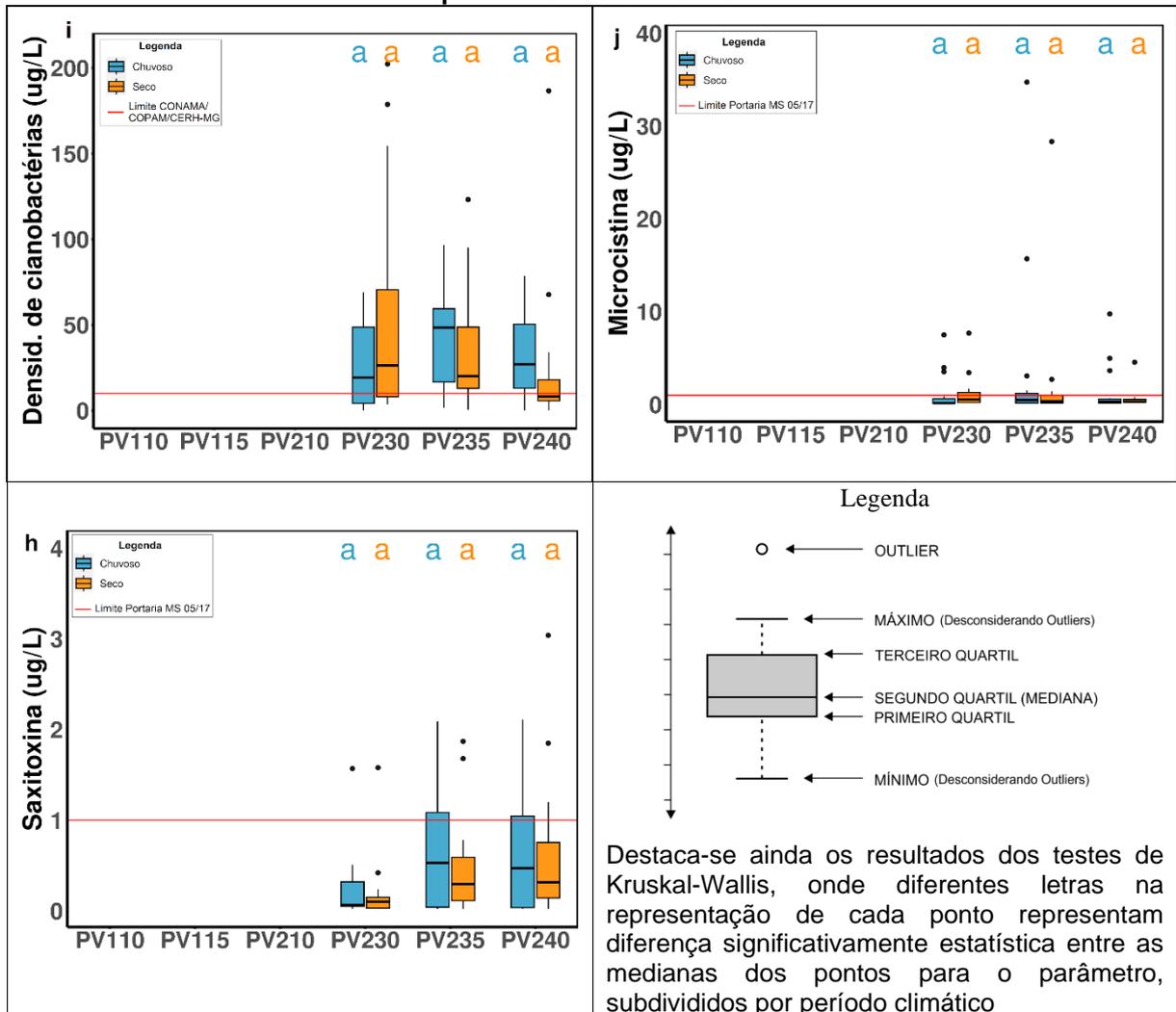
**Figura 9 – Box-plot com a distribuição dos dados analisados no presente estudo para os pontos PV110, PV115, PV210, PV230, PV235 e PV240 entre 2013 e 2021, subdivididos em período chuvoso e seco.**



**Figura 9 – Box-plot com a distribuição dos dados analisados no presente estudo para os pontos PV110, PV115, PV210, PV230, PV235 e PV240 entre 2013 e 2021, subdivididos em período chuvoso e seco.**



**Figura 9 – Box-plot com a distribuição dos dados analisados no presente estudo para os pontos PV110, PV115, PV210, PV230, PV235 e PV240 entre 2013 e 2021, subdivididos em período chuvoso e seco.**



Fonte: Autor (2023)

A resolução CONAMA 357/2005 impõe o padrão entre 6 e 9 ao pH. Poucas amostras ultrapassaram este limite nos dados observados, uma no ponto PV230 apresentou caráter ácido e nos pontos PV235 e PV240 três amostras em cada ponto apresentam caráter básico. Na Figura 9b podemos verificar a posição das amostras analisadas em relação a estes limites normativos. Estes resultados são importantes, pois níveis elevados de pH são associados à ocorrência de amônia, substância que apresenta toxicidade à biota e representa uma fração do N. Am., parâmetro presente em altas concentrações no reservatório (Figura 9d). Além disso, variações dos índices de pH além do limite estabelecido podem provocar efeitos adversos diretamente na biota aquática, em seus processos biológicos e no aparecimento de íons livres de metais pesados em condições de acidez do meio.

A resolução CONAMA 357/2005 estabelece limites variados de N. Am. a depender do pH, 13,3 mg/L de N ( $\text{pH} \leq 7,5$ ), 5,6 mg/L de N ( $7,5 < \text{pH} \leq 8,0$ ), 2,2 mg/L de N ( $8,0 < \text{pH} \leq 8,5$ ) e 1,0 mg/L de N ( $\text{pH} > 8,5$ ). Esta subdivisão ocorre porque a concentração das frações de amônia e o íon amônio estão relacionados com o pH do meio, onde há aumento da concentração de amônia devido ao aumento do pH (VON SPERLING, 2005). Ao se comparar as concentrações individuais das concentrações de N.Am. e de pH foram encontradas pares de amostras destes parâmetros que excedem os limites estabelecidos, citados acima, em 17% das amostras. Desta forma, uma análise mais detalhada é necessária para entender o comportamento dos componentes de N. Am. total no reservatório.

Na Figura 9b, podemos observar que o pH em todo o reservatório tende a alcalinidade. O ponto PV230 apresentou resultados próximos à normalidade e os pontos PV235 e PV240 mais básicos, com medianas próximas a 8. Nos testes de hipóteses, Kruskal-Wallis e Dunn, representado na Figura 9b, observa-se que os pontos representando os tributários, PV110, PV115 e PV210 se assemelham ao PV230, Bezerra et al. (2022) citam que a qualidade da água do reservatório na região próxima aos córregos Ressaca e Sarandi é consideravelmente influenciada por estes, o que diminui a eficiência dos tratamentos realizados no reservatório em função da inserção constante de efluentes poluídos.

Dentro da Lagoa da Pampulha, foi possível constatar diferença significativa entre o ponto PV230 e os outros dois pontos no período chuvoso e entre o ponto PV230 e PV235 no período seco, sendo que os pontos PV230 e PV240 não apresentaram diferença neste período climático. Maiores valores de pH nos pontos PV235 e PV240 podem estar relacionados com a fotossíntese, nestes pontos há menor influência dos tributários e altas taxas de produção primária (Figura 9h), que, segundo Esteves (2011), sequestra o  $\text{CO}_2$  e aumenta o pH meio.

A resolução CONAMA 357/2005 estipula que o limite para a turbidez na classe 3 é de 100 Unidades Nefelométricas de Turbidez (UNT). O ponto PV230, a montante do reservatório, apresentou 6 observações acima do limite previsto na legislação (17%) e o ponto PV235 apenas uma amostra ultrapassou o limite legal, já no ponto PV240 nenhuma amostra acima de 100 UNT foi observada. Além disto, este ponto apresentou as menores medianas, Figura 9d, este comportamento se deve ao fato do

ponto se localizar a jusante do reservatório, próximo à barragem, onde a maior parcela dos sólidos com maior granulometria, que conferem turbidez à água, se sedimentaram (LOPES et al., 2019).

Não foi encontrada diferença estatística significativa no teste de Mann-Whitney entre os períodos seco e chuvoso nas amostras analisadas para turbidez, este fato pode estar associado com o aumento da concentração de algas, que é mais intensa no reservatório durante o período seco (Figura 9h) e que confere aos corpos d'água maiores índices de turbidez.

As concentrações médias de PT (Figura 9g), nitrato (Figura 9f) e N. Am. (Figura 9d), foram significativamente superiores nos tributários dos córregos Ressaca e Sarandi, pontos PV110 e PV115, respectivamente. O ponto PV230 (na região à montante do reservatório), que recebe as contribuições dos córregos Ressaca e Sarandi, apresentou altas concentrações de PT, N. Am. e orgânico, apresentando semelhança nas concentrações destes parâmetros com tais tributários. Com isto, no ponto PV230 foram observadas altas taxas de produção primária no ponto (Figuras 9h e 9i, que representam as concentrações de chl-a e DC, respectivamente). Bezerra et al. (2022), observaram comportamento semelhante ao analisarem as concentrações de PT e nitrogênio total nestes tributários e na lagoa, comprovando que há forte influência destes córregos na qualidade da água da lagoa e que as ações para revitalização são prejudicadas devido ao contínuo aporte de nutrientes desses tributários.

O íon amônio e o nitrato são as principais formas assimiladas pelos produtores primários, sendo de grande importância ecológica. O nitrogênio orgânico em sua forma dissolvida, como a ureia e aminoácidos, também pode ser utilizado como fonte deste nutriente pelos organismos (ESTEVES, 2011). O nitrogênio pode ainda ser um indicador da distância da fonte de poluição em um corpo d'água, formas reduzidas são encontradas em locais próximos aos despejos poluidores e as formas oxidadas em locais distantes da fonte inicial de contaminação, sendo que o nitrato é o último composto desta cadeia (VON SPERLING, 2005).

A Figura 9e e 9f demonstram tal comportamento, as principais fontes de poluição, os córregos Ressaca e Sarandi, PV110 e PV115, apresentam altos índices de nitrato e nitrito (depois de percorrer toda a bacia destes córregos). Ao ponto que maiores concentrações de N. Am. no reservatório (Figura 9d) se encontram no ponto PV230,

próximos a estes tributários, decaindo ao longo do reservatório. A Figura 9g mostra que as concentrações de PT apresentam características similares, nos tributários e no reservatório, aos observados para N. Am. que é uma das principais fontes de nitrogênio para a produção primária. Com isto, potencializa-se a ocorrência de altas concentrações de fitoplâncton na região do ponto PV230, principalmente no período seco (Figuras 9h e 9i).

Segundo Plec et al. (2021) a Lagoa da Pampulha apresenta tempo de residência da água de 89 dias (média anual), desta forma, para águas de classe 3 o padrão da resolução 357/2005 do CONAMA impõe para o PT em ambiente lântico um limite de 0,05 mg/L. O *Box-plot* da Figura 9c mostra que no ponto PV230 todas as amostras coletadas durante o período analisado superaram este limite. Apesar de apresentarem menores medianas, os pontos PV235 e PV240 também ultrapassaram o limite legal na maioria das amostras (94% em ambos os pontos), considerando ambos os períodos climáticos da lagoa.

Em relação aos tributários, observou-se altas concentrações do parâmetro nos pontos PV110 e PV115, córregos Ressaca e Sarandi, respectivamente. O ponto PV230 situado próximo ao ponto de afluição destes tributários apresentou semelhança estatística significativa com eles no período seco, já no período chuvoso, apenas com o ponto PV115 apresentou tal característica, com o ponto PV110 apresentando concentrações mais elevadas que ambos. No reservatório, a diferença estatística entre os pontos ocorreu apenas no período seco, entre os pontos PV230 e PV240 (à jusante do reservatório)

Estes resultados comprovam as condições de hipertrofia no reservatório, que nas últimas décadas teve a dinâmica do fósforo modificada como resultado da intervenção humana em sua bacia hidrográfica, introduzindo altas cargas de esgoto sanitário sem tratamento. Constantes florações de algas vêm sendo reportadas neste período, incluindo as cianobactérias, como reportado em diversos estudos que analisaram o reservatório (BEZERRA et al., 2022; BRANDÃO et al., 2016; GIANI et al., 2020). Na Lagoa da Pampulha o incremento constante de altas concentrações de fósforo pelos seus tributários, que ultrapassaram os limites estabelecidos pela CONAMA 357/05 para ambientes lóticos (0,15 mg/L), contribui substancialmente na manutenção das

altas concentrações do parâmetro, principalmente no ponto PV230 (local próximo aos principais tributários), onde a concentração do PT foi maior (BARÇANTE et al., 2020).

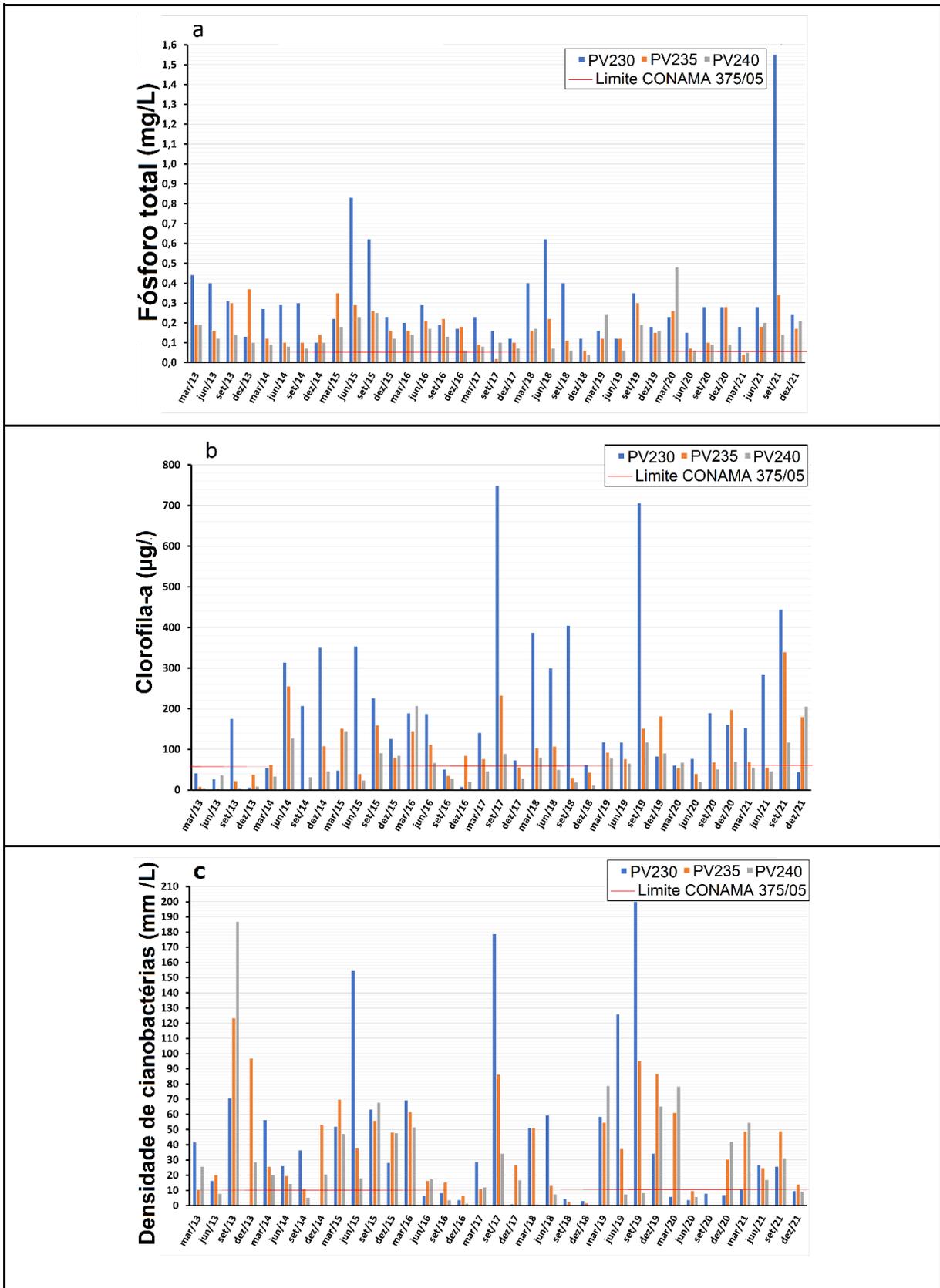
Bezerra et al. (2022) analisaram as concentrações de PT no reservatório e nos tributários entre janeiro e abril de 2018, foram observadas maiores concentrações do parâmetro nos tributários, porém com menor diferença em relação aos pontos monitorados no reservatório, médias para todos os pontos monitorados de 0,05 mg/L e 0,04 mg/L, respectivamente. Desta forma, os autores não observaram diferenças significativas entre os dados observados entre as concentrações medianas de PT no reservatório e nos tributários, o que denota, segundo eles, a influência dos córregos nas concentrações do parâmetro na Pampulha.

A Figura 10a expõe a variação temporal de fósforo total na Lagoa da Pampulha nos três pontos que o IGAM monitora no reservatório. Observa-se que as concentrações de fósforo no reservatório se mantêm em níveis elevados durante todo o período analisado. Este cenário se mantém mesmo com as ações realizadas pelos órgãos responsáveis por gerir o reservatório, como a aplicação dos produtos comerciais Phoslock® e *Enzilimp*®, que tem como objetivo reduzir as frações de fósforo e nitrogênio na lagoa.

Barçante et al. (2020) observaram que após o tratamento com Phoslock® entre os anos de 2016 e 2018 houve diminuição dos níveis de fósforo e da biomassa de cianobactérias, provocando o aparecimento de uma comunidade fitoplanctônica mais diversificada em alguns meses ao final do tratamento. Porém, os autores relataram uma tendência de retorno das cianobactérias ao final do estudo, fato atribuído aos altos índices de fósforo no reservatório.

De fato, Bezerra et al. (2022), chegaram à conclusão de que após 2 anos de tratamento com o Phoslock® e *Enzilimp*®, apenas pequenos efeitos foram mensurados na qualidade da água. No entanto, mesmo menores, as concentrações de fósforo e nitrogênio ainda são expressivas devido ao contínuo aporte de efluentes sem tratamento oriundo da bacia hidrográfica. Segundo os autores, os córregos Ressaca e Sarandi configuram as principais fontes de nutrientes e matéria orgânica e a falta de ações que controlem a influência dos tributários faz com que haja maior resistência para reverter as condições de hipertrofia ao qual o reservatório se encontra.

**Figura 10 - Gráfico temporal das concentrações de fósforo total (a), clorofila-a (b), e densidade de cianobactérias (c) nos pontos PV230, PV235 e PV240 para o período analisado (2013 a 2021), subdivididos em período chuvoso e seco**



Fonte: Autor (2022)

A variação de chl-a no período estudado apresentou grande amplitude, 0,01 a 748,09 µg/L, a Figura 9h mostra que as medianas dos pontos PV230 e PV235 ultrapassaram o limite legal da Resolução CONAMA 357/05 para águas de classe 3 que é de 60 µg/L, o ponto PV240 apresentou mediana pouco superior a este limite no período chuvoso (60,90 µg/L) e no período seco a mediana foi inferior ao limite estabelecido. Sendo que, neste ponto, 45% das observações ultrapassaram tal limite, nos pontos PV230 e PV235 essas porcentagens foram de 68% e 57%, respectivamente. Estudos anteriores, Barçante et al. (2020); Belisco (2017), reportaram altos níveis de chl-a na Lagoa da Pampulha, os autores concluíram que a alta concentração de fósforo é um dos principais fatores para este resultado. Resultados apresentando menores concentrações de chl-a nos pontos PV235 e PV240 podem ser explicados pelos processos de sedimentação dos nutrientes e decomposição da matéria orgânica ao longo do reservatório, que diminuem o substrato disponível para o crescimento das algas (LOPES et al., 2019).

Em relação ao reservatório, o teste de Kruskal-Wallis (Figura 9h) apontou que no período seco houve diferença significativa entre as medianas dos pontos amostrais. O teste post-hoc de Dunn indicou que a diferença ocorreu entre os pontos PV240 e PV 230, que apresentou concentrações de chl-a muito superiores. A alta concentração do parâmetro neste ponto no período seco impactou na comparação em relação à sazonalidade e somente o ponto PV230 apresentou diferença significativa no teste de Mann-Whitney (Apêndice C). A proximidade deste ponto com os principais afluentes, Ressaca e Sarandi, que possuem altas cargas de efluentes sem tratamento e a influência do tempo seco podem ser responsáveis por estes resultados (BARÇANTE et al., 2020). Boëchat et al. (2019); Giani et al. (2020) enfatizam em seus estudos que em regiões tropicais os baixos índices de pluviosidade característicos do período de estiagem provocam menor diluição da poluição incidente, como nutrientes e matéria orgânica, o que aumenta a concentração dos poluentes e favorece o aumento da produção primária.

A Figura 10b mostra a variação temporal de Chl-a no período estudado, nota-se tendência de aumento em todos os pontos a partir do ano de 2019, observa-se também que o ponto PV230 apresenta concentrações do parâmetro acima do limite legal para o corpo d'água na maioria das observações, principalmente no período

seco, comportamento observado na Figura 9h, ponto este que também apresentou as maiores concentrações de PT no reservatório.

O limite para a DC para a classe 3, orientado pela CONAMA 357/05, é de 100.000 cel/ml ou 10 mm<sup>3</sup>/L. Na Figura 9i podemos observar que a maioria das amostras ultrapassaram o limite legal. No ponto PV235, 85% dos resultados foram superiores ao estabelecido pelo CONAMA, já os pontos PV230 e PV240 apresentaram 68% e 60% de amostras acima do padrão, respectivamente (valores para os dados agrupados). Somente o ponto PV240, no período seco, teve mediana inferior ao limite para o parâmetro e mostrou um padrão similar a chl-a.

O monitoramento do IGAM não contempla a DC e suas cianotoxinas nos tributários da Lagoa da Pampulha. Não foram observadas diferenças significativas entre os pontos amostrais (Figura 9i) e entre os períodos seco e chuvoso (Apêndice B). Estes resultados mostram o comportamento homogêneo da biomassa de cianobactérias em todo o reservatório em todos os períodos do ano, que representa a capacidade de adaptação das cianobactérias aos diferentes ambientes do reservatório. Figueiredo et al. (2016) constataram que as florações de cianobactérias na Lagoa da Pampulha passaram de sazonais para persistentes nas últimas décadas, sendo que o PT foi o principal determinante para estas florações. Os autores observaram que os níveis deste parâmetro aumentaram 3 vezes no período estudado (entre 1993 e 1996 e 2010 e 2011), favorecendo as cianobactérias que possuem alta adaptabilidade a ambientes eutrofizados.

No estudo de Batista, Figueiredo e Giani (2018), este comportamento também foi observado, os autores determinaram as espécies dominantes em cada período climático, *Cylindrospermopsis raciborskii* na estação chuvosa e *Microcystis spp* na estação seca. Além disso, os fatores determinantes para as florações destas espécies foram mensurados, sendo que na estação chuvosa *Cylindrospermopsis raciborskii* se correlacionou fortemente com a estratificação e com a precipitação e na estação seca *Microcystis spp* se relacionou fortemente com o PT.

A Figura 10c mostra a variação temporal de DC no período estudado, foi possível perceber uma tendência de queda para o parâmetro, somente o ponto PV240. Estes resultados associados ao aumento dos índices de chl-a podem ser um indicador do aumento da biodiversidade na produção primária, que pode ser provocado pelas

intervenções no reservatório ou por mudanças nos fatores que controlam a dominância destas espécies, como alterações no regime climático. No entanto, os altos índices de fósforo observados mantêm o reservatório em condições hipereutróficas, criando condições para que florações destas espécies continuem a acontecer no reservatório. Concentrações acima de 0,1 mg/L foi sugerida como ponto em que há tendência à dominância de cianobactérias em ecossistemas aquáticos (BARÇANTE et al., 2020), na Figura 10a podemos observar que em sua maioria, as concentrações de PT ultrapassam este limite, o que pode explicar os altos índices de DC observados.

A Resolução CONAMA 357/05 não cita a microcistina em sua resolução, no entanto há um limite estabelecido pela portaria número 5 de 2017 do Ministério da Saúde, alterada pela portaria GM/MS número 888 de 2021 (MS Nº 5/2017), que prevê o limite máximo permitido em águas para consumo humano de 1 µg/L. Mesmo não sendo utilizada como fonte de abastecimento público, a MS Nº 5/2017 foi utilizada como padrão para comparação neste estudo. No geral, não foram observados muitos valores acima deste limite (Figura 9j), porém foi possível constatar a presença desta toxina em todo período analisado.

Segundo Barçante (2019), que monitorou a presença de cianotoxinas na lagoa da Pampulha entre abril de 2016 e maio de 2018, assim como no presente estudo, foi constatada a presença de microcistina no reservatório em todo o período avaliado, se associando à presença da espécie *Microcystis aeruginosa*, predominante nas florações na Lagoa da Pampulha, principalmente no período seco (BATISTA; FIGUEIREDO; GIANI, 2018). Como se pode observar na Figura 9j, a maioria das ocorrências extremas de microcistina ocorreu no período chuvoso, período no qual *Microcystis aeruginosa* não é dominante, estes resultados podem estar associados a fatores como a coexistência de genótipos tóxicos e não tóxicos de uma mesma espécie e as condições ambientais no momento da floração (BARÇANTE, 2019).

Assim como a microcistina e DC não há monitoramento de saxitoxina pelo IGAM nos tributários da Lagoa da Pampulha. Este parâmetro também não é contemplado na resolução 357/2005 do CONAMA, na portaria MS Nº 5/2017 está previsto para este parâmetro o limite de 3,0 µg/L. Apenas uma amostra ultrapassou este limite, no ponto PV240 (Figura 9h). O parâmetro não apresentou variações entre os períodos seco e

chuvoso. De qualquer forma, observou-se que esta toxina também esteve presente no reservatório durante todo o período avaliado. Barçante (2019) correlacionou a presença desta toxina com o gênero *Planktothrix* spp e *Cylindrospermopsis raciborskii*, e segundo Batista, Figueiredo e Giani (2018) a espécie *Cylindrospermopsis raciborskii* apresentam predominância na estação chuvosa, associada fortemente com a estratificação térmica e precipitação, período no qual foram observadas maiores concentrações de saxitoxina na lagoa no presente estudo.

No período chuvoso não foram encontradas variações significativas entre os pontos amostrais. No período seco foi verificada diferença significativa entre os pontos PV230 e PV240, sendo que estes são semelhantes ao ponto PV235 (Figura 9h). Estes resultados podem ter sido originados pela maior adaptabilidade dos gêneros de cianobactérias que produzem saxitoxina nos pontos a jusante do reservatório, principalmente o PV240, que apresentou menor DC e maior concentração de saxitoxina.

## **5.2 Caracterização da qualidade da água na lagoa da Pampulha através de medições obtidas com a sonda multiparâmetros**

Na análise dos dados primários, obtidos em coletas mensais entre março de 2022 e fevereiro de 2023, foi possível observar que a região a montante do reservatório, pontos 1 a 10, detém os piores índices de qualidade da água. A mesma tendência foi observada nos dados secundários, onde o ponto PV230 que está localizado nesta região do reservatório, apresentou maior quantidade de parâmetros com índices de qualidade não satisfatórios.

Foi realizado o teste não paramétrico de Mann-Whitney na análise estatística dos dados coletados com a sonda para a comparação entre os dados subdivididos em antes e depois da cortina retentora de resíduos ter sido introduzida na lagoa em abril de 2016, os resultados estão dispostos na Tabela 7. Todos os parâmetros apresentaram diferença estatisticamente significativa entre as regiões e as medianas da região antes desta cortina, a montante da lagoa, apresentaram pior qualidade da água.

**Tabela 7 - Resultado para o teste de Mann-Whitney comparando as concentrações dos parâmetros observados antes e depois da cortina retentora de resíduos na Lagoa da Pampulha**

Parâmetro	p Valor	Diferença significativa	Mediana antes (n = 120)	Mediana depois (n = 240)
chl-a (µg/L)	$2,20 \times 10^{-16}$	Sim	58,51	28,10
Ficocianina (µg/L)	0,0001	Sim	9,86	6,10
pH	$6,79 \times 10^{-16}$	Sim	7,36	7,85
OD (mg/L)	$4,19 \times 10^{-5}$	Sim	6,48	7,20
Turbidez (FNU)	$1,07 \times 10^{-15}$	Sim	36,25	14,55
Disco de Secchi (m)	$2,20 \times 10^{-16}$	Sim	0,2	0,5

Fonte: Autor (2023)

Como possíveis causas destes resultados podemos citar a proximidade com os córregos Ressaca e Sarandi, que tem aporte contínuo de altas cargas de esgoto doméstico e industrial (BEZERRA et al., 2022) e o aporte de sedimentos carreados da bacia hidrográfica (SILVA et al., 2019). A poluição incidente provoca piora significativa da qualidade da água nesta região, com efeito em toda o reservatório, provocando problemas como constantes florações de algas, principalmente cianobactérias (FIGUEIREDO et al., 2016) e o assoreamento (MELO, 2018). Outro fator que pode explicar estes resultados é a profundidade da região anterior à cortina. Segundo Resck; Bezerra-Neto; Coelho, (2007), que mapearam a batimetria da Lagoa da Pampulha, nesta porção da lagoa, as profundidades são inferiores a 3 metros. Com isto, potencializa-se a produtividade biológica que é maior em águas mais rasas. Ainda segundo os autores, esta região ainda é suscetível a ressuspensão do sedimento, que alimenta a coluna d'água com cargas extras de fósforo total.

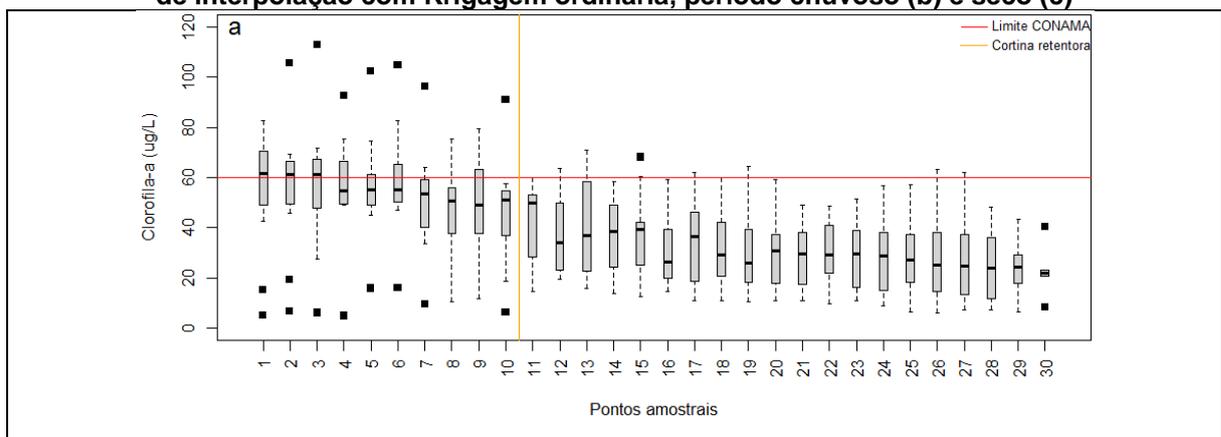
O teste de Mann-Whitney foi utilizado também para comparação entre os dados primários (Sonda) e secundários (IGAM), também subdivididos em seco e chuvoso. Na comparação entre os dados primários coletados entre mar/2022 e fev/2023 e os dados secundários (2013 a 2021) foi possível constatar redução estatisticamente significativa entre as concentrações medianas de chl-a nos períodos seco e chuvoso, em ambos os casos as medianas dos dados primários apresentaram menores valores. Diferentemente do observado nos dados históricos do IGAM, a concentração de chl-a foi superior no período chuvoso, que compreendeu o mês de março de 2022 e os meses de outubro de 2022 a fevereiro de 2023. No período chuvoso a mediana foi 38,48 µg/L e no período seco 32,68 µg/L, ambos inferiores ao limite estabelecido pela Resolução CONAMA 357/05.

Buscou-se representar a espacialização dos parâmetros monitorados através do método de interpolação de Krigagem ordinária, os dados foram subdivididos para os períodos de chuva e de estiagem. A análise estatística foi demonstrada em gráficos *box-plot*, onde podem ser visualizados as medianas, valores mínimos, primeiro e terceiro quartis, valores máximos e os *outliers*, nesta análise foram considerados os dados agrupados de todos os meses amostrados. As próximas seções mostram os resultados para estas análises.

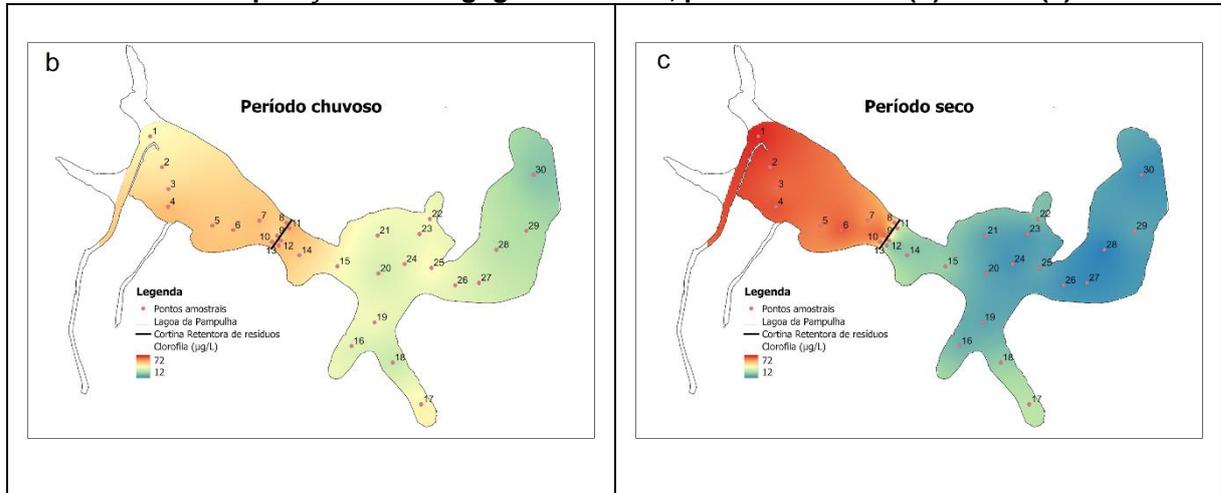
### 5.2.1 Clorofila-a

Através da Figura 11a pode-se observar a distribuição mediana das concentrações de chl-a ao longo de todo o reservatório da Pampulha, em vários pontos alguma observação ultrapassou o limite indicado para corpos d'água de classe 3 do CONAMA, que é de 60 µg/L. Em sua maioria, este fato ocorre a montante da lagoa, localizada próximo aos principais tributários. Esta porção do corpo d'água, historicamente, também concentra as maiores concentrações de chl-a e fósforo, de acordo com os dados analisados na seção 5.1, que analisa os dados secundários de chl-a do IGAM entre 2013 e 2021. Desta forma, os dados primários, obtidos com a sonda multiparâmetros, apresentam conformidade com a divisão espacial verificada nos dados observados na lagoa para os últimos anos.

**Figura 11 - *Box-plot* com a distribuição dos dados de chl-a nos 30 pontos amostrais (entre março/2022 e fevereiro/2023) (a) e espacialização dos valores medianos realizada por meio de interpolação com Krigagem ordinária, período chuvoso (b) e seco (c)**



**Figura 11 - Box-plot com a distribuição dos dados de chl-a nos 30 pontos amostrais (entre março/2022 e fevereiro/2023) (a) e espacialização dos valores medianos realizada por meio de interpolação com Krigagem ordinária, período chuvoso (b) e seco (c)**



Fonte: Autor (2022)

A cortina retentora de resíduos está representada na Figura 11a, entre os pontos 10 e 11. As concentrações de chl-a diminuem após esta barreira e os *outliers*, que representam dados discrepantes com o resto da amostragem, ocorrem predominantemente antes da cortina. Outliers observados na região anterior à cortina, menores que o restante da amostragem, foram originados da amostragem de janeiro que teve concentrações do parâmetro influenciados por um período de clima nublado, nos dias anteriores à coleta dos dados, que reduziu a produção primária.

Nota-se que no período seco são verificadas as maiores concentrações de chl-a a montante do reservatório (Figura 11b), que diminuem consideravelmente à medida que se avança para seu barramento. Isto ocorre pela alta inserção e disponibilidade de nutrientes a montante e o conseqüente declínio ao longo do reservatório, tanto pela assimilação biológica, quanto pela sedimentação (LOPES et al., 2020).

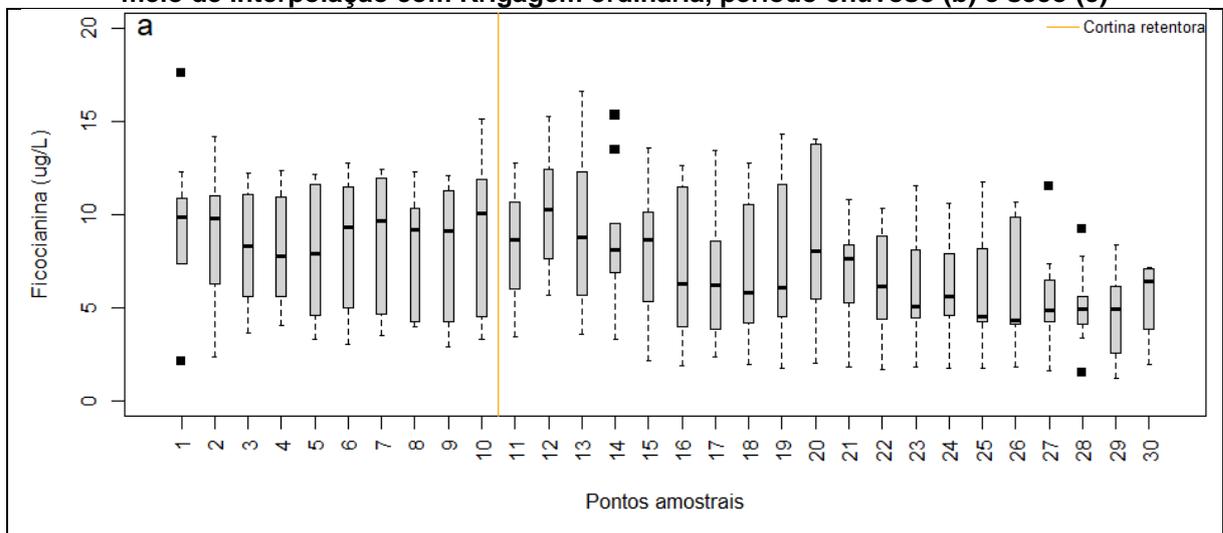
Houve diferença significativa em relação à sazonalidade para os dados de chl-a, sendo que o período chuvoso apresentou maiores concentrações do parâmetro. Neste período, a concentração de chl-a se torna mais homogênea, sendo que a região a montante ainda apresenta os piores índices para o parâmetro. Observa-se também que a região próxima aos pontos 16 e 17 apresentam maiores concentrações de chl-a em relação aos outros pontos da região, possivelmente em função da entrada de efluentes por meio dos córregos Mergulhão e Tijuco.

Bezerra et al. (2022), analisaram 10 pontos em seu estudo sobre a qualidade da água na Lagoa da Pampulha, houve comparação entre amostras coletadas nos anos de 2013 e 2018. Os resultados mostraram que não houve diferença significativa entre as amostras destes períodos e que as médias observadas não ultrapassaram em nenhuma das coletas o limite de 60 µg/L estabelecido pelo CONAMA. Estes resultados diferem do presente estudo, pois foram encontradas amostras em que a mediana ultrapassou este limite, nos pontos 1, 2, e 3 (Figura 11a), que se localizam próximos à afluência dos córregos Ressaca e Sarandi.

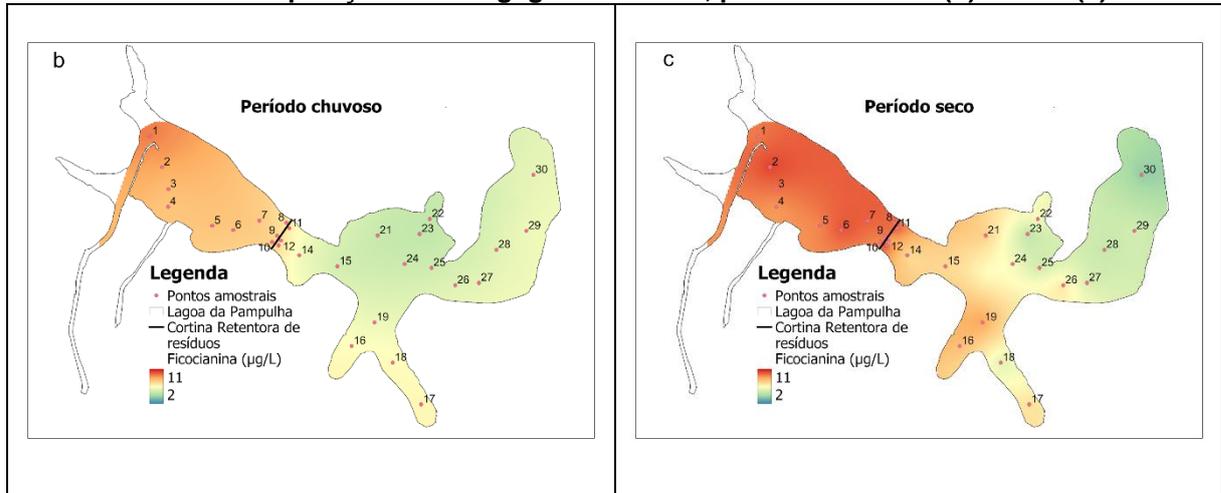
### 5.2.2 Ficocianina

A ficocianina é um pigmento encontrado em maior concentração nas cianobactérias e sua presença permite indicar a presença deste grupo do fitoplâncton na água (SALGADO, 2018). Na Figura 12a estão as concentrações medianas em cada ponto monitorado. A região a montante apresenta maiores concentrações, porém, assim como nos dados de chl-a.

**Figura 12 - Box-plot com a distribuição dos dados de ficocianina nos 30 pontos amostrais (entre março/2022 e fevereiro/2023) (a) e espacialização dos valores medianos realizada por meio de interpolação com Krigagem ordinária, período chuvoso (b) e seco (c)**



**Figura 12 - Box-plot com a distribuição dos dados de ficocianina nos 30 pontos amostrais (entre março/2022 e fevereiro/2023) (a) e espacialização dos valores medianos realizada por meio de interpolação com Krigagem ordinária, período chuvoso (b) e seco (c)**



Fonte: Autor (2022)

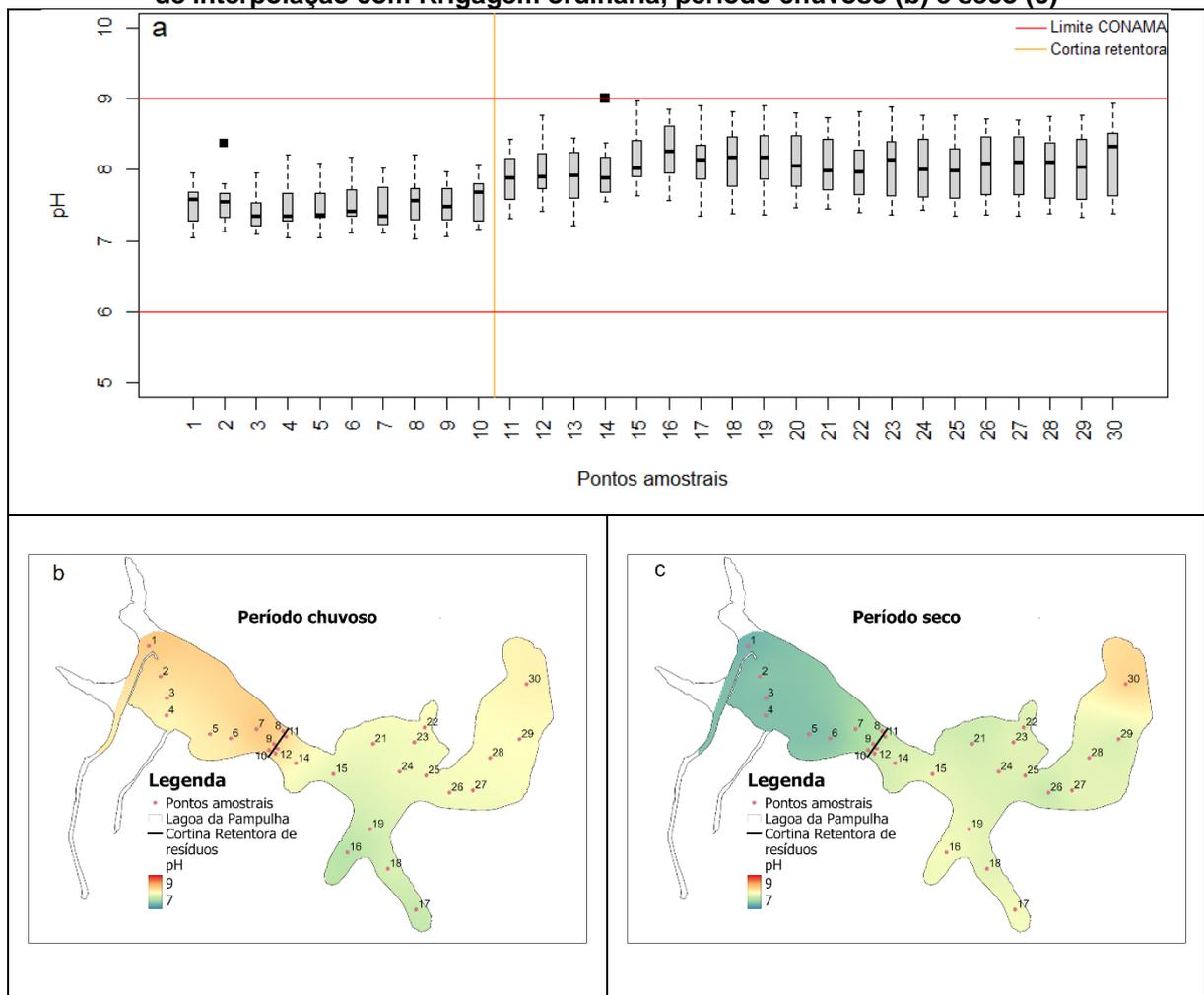
As Figuras 12b e 12c mostram a variação das concentrações de ficocianina no reservatório durante todo o período estudado (mar/2022 a fev/2023), subdivididos em período chuvoso e seco, respectivamente. A exemplo dos dados primários que avaliam as concentrações de DC (Figura 9b), os resultados denotam que no período chuvoso as concentrações de ficocianina, e em consequência a presença de cianobactérias, não apresentam variação significativa ao longo da Lagoa da Pampulha. Estes resultados indicam a boa adaptação das cianobactérias a todos os ambientes do reservatório no período citado.

Assim como para chl-a, a espacialização dos dados (Figuras 12c) demonstrou que o período seco apresentou maior variabilidade nas concentrações do parâmetro, influenciado pela maior produção primária na região a montante. Neste período as concentrações de ficocianina foram maiores próximo dos principais tributários (Ponto 1) e diminuíram ao se aproximar da barragem (Ponto 30). A região dos pontos 16 e 19 apresentou concentrações superiores que as proximidades, provavelmente sob influência das contribuições dos córregos Tijuco e Mergulhão, que deságuam nesta porção da Lagoa da Pampulha. No teste de sazonalidade este período apresentou maiores concentrações do parâmetro (8,27  $\mu\text{g/L}$ ) em relação ao período chuvoso (6,20  $\mu\text{g/L}$ ), valores que representam diferença significativa entre os períodos. Este padrão é contrário ao observado para as concentrações de chl-a para os dados primários e coincidentes com os resultados dos dados secundários avaliados na seção 5.1.

### 5.2.3 pH

O pH não apresentou observações que ultrapassam os limites indicados como ótimos para a meio aquático em corpos d'água de nível 3 do CONAMA, que se encontra entre 6 e 9 (Figura 13a). Este resultado se assemelha aos dados do IGAM entre 2013 e 2021, que apresentaram poucas amostras acima destes limites, Figura 9e. O ponto 16, localizado próximo a entrada do córrego Tijuco no reservatório, foi o único a apresentar diferença estatística da mediana em relação aos pontos 4 e 3, localizados próximo à Ilha dos Amores e ao ponto 7, próximo à cortina retentora de resíduos. Estes locais apresentam características físico-químicas, hidrodinâmicas, biológicas e de entrada de efluentes distintas, o que pode ter influenciado na diferença observada para os índices de pH.

**Figura 13 - Box-plot com a distribuição dos dados de pH nos 30 pontos amostrais (entre março/2022 e fevereiro/2023) (a) e espacialização dos valores medianos realizada por meio de interpolação com Krigagem ordinária, período chuvoso (b) e seco (c)**



Fonte: Autor (2022)

O pH apresentou diferença sazonal entre os períodos de chuva e de estiagem, com o período chuvoso apresentando pH mediano de 8,02 e o seco de 7,77. As Figuras 13b e 13c indicam que os valores de pH foram menores na porção superior da Lagoa da Pampulha, antes da cortina de retenção, nos dois períodos climáticos analisados, chuvoso e seco. Mesmo com as altas taxas de concentração de chl-a, que deveria aumentar o pH devido à fotossíntese, este ponto é rico em matéria orgânica, pela morte e senescência do próprio fitoplâncton e pela adição proveniente dos córregos Ressaca e Sarandi.

Os menores índices de pH podem estar associados à decomposição bacteriana da matéria orgânica que aumenta os níveis de CO<sub>2</sub> e diminui o pH (BRANDÃO, STAEHR E BEZERRA-NETO, 2016). Nota-se, que a região próxima à afluição dos córregos Tijuco e Mergulhão, pontos 16,17,18 e 19, maiores índices de pH são observados, a provável razão é a concentração de chl-a, que ainda é relativamente alta, e menor incidência de matéria orgânica em relação à região a montante do reservatório.

Souza (2019) analisou a qualidade da água do reservatório em abril de 2018, foram encontrados índices de pH inferiores ao do presente estudo, considerando o mesmo mês de amostragem, (7,8 ante 8,1, respectivamente). Mesmo com a diferença observada, os índices do parâmetro se encontram na faixa definida como limite pelo CONAMA, assim como observado nos dados históricos, entre 2013 e 2021 (Seção 5.1).

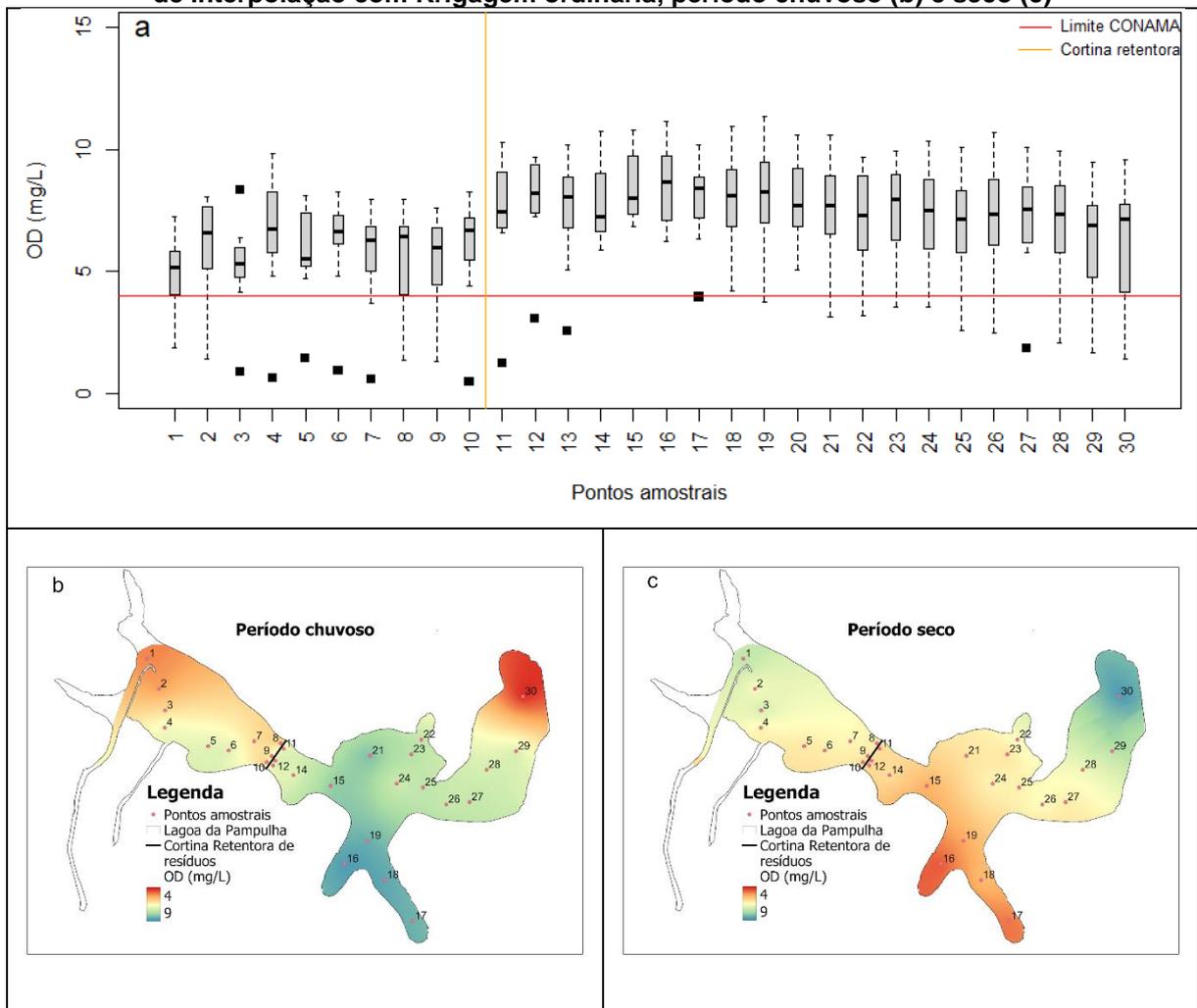
#### 5.2.4 Oxigênio dissolvido

A Figura 14a mostra que as concentrações de OD para os dados primários apresentaram menores concentrações a montante do reservatório. Estes resultados podem se relacionar com a decomposição de altas concentrações de matéria orgânica presente nesta parte da Lagoa que continuamente é abastecida com altas cargas de esgotos sanitários. Desta forma, mesmo com a intensa taxa de fotossíntese nesta porção do reservatório, que deveria elevar os índices de OD, estes processos citados tendem a promover a diminuição deste composto na água, efeito inverso ao citado na análise dos dados históricos (seção 5.1), onde o pH é afetado por esta dinâmica.

A campanha do mês de janeiro apresentou concentrações baixas de OD no reservatório, representadas por *outliers* em alguns pontos na Figura 14a. Esta campanha ocorreu em meio a episódios de precipitação e baixa incidência de radiação

que durou alguns dias. Segundo Belisco (2017) estes períodos têm efeito sobre a comunidade fitoplanctônica, pois ao cessar sua fonte de radiação não há renovação das comunidades. A decomposição da matéria orgânica oriunda da morte dos organismos algais gera a diminuição do OD, sem a recomposição gerada pela fotossíntese.

**Figura 14 - Box-plot com a distribuição dos dados de OD nos 30 pontos amostrais (entre março/2022 e fevereiro/2023) (a) e espacialização dos valores medianos realizada por meio de interpolação com Krigagem ordinária, período chuvoso (b) e seco (c)**



Fonte: Autor (2022)

As Figuras 14b e 14c demonstram os resultados da Krigagem para os dados de OD, para os dados subdivididos em período chuvoso e seco, respectivamente. Não houve variação significativa entre os períodos climáticos para o parâmetro, sendo que a mediana no período chuvoso foi superior. Neste período climático, a região central da Lagoa da Pampulha, entre os pontos 15 e 25, ainda possui intensa produção primária e, devido à decomposição microbiana a montante e à sedimentação, a matéria

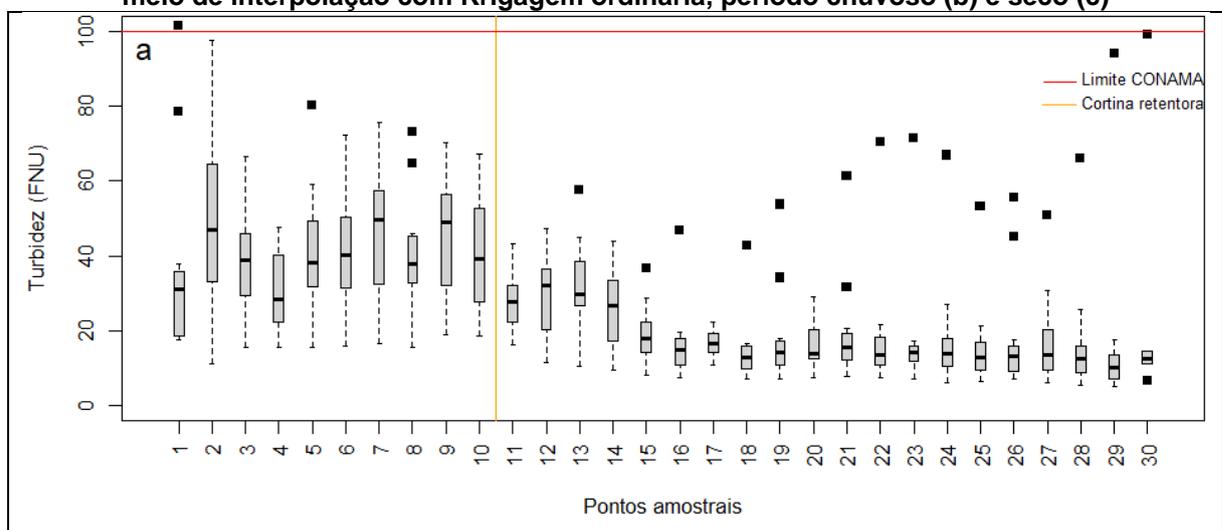
orgânica não induz a perda de O<sub>2</sub> nesta porção do reservatório. Sendo assim, as concentrações de OD nesta região, dita intermediária, foram maiores.

As regiões de entrada dos córregos Ressaca e Sarandi (ponto 1) e próxima à barragem (ponto 30), apresentaram os menores índices de OD, em razão da baixa qualidade da água e da baixa produtividade primária nas proximidades, respectivamente. No período seco a região próxima ao ponto 30, possivelmente influenciado pela decomposição e sedimentação da matéria orgânica ao longo do reservatório, o que diminui o sequestro de O<sub>2</sub> necessário para sua decomposição, os índices de OD foram superiores ao observado no restante da Lagoa da Pampulha. A região próxima aos pontos 16, 17, 18 e 19 apresentaram os menores índices de OD no período seco, influenciado pela menor produção primária na região (Figura 9h) e pela inserção de efluentes pelos córregos Tijuco e Mergulhão.

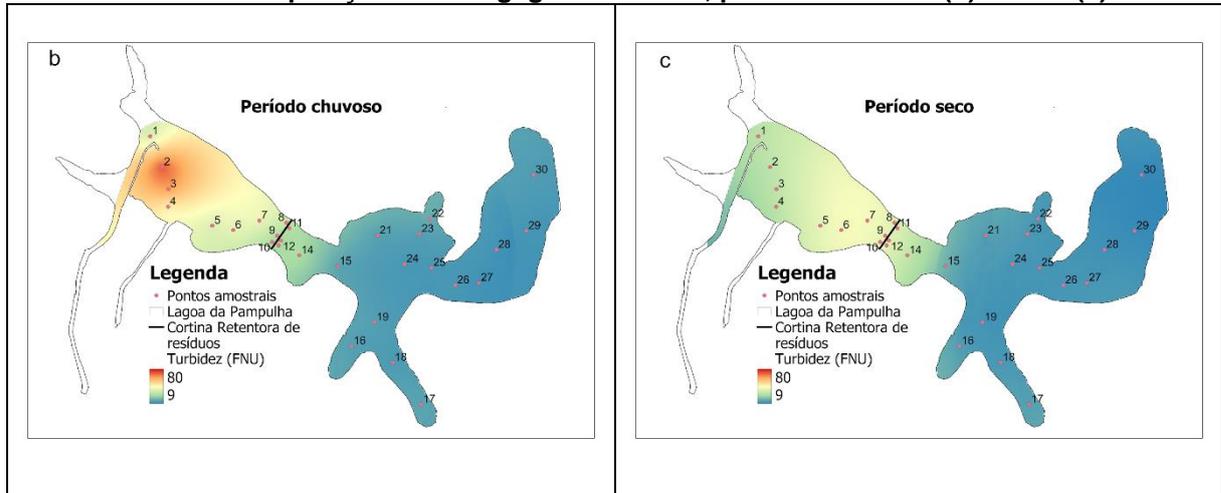
### 5.2.5 Turbidez

Na Figura 15a podemos observar que a turbidez foi maior nos pontos anteriores à cortina de retenção de resíduos (7, 9 e 10), em função da barreira imposta. Nota-se grande quantidade de *outliers* na região próxima ao barramento, provocados pela coleta no mês de outubro, em que as concentrações de chl-a e ficocianina foram superiores às demais amostragens, resultando em colônias de fitoplâncton, provavelmente cianobactérias, na superfície da água, que interferiram nos níveis de turbidez.

**Figura 15 - Box-plot com a distribuição dos dados de turbidez nos 30 pontos amostrais (entre março/2022 e fevereiro/2023) (a) e espacialização dos valores medianos realizada por meio de interpolação com Krigagem ordinária, período chuvoso (b) e seco (c)**



**Figura 15 - Box-plot com a distribuição dos dados de turbidez nos 30 pontos amostrais (entre março/2022 e fevereiro/2023) (a) e espacialização dos valores medianos realizada por meio de interpolação com Krigagem ordinária, período chuvoso (b) e seco (c)**



Fonte: Autor (2022)

Através das Figuras 15b e 15c se observa que a turbidez tem padrões similares nos dois períodos climáticos, não ocorrendo variação estatística na sazonalidade, não sendo observadas diferenças significativas entre as medianas dos períodos de chuva e estiagem. Em ambos os períodos a turbidez foi maior na região superior do reservatório em função da chegada de sedimentos pelos tributários (período chuvoso) e intensa produtividade primária (período seco).

À jusante do reservatório a turbidez foi menor devido à sedimentação dos sólidos suspensos na água que provocam a turbidez. Lopes et al. (2019), encontraram resultados semelhantes ao analisar as condições da Lagoa da Pampulha para prover atividades de contato primário utilizando dados do IGAM entre 2012 e 2017, apontando que existe a tendência de assentamento das partículas e decomposição microbiana, com a consequente diminuição da turbidez ao longo da lagoa.

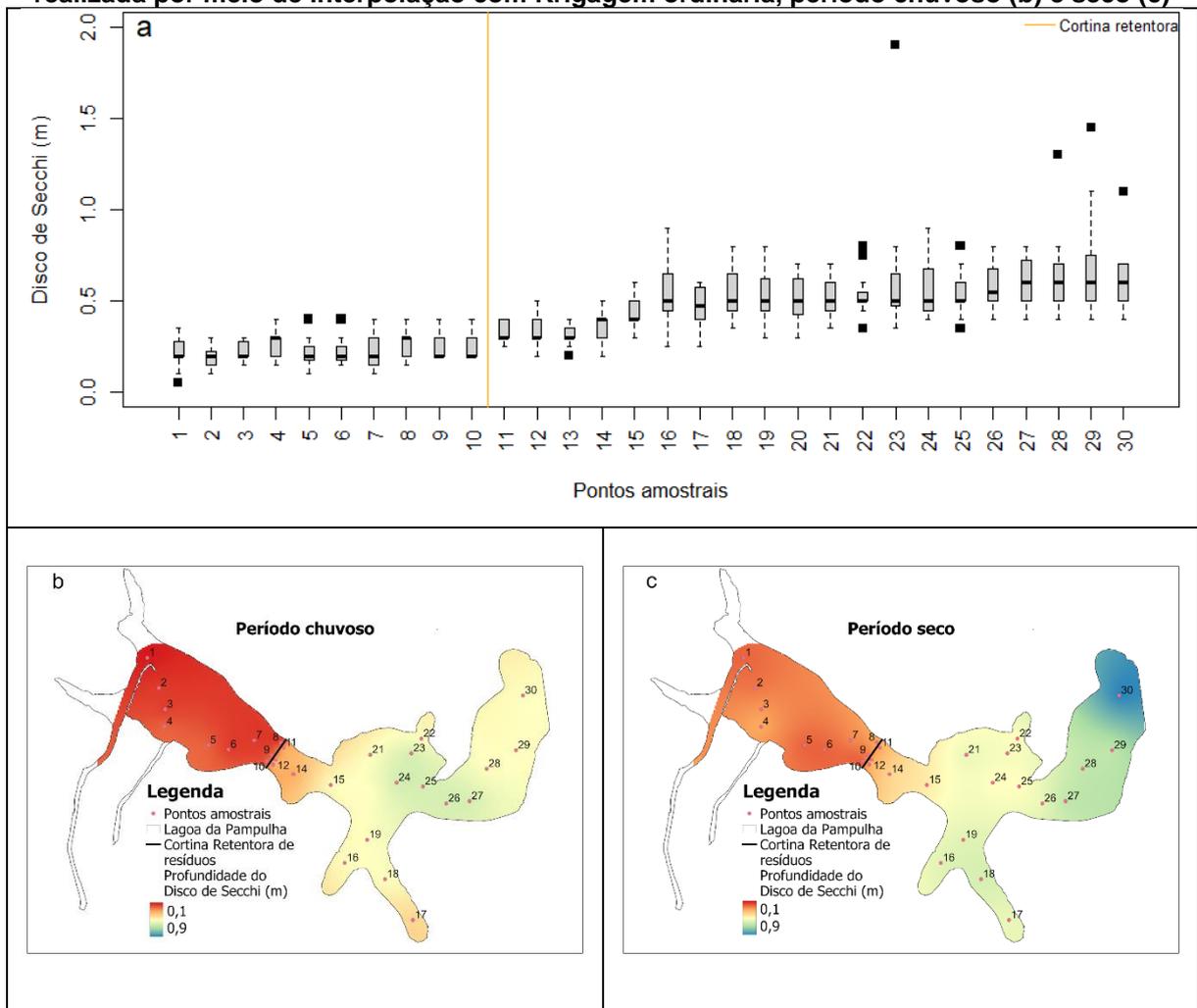
#### 5.2.6 Profundidade do disco de Secchi

O disco de Secchi é utilizado para medir a transparência da coluna d'água e mensurar a profundidade da zona fótica, que compreende a região da massa de água onde há até 1% da luz incidente. Esta região é fundamental para os processos físicos e biológicos em lagos e reservatórios, pois interfere nos padrões de circulação e estratificação da água e na fotossíntese (ESTEVES, 2011).

Como esperado, os valores medidos com o disco de Secchi (Figura 16a) apresentaram correlação inversa com a turbidez, Figura 15a, já que altos níveis de

turbidez impedem a passagem da luz no meio aquático, o que reduz a profundidade do disco de Secchi. A relação inversa também foi observada com os valores de chl-a (Figura 11a), isto ocorre em função de florações de algas e cianobactérias, comuns na Lagoa da Pampulha, que alteram os índices de iluminação para as camadas inferiores da coluna de água e interferem na passagem da luz na coluna d'água.

**Figura 16 - Box-plot com a distribuição dos dados de Disco de Secchi nos 30 pontos amostrais (entre março/2022 e fevereiro/2023) (a) e espacialização dos valores medianos realizada por meio de interpolação com Krigagem ordinária, período chuvoso (b) e seco (c)**



Fonte: Autor (2022)

Em ambos os períodos climáticos há claro aumento da profundidade do disco ao longo do reservatório (Figuras 16b e 16c), que é maior na região da barragem e não foram observadas variações significativas em relação à mediana. Assim como a turbidez estes resultados ocorrem pela maior concentração de partículas suspensas que interferem na passagem da luz a montante do reservatório e que se sedimentam a jusante.

Souza (2019) encontrou profundidades maiores para o disco de Secchi (0,55 m), indicando melhor qualidade da água no período analisado, abril de 2018. No presente estudo, a profundidade mediana foi de 0,30 m, para o mês de abril de 2022. A autora analisou os efeitos da cortina retentora de resíduos nos dados para todos os parâmetros. Concluiu-se, assim como no presente estudo, que a cortina divide a Lagoa da Pampulha em dois ambientes distintos, um com pior qualidade da água (antes da cortina) e outro com carga pouco menor poluição (após a cortina).

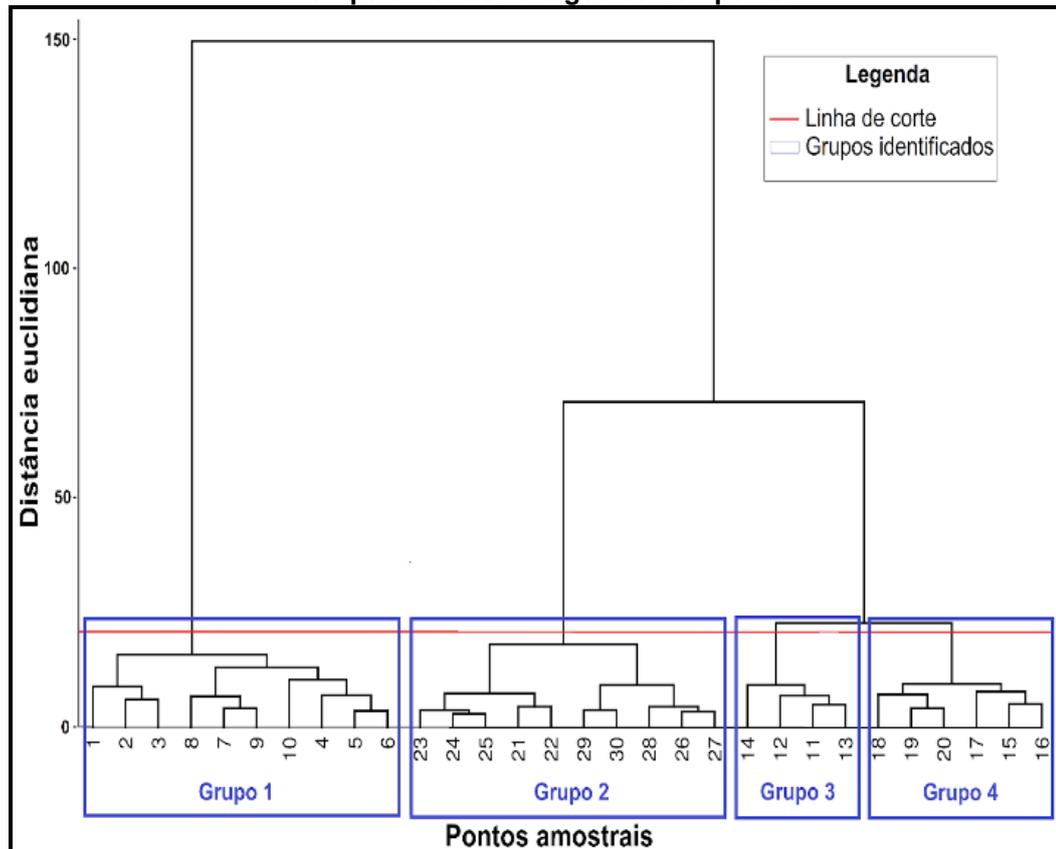
#### 5.2.7 Análise de cluster

A Figura 17 mostra o dendrograma para visualização dos agrupamentos obtidos por meio da análise de cluster, para os grupos de pontos em relação a todos os dados dos parâmetros monitorados com a sonda multiparâmetros (clorofila-a, ficocianina, pH, OD, turbidez e profundidade do disco de Secchi) na lagoa da Pampulha entre março de 2022 e fevereiro de 2023.

O método *gap statistic*, determinou como 4 o número ideal de agrupamentos (Apêndice E). O cálculo do coeficiente de correlação cofenética apresentou resultado de 0,71, demonstrando boa correspondência do método de agrupamento utilizado, já que valores acima de 0,7 representam baixas distorções e bom grau de ajuste nos agrupamentos obtidos (COSTA, 2021).

Percebe-se, na Figura 17, grande diferença entre o Grupo 1, pontos 1, 2, 3, 4, 5, 6, 7, 8, 9 e 10, todos situados antes da cortina de resíduos (região a montante e próxima aos principais tributários), e os demais. Estes resultados corroboram as análises realizadas no presente estudo, nos quais foram constatadas diferenças significativas para todos os parâmetros nos dados antes e depois desta barreira. O Grupo 1 apresentou, como demonstrado na análise dos dados primários e secundários, os piores índices de qualidade da água, influenciados pela entrada das águas dos córregos Ressaca e Sarandi.

**Figura 17 - Dendrograma da análise de agrupamento dos parâmetros coletados com a sonda multiparâmetros na lagoa da Pampulha**



Fonte: do Autor (2023).

O Grupo 2, que representa os pontos situados próximos à barragem, é caracterizado pelos pontos: 21, 22, 23, 24, 25, 26, 27, 28, 29 e 30. Este grupo apresentou os melhores indicadores de qualidade da água, favorecido pela menor quantidade de sólidos sedimentáveis e nutrientes verificados nas seções 5.1, 5.2.1 e 5.2.5. O Grupo 4 é representado pelos pontos 15, 16, 17, 18, 19 e 20, correspondendo ao compartimento intermediário da lagoa ou zona de transição.

Estes 3 grupos denotam semelhança espacial com os pontos monitorados pelo IGAM no reservatório. O ponto PV230 está inserido na região do Grupo 1, o ponto PV235 se localiza na mesma região do Grupo 4, compartimento intermediário da Lagoa da Pampulha e o ponto PV240 está localizado nas proximidades do barramento, mesma região em que se encontram os pontos do Grupo 2. Estes resultados evidenciam que o monitoramento executado pelo órgão representa bem essas porções do reservatório. Entretanto, a presença do Grupo 3 nos resultados de agrupamento, representado pelos pontos 11, 12, 13 e 14 e que estão situados logo após à cortina retentora de sedimentos, exprime a possível necessidade de se adicionar mais um

ponto de monitoramento que abranja a região ao qual estão inseridos estes pontos. A cortina foi alocada em 2016, após a definição dos 3 pontos amostrais monitorados pelo IGAM. De fato, ao analisar a espacialização dos parâmetros na Lagoa da Pampulha, através das Figuras 11b e 11c (chl-a), 12b e 12c (ficocianina), 15b e 15c (turbidez) e 16b e 16c (profundidade do disco de Secchi), percebe-se que na região logo após a cortina de retenção de resíduos ocorrem alterações abruptas nas concentrações dos parâmetros analisados, que se tornam gradativas no restante do reservatório.

Portanto, com a implementação da cortina houve diminuição da zona de influência do compartimento fluvial do reservatório, que detém os piores índices de qualidade de água. O compartimento de transição se subdividiu em duas porções, uma próxima à cortina com a qualidade da água influenciada pelas águas anteriores à mesma e outra a jusante (localizada próxima ao ponto PV235), onde observou-se melhores índices nas concentrações dos parâmetros analisados.

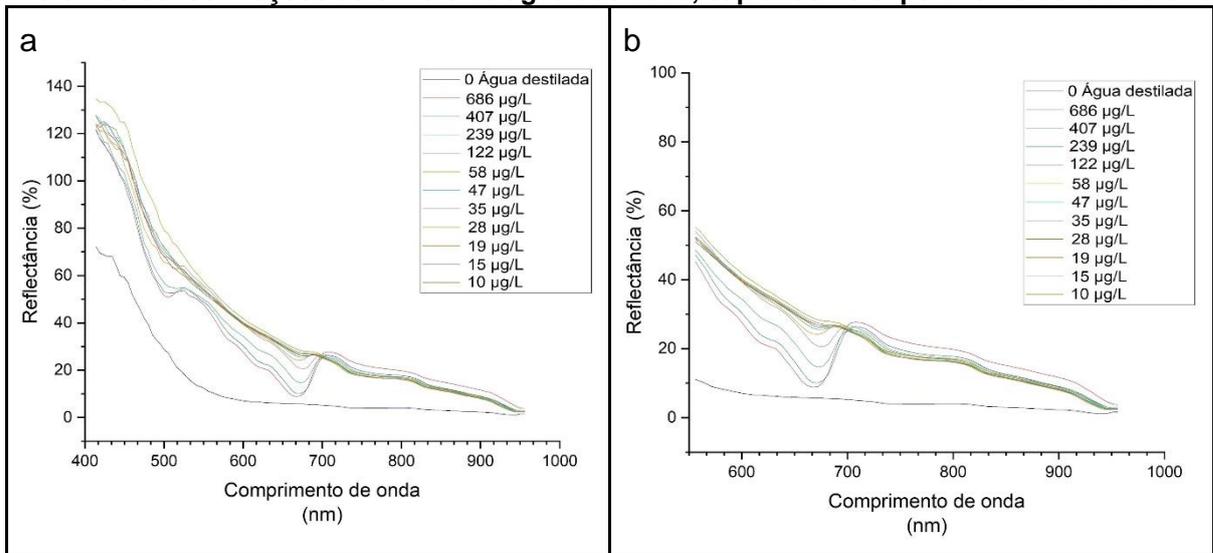
Desta forma, recomenda-se a realização de estudos com uma ampla variedade de parâmetros para analisar a necessidade de alocação de mais um ponto amostral na região em que se encontra o Grupo 3. Necessidade devida à suma importância para os gestores deste recurso hídrico em compreender toda a dinâmica espacial e efeitos da poluição incidente que ocorrem na qualidade da água em todos os compartimentos encontrados na Lagoa da Pampulha.

### **5.3 Desenvolvimento da metodologia de calibração e utilização do sensor hiperespectral em laboratório**

#### **5.3.1 Reflectância hiperespectral das imagens de laboratório**

A Figura 18 mostra as curvas de reflectância dos espectros das diferentes concentrações de chl-a em razão do comprimento de onda, mensuradas a partir do cultivo da alga *Raphidocelis subcapitata*, bem como de uma amostra de água destilada. Os dados foram extraídos das imagens capturadas pelo sensor hiperespectral OCI™-F e processadas pelo software *CubeCreator*. Foram obtidas 109 bandas entre 413 e 955 nm, com resolução espectral de aproximadamente 5 nm.

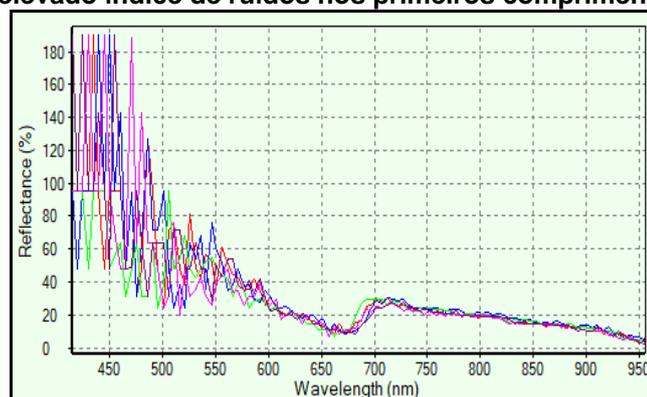
**Figura 18 - Espectros de reflectância obtidos através das imagens hiperespectrais das diferentes concentrações de chl-a e da água destilada, captadas no experimento de laboratório**



Fonte: Autor (2023)

A partir da Figura 18a, observa-se uma superestimação dos valores de reflectância nos comprimentos de onda entre 400 nm e 550 nm (azul e parte do verde). Essa superestimação resulta em valores de reflectância que ultrapassam 100% e ocorre devido à alta incidência de ruído (Figura 19). Como a metodologia utilizada propõe a aplicação de uma área de interesse, o espectro gerado desta área é a média dos pixels presentes na mesma. Conseqüentemente, os altos níveis de ruído afetam as características espectrais de reflectância da água e das amostras de algas nessa região, introduzindo incertezas nos resultados obtidos com o espectro de reflectância. Essas perturbações podem ser causadas por diversos fatores, como imperfeições do sensor, interferência atmosférica, reflexão e espalhamento de luz, efeitos de iluminação variável e erros no processo de aquisição e transmissão dos dados (RASTI et al., 2018; WANG et al., 2018)

**Figura 19 - Exemplos de 4 pixels extraídos de uma imagem do experimento de laboratório, expondo o elevado índice de ruídos nos primeiros comprimentos de ondas**

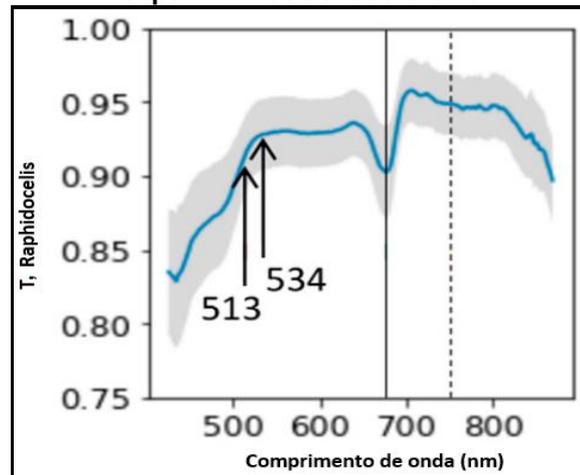


Fonte: Autor (2023)

Resultados semelhantes relacionados a ruídos foram observados por Ewing et al. (2020) ao utilizarem o sensor OCI-FTM em experimentos de detecção da gradação de solos em laboratório. Os autores utilizaram uma lâmpada de tipo similar (500 W) e constataram que as bandas entre 400 e 449 eram fortemente influenciadas pelos ruídos. Essas bandas foram particularmente afetadas devido à menor emissão de energia luminosa pelas lâmpadas halógenas nessa faixa do espectro (aproximadamente entre 350 e 475 nm), conforme observado nos experimentos de Kim et al. (2018). Além disso, variações na modulação da fonte de eletricidade e nas propriedades do material do filamento também podem ocorrer. Para investigar essa questão, foram realizadas medições da irradiância para a lâmpada de halogênio, conforme apresentado na seção 4.4.

As regiões afetadas pelos sinais de ruído são importantes para a detecção de chl-a (SAGAN et al., 2020), Salmi, Eskelinen e Leppanen (2021) conduziram estudos laboratoriais para a detecção de microalgas utilizando uma câmera hiperespectral. Na Figura 20, é possível observar o espectro gerado, indicando os comprimentos de onda que apresentaram maior correlação com análises de fluorescência. Nessa imagem, também é perceptível que a região entre 400 nm e 550 nm mostrou diferenças em relação ao presente estudo, no qual foram observadas interferências de ruído. No entanto, é importante ressaltar que o sensor e as metodologias de coleta foram diferentes. A emissão da lâmpada na região de alta incidência de ruídos nos espectros de reflectância foi reduzida (Figura 7b), comportamento semelhante aos experimentos realizados por Kim et al. (2018), que relacionaram a menor irradiância a oscilações no sinal, as quais podem gerar ruídos observados.

**Figura 20 - Espectro de transmitância hiperespectral (T) da alga *Raphidocelis*, captada em experimentos de laboratório**



Fonte: Salmi, Eskelinen, Leppanen (2021)

No caso da água destilada, observa-se um aumento da absorção à medida que o comprimento de onda aumenta. Comprimentos de onda acima de 700 nm absorvem praticamente toda a luz que incide na superfície da água (ESTEVES, 2011). Para a chl-a, podemos observar a presença de vales na região entre 672 nm e 690 nm, resultantes de uma maior absorção da luz pela clorofila-a nessa faixa de comprimento de onda (YANG et al., 2022), e picos após 700 nm, devido a uma maior reflectância da luz nessa faixa espectral (ZHOU et al., 2019). Essas características específicas do pigmento são amplamente descritas na literatura.

Zhou et al. (2019) estudaram a resposta de reflectância hiperespectral de três tipos de algas em laboratório, *Anabaena cylindrica*, *Nannochloropsis gaditana* e PW-95. As regiões de maior pico foram encontradas entre 700 nm e 730 nm para *A. cylindrica* e *N. gaditana*, que também apresentaram vales semelhantes entre 660 nm e 700 nm. No caso da PW-95, os vales foram observados entre 660 nm e 708 nm, e os picos entre 708 nm e 735 nm. Salmi, Eskelinen e Leppanen (2021) também constataram a presença desses picos e vales em um cultivo de *T. Raphidocelis*, nas regiões em torno de 700 nm (picos) e 667 nm (vales), conforme mostrado na Figura 20.

### 5.3.2 Estimativa da concentração de algas a partir da reflectância hiperespectral em laboratório

Para a construção do modelo, foram selecionadas as bandas mais significativas utilizando o método Forward de seleção automática. Para análises do modelo, foram excluídas as faixas espectrais em que os pixels apresentaram altos índices de ruído

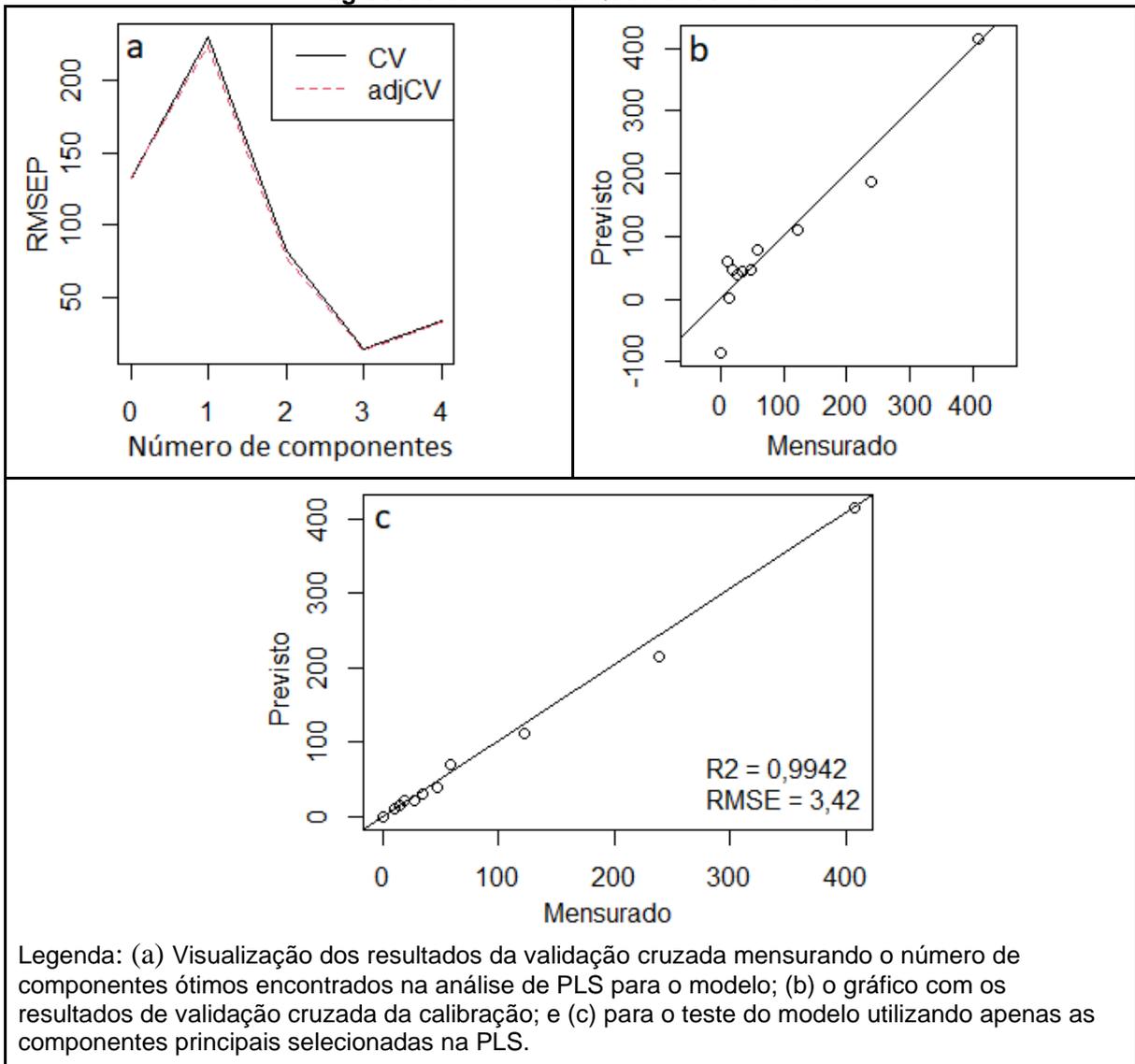
(413 a 550 nm). As bandas nos comprimentos de onda de 672 nm, 792 nm, 692 nm e 712 nm foram identificadas como as principais preditoras nesse contexto, mostrando similaridade com os resultados observados na Tabela 4, que lista as principais bandas encontradas em alguns estudos.

Os resultados da regressão linear múltipla demonstraram uma alta correlação dos dados, com  $R^2 = 0,999$  e  $RMSE = 7,076 \mu\text{g/L}$ . No entanto, a análise de resíduos revelou a presença de multicolinearidade entre as variáveis, um fenômeno comum em dados hiperespectrais, devido ao elevado número de bandas, que pode levar a problemas de sobreajuste e superestimação do modelo (ZHAO et al., 2021).

Para contornar esse problema, adotou-se a regressão de mínimos quadrados parciais (PLS, do inglês Partial Least Squares). Foram utilizadas apenas as bandas preditoras selecionadas pelo método Forward, pois apresentaram um RMSE melhor em comparação ao PLS com todas as bandas (556 nm a 955 nm). Inicialmente, determinou-se a quantidade de componentes que melhor explicavam o ajuste do modelo por meio da validação cruzada k-fold, o que resultou na utilização de 3 componentes para otimizar o modelo.

Os resultados estão ilustrados na Figura 21a, na qual podemos observar que, após a terceira componente (RMSE de  $10,8 \mu\text{g/L}$ ), o valor do RMSE aumenta, indicando que a utilização de 3 componentes PLS é ideal para o modelo final. A Figura 21b apresenta os resultados da validação cruzada para os dados de calibração utilizando todas as componentes, evidenciando uma boa concordância entre os valores medidos e previstos, com um  $R^2$  de 0,9981.

**Figura 21 – Resultados do modelo aplicado ao experimento de laboratório utilizando a regressão de Mínimos Quadrados Parciais**



Fonte: Autor (2023)

Os resultados da PLS demonstraram uma alta porcentagem de variância explicada pelos 3 componentes principais selecionados para aplicação do modelo de teste (Figura 21c). Além disso, foi observada uma redução no RMSE, que alcançou um índice de 3,42  $\mu\text{g/L}$ , enquanto o  $R^2$  foi de 0,9942. Esses resultados são similares aos estudos encontrados na literatura. Por exemplo, Zhou et al. (2019) analisaram a concentração de 4 tipos diferentes de algas em laboratório usando imagens hiperespectrais e obtiveram resultados com  $R^2$  acima de 0,99 para todas as algas estudadas. Salmi et al. (2021) encontraram coeficientes de determinação entre 0,85 e 0,96 ao correlacionar proporções de bandas de imagens hiperespectrais com concentrações de clorofila-a para monitorar o crescimento de 6 tipos diferentes de

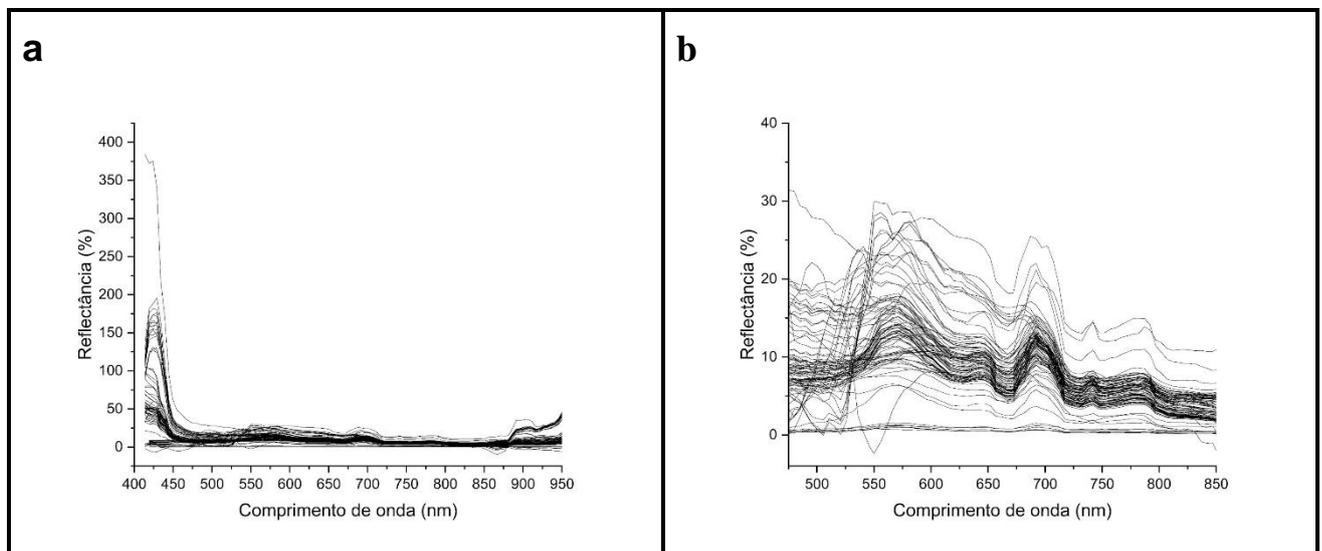
algas em laboratório. Osanen (2022) também obteve resultados semelhantes ( $R^2$  acima de 0,97) ao analisar amostras de algas de gelo marinho em laboratório. Portanto, conclui-se que o sensor OCI-FTM tem a capacidade de reproduzir resultados observados em outras metodologias que utilizam sensores hiperespectrais para estimar a concentração de biomassa de algas através da chl-a.

#### 5.4 Estimativa da concentração de clorofila-a através do sensor hiperespectral em campo através do desenvolvimento de modelos empíricos

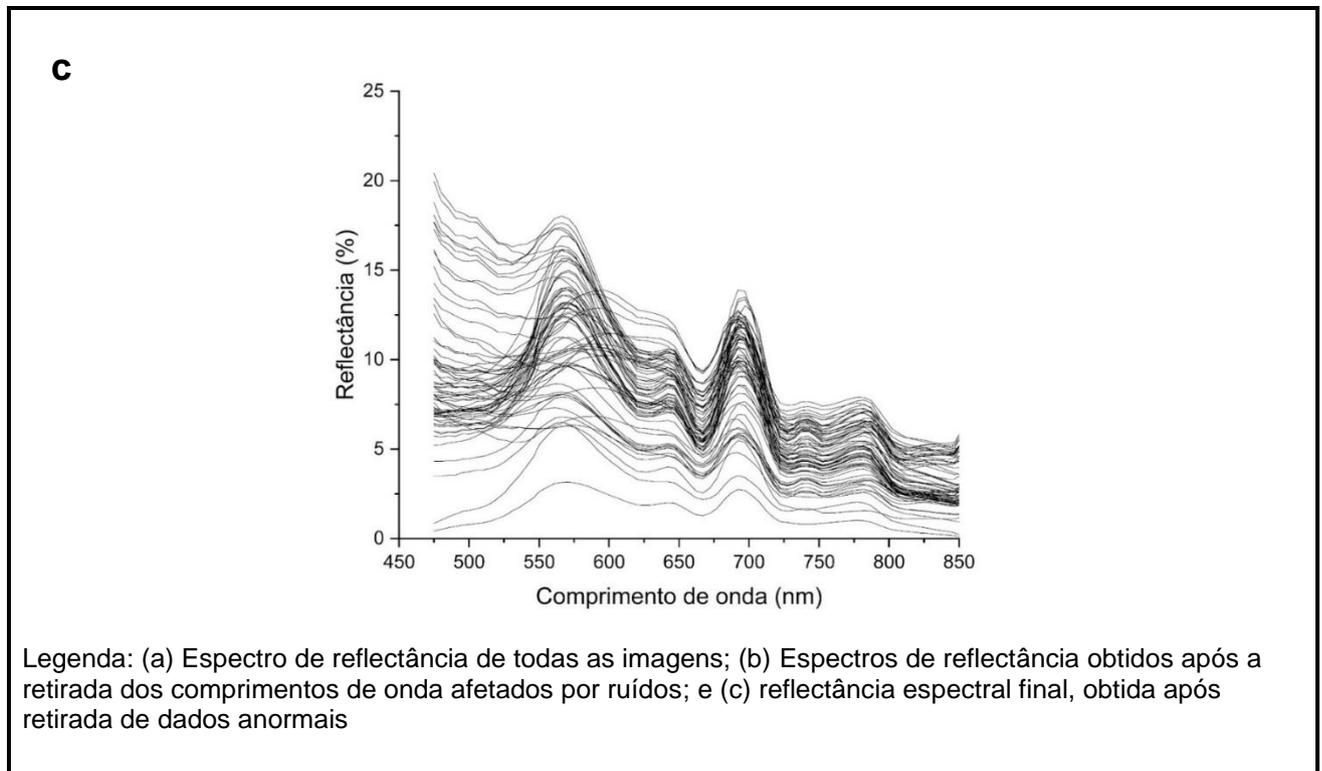
##### 5.4.1 Reflectância hiperespectral das imagens coletadas em campo

A Figura 22 apresenta os espectros de reflectância das imagens hiperespectrais coletadas em campo. Na Figura 22a, podem ser observadas distorções nos comprimentos de onda iniciais e finais, causadas pelos ruídos presentes nas imagens, semelhantes aos observados nas imagens de laboratório (Figura 19, seção 5.3.1). Keller et al. (2018) mencionam que, em geral, imagens hiperespectrais são suscetíveis a ruídos nestes comprimentos de onda. Nos dados do presente estudo, foi observado que os comprimentos de onda até 470 nm e acima de 850 nm foram mais afetados pela presença de ruídos. Na Figura 23c (seção 5.4.2), é apresentado o espectro contínuo de 1280 bandas com 1 nm de comprimento de onda, coletadas pelo OCI-FTM a partir de uma amostra do banco de dados, evidenciando a maior magnitude dos ruídos nas faixas extremas do espectro.

**Figura 22 - Espectros de reflectância hiperespectral obtidos através das imagens coletadas em campo na lagoa da Pampulha**



**Figura 22 - Espectros de reflectância hiperespectral obtidos através das imagens coletadas em campo na lagoa da Pampulha**



Fonte: Autor (2023)

Devido à presença desses ruídos, foram excluídos da análise os comprimentos de onda até 470 nm e acima de 850 nm, a fim de evitar incertezas nos resultados dos modelos. A Figura 22b apresenta os espectros resultantes após a remoção dos comprimentos de onda extremos, onde podem ser observados espectros anômalos que não representam características de reflectância típicas da superfície da água. Esses espectros anômalos são causados pela reflexão especular da luz solar, resultando em uma maior magnitude do espectro (CHONG et al., 2021), e por imagens coletadas de forma inadequada, sem os ajustes adequados de tempo de exposição (espectros abaixo do normal). Após a exclusão dos dados mencionados, obtemos os espectros de reflectância que foram utilizados como entrada nos modelos propostos (Figura 22c).

A forma dos espectros resultantes das imagens hiperespectrais coletadas na Lagoa da Pampulha se assemelha a estudos similares encontrados na literatura que analisaram corpos d'água eutróficos com o uso de sensores hiperespectrais em campo (WANG et al., 2017; CHENG et al., 2019; YANG et al., 2022; CAI et al., 2022).

Características típicas observadas incluem o vale de absorção entre os comprimentos de onda de 660 nm e 680 nm, que é característico da chl-a (GHOLIZADEH; MELESSE; REDDI, 2016; SAGAN et al., 2020). Além disso, foram identificados picos de reflexão nas faixas espectrais entre 550 nm e 580 nm e próximo a 700 nm, que são relatados como intervalos espectrais importantes para o cálculo das concentrações de chl-a em águas do tipo 2 (YANG et al., 2022). Cui et al. (2021) encontraram picos de reflexão na faixa entre 550 nm e 580 nm ao analisarem a concentração de chl-a por meio de um sensor hiperespectral em um lago na China. Os autores explicam que esse pico é causado pela baixa absorção da chl-a e do caroteno, além dos efeitos de dispersão da luz provocados pelas células dos organismos. O pico encontrado na faixa de 700 nm é comumente utilizado na construção de índices entre bandas para a estimativa da concentração de chl-a (CHENG et al., 2019; GHOLIZADEH; MELESSE; REDDI, 2016).

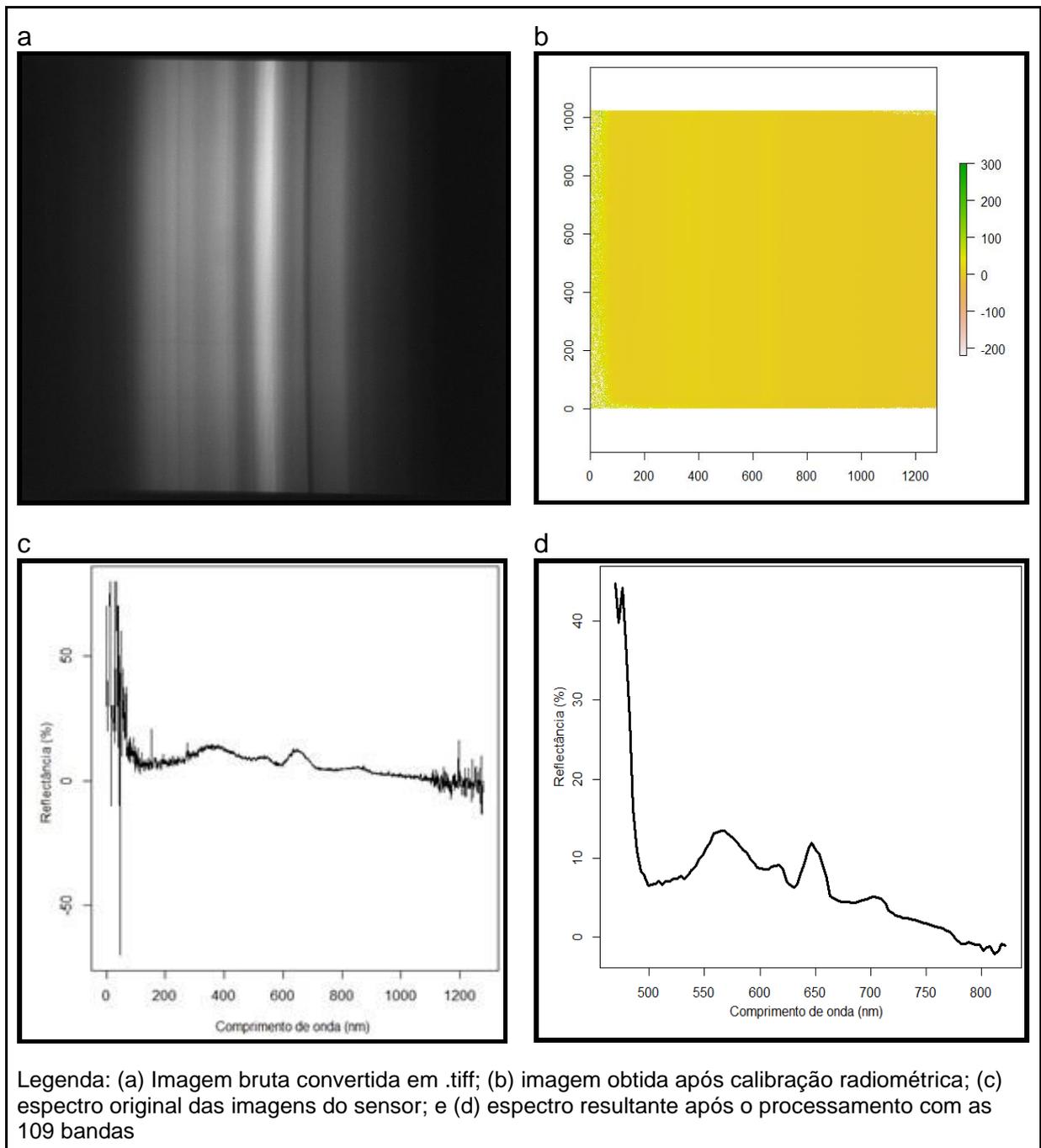
#### 5.4.2 Estimativa da concentração de clorofila-a a partir das imagens de campo do sensor hiperespectral

Foram realizadas coletas de campo mensais, resultando em 107 pares de dados. No entanto, na coleta de outubro, algumas amostras de imagens foram perdidas devido a um erro na aquisição. Para compensar as informações não coletadas em dezembro, devido a eventos climáticos adversos, foram extraídas imagens de todos os pontos amostrados pela sonda multiparâmetros em fevereiro. Após a exclusão das imagens mencionadas na seção 5.4.1, restaram 82 pares de dados que foram inseridos no algoritmo para estimar a concentração de chl-a. As campanhas de novembro e janeiro foram as principais responsáveis pelos dados de reflectância que apresentaram índices mais altos, possivelmente devido às condições de tempo nublado durante as coletas. Devido a essas condições climáticas, foi necessário aumentar o ganho de captura das imagens nesses dois meses, o que, combinado com alterações no tempo de exposição, pode resultar em maiores proporções de ruído, conforme descrito por Gowen et al. (2015).

Como resultado das extrações das imagens, os métodos utilizados no presente estudo, assim como o *CubeCreator*, criam um arquivo com a curva espectral das 109 bandas captadas pelo sensor para cada imagem. A sequência de trabalho empregada foi a seguinte: (i) Transformação da extensão da imagem *raw* para *.tiff* no ImageJ

(Figura 23a); (ii) Inserção das imagens (imagem referente ao ponto amostral, de referência branca e escura) e calibração radiométrica com a obtenção da imagem com os valores de reflectância, através da Equação 5 (Figura 23b); (iii) visualização das curvas hiperespectrais adquiridas com resolução espectral de 1280 bandas de acordo com o arquivo *raw* (Figura 23c) e arquivamento da curva com dados das 109 bandas, empregando a redução de dimensionalidade descrita na seção 4.1.1, utilizando os mesmos parâmetros do arquivo de calibração usado pelo *CubeCreator*, Figura 23d.

**Figura 23 - Resultados para as etapas das rotinas computacionais do processamento das imagens hiperespectrais.**



Em relação às concentrações de chl-a, as seções anteriores demonstraram a existência de uma grande variabilidade espacial e temporal na Lagoa da Pampulha, o que é um fator importante para a calibração e treinamento dos algoritmos, resultando em uma melhor eficiência do modelo (CHENG et al., 2019). A Tabela 8 apresenta as estatísticas descritivas das concentrações de chl-a no reservatório nos meses e pontos considerados para o desenvolvimento dos modelos de previsão. Observa-se que os valores máximos estão abaixo dos valores históricos observados na avaliação dos dados secundários do IGAM e dos dados primários obtidos com a sonda, que foram de 748,09 µg/L e 112,96 µg/L, respectivamente. No entanto, esses valores ainda estão dentro do limite estabelecido pela Resolução CONAMA 357/05, que é de 60 µg/L, e estão próximos das medianas observadas nos dados históricos para a Lagoa da Pampulha (tanto os primários quanto os secundários), com exceção do Ponto PV230 durante o período seco, onde a mediana foi superior a 200 µg/L (Figura 9h) nos anos de 2013 a 2021.

**Tabela 8 - Estatística descritiva para os dados de chl-a em µg/L coletados simultaneamente com as imagens hiperespectrais para construção dos modelos**

Chl-a (µg/L)	n	Mediana	Mínimo	Máximo	Desvio Padrão	Variância
Mês da coleta						
Setembro de 2022	15	20,02	10,87	62,35	19,56	382,65
Outubro de 2022	8	55,79	45,97	68,13	6,07	36,89
Novembro de 2022	16	31,81	15,42	54,37	10,51	110,56
Janeiro de 2023	13	37,8	4,98	46,23	15,15	229,79
Fevereiro de 2023	30	34,92	20,53	61,69	14,54	211,44

Fonte: Autor (2023)

Ao todo, foram 16 pontos em que as imagens hiperespectrais foram coletadas. A campanha de dezembro não foi realizada devido a condições climáticas e a campanha de outubro apresentou menores amostras por falha no momento de aquisição das imagens. Com isto, na campanha de fevereiro, optou-se pela aquisição das imagens em todos os pontos monitorados pela sonda multiparâmetros para suprir as perdas dos meses citados.

Foram realizados 4 Testes: (A) apenas com os meses de setembro, outubro e fevereiro; (B) apenas com o mês de novembro e de janeiro; (C) com os dados de todos os meses; (D) com os dados de todos os meses, aplicado após a seleção das

melhores bandas preditivas (método de seleção: *forward inclusion stepwise*). O intuito foi o de verificar os efeitos do maior índice de ruídos observados nos meses de novembro e janeiro e suas influências nas incertezas associadas à captura das imagens.

O Teste C apresentou os resultados mais satisfatórios, conforme apresentado na Tabela 9, indicando que não são necessárias manipulações dos dados para se obter bons resultados utilizando o sensor hiperespectral OCI-F™ e os modelos de aprendizagem de máquina propostos. Os resultados citando os resultados para cada teste realizado estão dispostos no Apêndice F. É importante ressaltar que os melhores resultados encontrados para R<sup>2</sup> e RMSE foram para o teste que compreendeu apenas os meses de setembro, outubro e fevereiro. No entanto, esse teste não foi considerado devido à possibilidade de overfitting, ou seja, um ajuste excessivo do modelo devido à menor quantidade de dados disponíveis (SAGAN et al., 2020), o que poderia resultar em baixo desempenho ou previsões imprecisas para dados de outros meses.

**Tabela 9 - Resultados dos modelos para o teste C, que utilizou todos os meses, demonstrando os valores de R<sup>2</sup> para o treino, teste e validação e RMSE de validação para cada tipo de entrada dos dados utilizada, com destaque para o melhor resultado observado**

Modelo	Tipo utilizado	R <sup>2</sup> Treino	R <sup>2</sup> Teste	R <sup>2</sup> Validação	RMSE Validação (µg/L)
Regressão Linear Multivariada (RLM)	Linear	0,16	0	0,19	16,29
	Polinomial	0,17	-0,22	0,19	16,35
	Exponencial	0,08	-0,18	0,17	16,49
	Padronizado	0,10	-0,05	0,18	16,42
	Min/Máx	0,16	0,01	0,19	16,29
Random Forest Regressor (RFR)	Linear	0,91	0,29	0,62	11,19
	Polinomial	0,90	0,27	0,65	10,65
	Exponencial	0,89	0,29	0,68	10,29
	Padronizado	0,91	0,33	0,66	10,57
	Min/Máx	0,91	0,32	0,66	10,64
Support Vector Regressor (SVR)	Linear	0,82	0,33	0,78	8,53
	Polinomial	0,78	0,34	0,73	9,33
	Exponencial	-0,57	-0,75	-0,41	21,55
	Padronizado	<b>0,76</b>	<b>0,22</b>	<b>0,86</b>	<b>6,89</b>
	Min/Máx	0,77	0,56	0,82	7,65

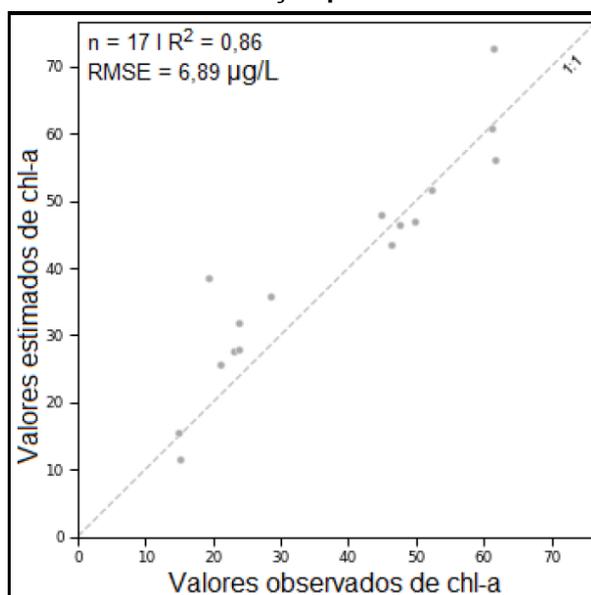
**Tabela 9 - Resultados dos modelos para o teste C, que utilizou todos os meses, demonstrando os valores de  $R^2$  para o treino, teste e validação e RMSE de validação para cada tipo de entrada dos dados utilizada, com destaque para o melhor resultado observado**

XGBooster Regressor (XGB)	Linear	1	0,28	0,62	11,13
	Polinomial	1	0,23	0,61	11,32
	Exponencial	1	0,26	0,57	11,94
	Padronizado	1	0,26	0,64	10,82
	Min/Máx	1	0,29	0,62	11,13

Fonte: Autor (2023)

O modelo SVR apresentou as melhores métricas de validação para estimativa da concentração da chl-a,  $R^2 = 0,86$  e  $RMSE = 6,89 \mu\text{g/L}$  (Tabela 9) no Teste C. A linha de ajuste que deteve melhor otimização em relação aos dados foi do tipo padronizado, método bastante utilizado para ajustes nas escalas de grandeza nas variáveis utilizadas. A Figura 24 exibe o diagrama de ajuste dos valores observados e as estimativas de concentração de chl-a, utilizando 16 pares de dados para validação. Observou-se que duas amostras apresentaram maior distorção, uma com valores observados de aproximadamente  $20 \mu\text{g/L}$  e outra com  $60 \mu\text{g/L}$ , nas quais o algoritmo superestimou as estimativas. No entanto, o gráfico da Figura 24 mostra uma boa dispersão dos dados em relação à linha de tendência 1:1 na maioria das comparações entre os valores medidos e observados, confirmando os resultados do coeficiente de determinação e indicando um bom ajuste do modelo aos dados.

**Figura 23 - Gráfico de dispersão entre os valores mensurados e estimados e os resultados das métricas de validação para o modelo SVR**



Fonte: Autor (2023)

Assim como no presente estudo, Peterson et al. (2019) obtiveram resultados semelhantes ao analisar chl-a por meio de um espectroscópio posicionado nas margens de vários lagos nos Estados Unidos. Foram analisadas um total de 105 amostras de pares de dados utilizando diversos modelos. Em contraste com o presente estudo, eles empregaram técnicas de redução de dimensionalidade, e o modelo Support Vector Regressor (SVR) demonstrou o melhor desempenho, com um  $R^2$  de 0,88 e RMSE de 5,87  $\mu\text{g/L}$ . Sagan et al. (2020) também utilizaram o SVR, entre outros modelos testados, para prever parâmetros de qualidade da água em um extenso banco de dados. Os resultados para o SVR foram um  $R^2$  de 0,78 e RMSE de 6,07  $\mu\text{g/L}$ . Em um estudo onde um sensor hiperespectral foi posicionado próximo a um experimento de mesocosmos com águas do Rio Mississippi nos Estados Unidos, os autores encontraram melhores resultados com modelos baseados em redes neurais.

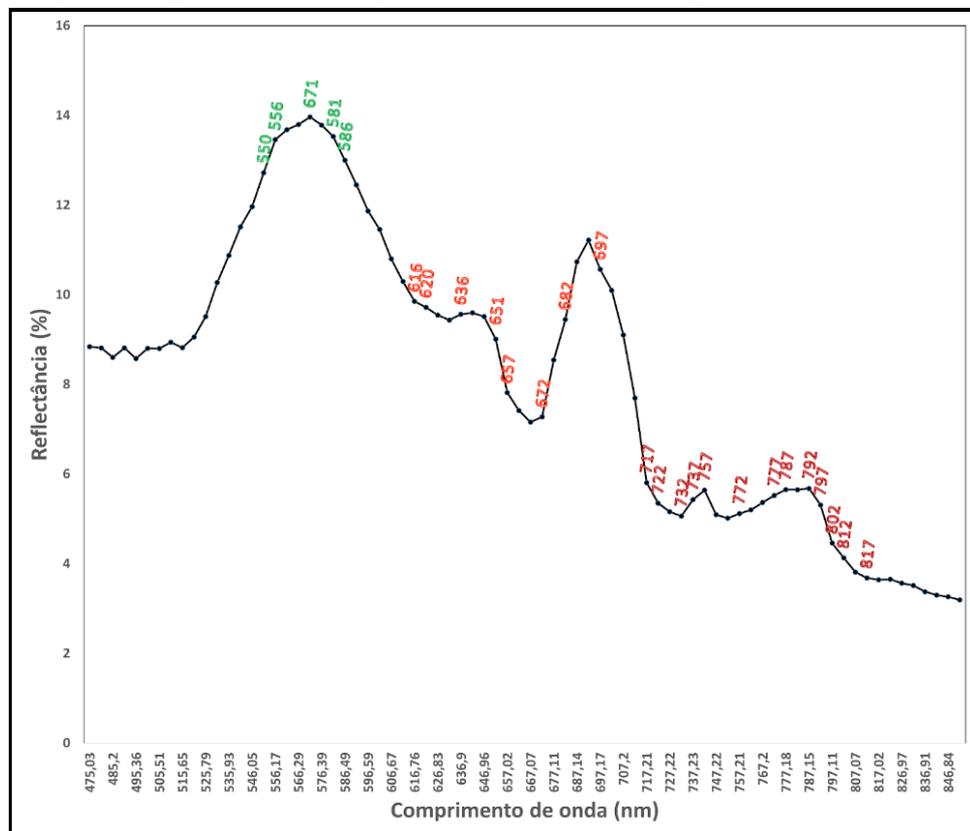
Assim como os estudos mencionados anteriormente, outros modelos também têm obtido bons resultados ao analisar chl-a em corpos d'água em ambientes eutróficos. Cui et al. (2021) alcançaram um  $R^2$  de 0,88 usando o XGBooster em 38 pontos amostrais de um rio altamente poluído na China, enquanto Keller et al. (2018) obtiveram  $R^2$  de 0,88 e 0,91 usando o modelo Support Vector Machine antes e após o tratamento dos dados com PCA, respectivamente. Os pares de dados utilizados foram coletados do Rio Elba, na Alemanha, totalizando 111 amostras. Wang et al. (2017) encontraram um  $R^2$  de 0,98 e RMSE de 6,15  $\mu\text{g/L}$  ao estimar chl-a usando a regressão de Partial Least Square (PLS) juntamente com a primeira derivada da reflectância em um lago com regiões eutróficas e não eutróficas no Japão. Vários estudos também têm utilizado índices de razão espectral para estimar chl-a em ambientes aquáticos, como mostrado por Cheng et al. (2019), Menon e Adhikari (2018), e Lu et al. (2020), que alcançaram  $R^2$  de 0,75, 0,85 e 0,94, respectivamente.

A Figura 25 exibe as 26 bandas selecionadas pelo método forward no teste D realizado neste estudo. Embora não tenha sido a melhor combinação entre as analisadas, esse teste apresentou boas métricas de validação ( $R^2 = 0,81$  e RMSE = 6,9  $\mu\text{g/L}$ , conforme mostrado no Apêndice F.3). As bandas estão dispostas no espectro de reflectância médio das imagens hiperespectrais utilizadas como entrada no modelo (Figura 22c). Essas bandas podem ser utilizadas para construir índices

que, ao serem inseridos no modelo, têm a capacidade de otimizar os resultados. As bandas selecionadas abrangem uma ampla faixa do espectro analisado.

No entanto, não foram incluídas bandas entre 470 nm e 550 nm, o que pode indicar que a influência do ruído afeta essa região do espectro. Em resumo, as bandas preditoras estão localizadas nos vales e picos mencionados na seção anterior (5.4.1) e na Tabela 4, que são de maior interesse para a chl-a. Comprimentos de onda entre 700 nm e 900 nm são frequentemente relatados como relevantes para a previsão de sólidos em suspensão (YANG et al., 2022).

**Figura 24 - Reflectância média dos espectros utilizados nos algoritmos (Figura 22c) e a representação das bandas preditoras selecionadas pelo método *forward***



Fonte: Autor (2023)

Os modelos XGB e SVR apresentaram resultados excelentes nas estimativas com os dados de setembro, outubro e fevereiro (Teste A, Apêndice F.1). Porém, os bons resultados obtidos,  $R^2 = 0,96$  e  $RMSE = 3,38 \mu\text{g/L}$  (XGB) e  $R^2 = 0,89$  e  $RMSE = 5,56 \mu\text{g/L}$  (SVR), podem sugerir a presença de overfitting, efeito ao qual os algoritmos de aprendizado de máquina são susceptíveis, que superestima os resultados obtidos nas análises de regressão (JUST et al., 2020). O Apêndice F.3 (Teste D) mostra que o modelo SVR também apresentou os melhores resultados quando houve redução da

dimensionalidade extraíndo as bandas predictoras. Este algoritmo obteve bons desempenhos em todos os testes ( $R^2 > 0,80$ ) com exceção do Teste B, que contou apenas com os meses de janeiro e novembro (onde o SVR não apresentou resultados). Portanto, recomenda-se o uso desse modelo para futuras análises relacionadas à Lagoa da Pampulha.

Ao analisar os resultados, verifica-se que a metodologia proposta em nosso estudo apresenta resultados semelhantes aos encontrados na literatura, sendo capaz de gerar dados para avaliar os níveis de chl-a na Lagoa da Pampulha por meio do sensor hiperespectral OCI-F™, utilizando algoritmos de aprendizado de máquina para modelar as variáveis.

Para construção do modelo foram utilizadas amostragens de pares de dados (coleta das imagens e concentração de chl-a) referentes a diferentes meses, representando os períodos climáticos em que a Lagoa da Pampulha está inserida. Este fator é importante para a construção de um banco de dados que garanta maior robustez aos algoritmos utilizados. No entanto, a faixa de dados utilizada, entre 5 µg/L e 68 µg/L, é inferior aos índices de chl-a encontrados no reservatório, especialmente durante o período seco e na região próxima à montante da Lagoa da Pampulha, onde a mediana é superior a 200 µg/L (Figura 8a).

Futuras amostragens realizadas durante os períodos em que ocorrem maiores concentrações de fitoplâncton na Lagoa da Pampulha, entre abril e setembro, podem suprir a falta de dados mencionada e melhorar a previsão do modelo. No entanto, o modelo obtido mostrou-se útil para analisar a faixa de concentração limite indicada para a classe 3, à qual a Lagoa da Pampulha está classificada pela legislação (60 µg/L).

Uma metodologia adicional que pode resultar em melhores resultados para o modelo proposto é a partição dos dados, conforme proposto por estudos como Cheng et al. (2019). Como demonstrado neste estudo (seção 5.2), a Lagoa da Pampulha apresenta características espaciais distintas em relação à qualidade da água. A região a montante (pontos 1 a 10 na Figura 4) é fortemente influenciada pelos principais tributários, os córregos Ressaca e Sarandi, enquanto as regiões intermediária e lacustre (pontos 11 a 30 na Figura 4) apresentam características diferentes, com

menores teores de sólidos em suspensão e maior transparência da água. Portanto, a interação da luz incidente com cada compartimento do reservatório é diferente, resultando em respostas espectrais distintas. Ao considerar a inclusão compartimentada dos dados de entrada nos modelos, é possível otimizar o desempenho.

Em um estudo realizado por Souza (2019), foi utilizado um espectrorradiômetro ASD FieldSpec para monitorar chl-a e turbidez na Lagoa da Pampulha. A autora empregou índices de razão espectral para estimar os parâmetros e coletou pares de dados em 26 pontos amostrais, que foram subdivididos em grupos representando as regiões antes e depois da cortina retentora de resíduos (grupo 1 e 2, respectivamente). As amostras de chl-a apresentaram valores mais altos do que os encontrados no presente estudo, com valores mínimos e máximos de 40,1 µg/L e 664,0 µg/L, respectivamente. A autora relata que os modelos propostos para chl-a apresentaram valores de  $R^2 = 0,67$  para o grupo 1, porém o RMSE teve uma grande discrepância em relação aos valores previstos (110,07 µg/L), e um  $R^2$  de 0,56 para o grupo 2 com  $RMSE = 6,38$  µg/L. Os resultados foram obtidos utilizando o índice de chl-a por Diferença Normalizada (NDCI, na sigla em inglês). Foram utilizadas as bandas 665 nm e 710 nm para o grupo 1 e 665 nm e 708 nm para o grupo 2. Concluiu-se que o estudo poderia obter melhores resultados ao utilizar diferentes algoritmos e aumentar a quantidade de amostragens, abrangendo diferentes períodos do ano para incluir as estações de chuva e seca.

Os resultados do teste B, que incluiu apenas os meses de novembro e janeiro e foi realizado com alterações no ganho das imagens, demonstraram os piores desempenhos, com  $R^2 = 0,50$  e  $RMSE = 6,54$  µg/L, obtidos pelo modelo XGB. Por outro lado, os modelos SVR e RFR não geraram resultados nesse teste. Essa observação ressalta a importância de padronizar a metodologia de aquisição de imagens, priorizando a captura pelo sensor com o mesmo ganho, mesmo em coletas realizadas em meses diferentes. Dessa forma, ajustes devem ser feitos apenas no tempo de exposição, evitando a introdução de incertezas nos resultados. Além disso, é essencial ajustar o tempo de exposição no momento da captura de cada ponto amostral para evitar a saturação ou subexposição dos pixels, garantindo a

preservação das informações. Essas práticas podem otimizar ainda mais os resultados de estimativa dos modelos propostos.

Os modelos RFR e RLM apresentaram os melhores resultados de estimativas, com  $R^2$  de 0,76 (Teste A) e 0,74 (Teste D), respectivamente. Destaca-se que o modelo RLM obteve resultados satisfatórios no Teste D, que envolveu a seleção de bandas preditoras. Esse modelo se destaca por demandar menor poder computacional, tempo reduzido para obter os resultados e facilidade de implementação, o que pode contribuir para otimizar a obtenção de informações sobre a qualidade da água na Lagoa da Pampulha.

## 6 CONCLUSÕES

A lagoa da Pampulha é um corpo d'água cuja bacia hidrográfica vem sofrendo modificações demográficas que influenciam no uso e ocupação do solo, em grande proporção sem a devida infraestrutura, provocando, ao longo das últimas décadas, perceptível piora na qualidade da água.

Nos resultados obtidos no presente estudo, percebeu-se, na avaliação histórica utilizando dados secundários, que os níveis de clorofila-a e densidade de cianobactérias apresentaram níveis de concentração superiores aos estabelecidos para o curso d'água durante todo o período analisado, comprovando as condições de hipertrofia da lagoa da Pampulha. Estes resultados podem ter influência

direta das concentrações de fósforo total observadas, sempre acima do limite estabelecido. Mesmo com o tratamento químico implementado pela prefeitura de Belo Horizonte, com a função específica de controle das concentrações de PT, o parâmetro se comporta de forma semelhante ao observado nos principais tributários, os córregos Ressaca e Sarandi, demonstrando a influência do aporte da poluição destes no reservatório e a necessidade de intervenções na bacia hidrográfica para reduzir os impactos observados.

Na análise dos dados primários, coletados entre março de 2022 e abril de 2023, observou-se a influência da cortina retentora de resíduos na espacialização da qualidade das águas da Pampulha, interferindo, de forma abrupta, na espacialização de todos os parâmetros e dividindo o reservatório em dois ambientes distintos, um a montante com qualidade da água inferior e outro a jusante da cortina com menor carga de poluição. A análise de cluster implementada aos dados primários resultou em 4 distintos grupos, 3 com espacialização dos pontos agrupados inseridos nas regiões próximas aos pontos monitorados pelo IGAM, compreendendo a compartimentação do reservatório nas regiões com características fluvial, lacustre e de transição e um quarto grupo situado logo após a cortina retentora de resíduos, região não amostrada pelo IGAM, denotando a possível necessidade de implementação de mais um ponto de amostragem pelo órgão responsável pelo monitoramento após a cortina retentora, com o intuito de captar todas as especificidades da qualidade da água da lagoa da Pampulha.

O sensor hiperespectral utilizado nas análises de estimativa de clorofila-a apresentou boa performance, tanto nos dados coletados em laboratório quanto na metodologia desenvolvida para coleta e processamento de dados obtidos a nível de campo na Lagoa da Pampulha. A utilização de dados de diferentes meses, que representam a variação sazonal do fitoplâncton, a coleta de dados em toda a área do reservatório e o uso de modelos utilizando aprendizado de máquina retornou resultados promissores na detecção da chl-a na Lagoa da Pampulha. Os resultados obtidos por meio dos testes realizados demonstraram a eficácia dos modelos XGB, SVR, RFR e RLM, que podem auxiliar no monitoramento inteligente da qualidade da água na Lagoa da Pampulha por meio do sensoriamento remoto. O modelo SVR apresentou excelentes resultados em diferentes testes, indicando seu potencial para estimativas precisas dos níveis de chl-a em águas interiores, com elevadas concentrações de poluentes. No entanto, foi observado que o modelo XGBooster pode estar sujeito a overfitting, ressaltando a importância de medidas para evitar esse fenômeno.

Ressalta-se que o refinamento da metodologia, com a aplicação de ferramentas de processamento dos dados obtidos para correções de incertezas e a aplicação da mesma metodologia de aquisição das imagens em meses diferentes podem reduzir os erros, otimizando os resultados. Além disso, outros parâmetros de qualidade da água opticamente ativos podem ser investigados para análise da performance do modelo, implementando o uso do sensor para o monitoramento amplo da Lagoa da Pampulha.

Portanto, os resultados obtidos reforçam a viabilidade de integração do sensoriamento remoto hiperespectral, utilizando algoritmos de aprendizado de máquina como suporte ao monitoramento inteligente da qualidade da água na Lagoa da Pampulha. Essa abordagem pode contribuir para o desenvolvimento de estratégias mais eficientes de gestão e preservação dos recursos hídricos, possibilitando uma tomada de decisão mais informada e sustentável.

## REFERÊNCIAS

AGUIAR, V. M. DE C.; NETO, J. A. B.; RANGEL, C. M. Eutrophication and hypoxia in four streams discharging in Guanabara Bay, RJ, Brazil, a case study. **Marine Pollution Bulletin**, v. 62, n. 8, p. 1915–1919, 2011.

ASSUNÇÃO, A. F. Dinâmica espaço-temporal de um reservatório urbano em períodos chuvosos e secos: uso do sensoriamento remoto para obtenção de parâmetros associados à qualidade da água. Dissertação (Mestrado em Saneamento, Meio Ambiente e Recursos Hídricos) - Escola de Engenharia, Universidade Federal de Minas Gerais, Belo Horizonte, 2023.

BATISTA, A. M. M.; FIGUEIREDO, C. C.; GIANI, A. Variability in a permanent cyanobacterial bloom: species-specific responses to environmental drivers. **FEMS Microbiology Ecology**. v. 94, 2018.

BARÇANTE, B. **Cianotoxinas e sua relação com espécies de cianobactérias potencialmente tóxicas em um reservatório urbano hipereutrófico**; capítulo 2. Dissertação (Mestrado em Biologia Vegetal) – Departamento de Botânica do Instituto de Ciências Biológicas, Universidade Federal de Minas Gerais, Belo Horizonte, 2019. p. 24-65.

BARÇANTE, B.; NASCIMENTO, N. O.; SILVA, T. F. G.; LEANDRO, A. R.; GIANI, A. Cyanobacteria dynamics and phytoplankton species richness as a measure of waterbody recovery: Response to phosphorus removal treatment in a tropical eutrophic reservoir. **Ecological Indicators**, v. 117, 1 out. 2020.

BAYSPEC. OCI™-F Ultra-Compact Hyperspectral Imager User Manual. n. 94, p. 1–62, 2016.

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

BELISCO, J. C. B. Impactos de eventos chuvosos na dinâmica físico-química e biológica em reservatórios urbanos. Estudo de caso: Lagoa da Pampulha. Dissertação (Mestrado em Saneamento, Meio Ambiente e Recursos Hídricos) - Escola de Engenharia, Universidade Federal de Minas Gerais. Belo Horizonte. 203p. 2017.

BEZERRA, M. P.; VIANA, E. A. P.; BRANDÃO, L. P. M.; MCGINNIS, D. F.; BEZERRA-NETO, J. F.; BARBOSA, F. A. R. Water quality evaluation and dissolved organic matter characterization of a tropical hypereutrophic reservoir and its streams treated with Phoslock® and microbial bioremediation Enzilimp®. **Environmental Science and Pollution Research**, v. 29, n. 1, p. 1375–1390, 2022.

BOECHAT, I. G.; PAIVA-MAGELA, A. B. M.; REIS, H. R.; GUCKER, B. Dissolved organic N in shallow tropical lakes and reservoirs: Contribution to total dissolved N and relationships with eutrophication. **International Review of Hydrobiology**. V. 104, p. 106-115, 2019.

BRANDÃO, L. P. M.; STAEHR, P. A.; BEZERRA-NETO, J. F. Seasonal changes in optical properties of two contracting tropical freshwater systems. **Journal of Limnology**. v. 75, ed. 3, p. 508-519, 2016.

BRASIL. Resolução CONAMA 357, de 17 de março de 2005. Conselho Nacional de Meio Ambiente. Disponível em: <<http://www.siam.mg.gov.br/sla/download.pdf?idNorma=2747>> Acesso em: 20 nov. 2022.

CAMPOS, M. O. Fatores que influenciam a distribuição espacial do fitoplâncton na Lagoa da Pampulha - BH, MG. 2010. 75 f. Dissertação (Mestrado em Saneamento, Meio Ambiente e Recursos Hídricos) - Escola de Engenharia, Universidade Federal de Minas Gerais, Belo Horizonte, 2010.

CAI, J.; CHEN, J.; DOU, X.; XING, Q. Using machine learning algorithms with in situ hyperspectral reflectance data to assess comprehensive water quality of urban rivers. **Transactions on geoscience and remote sensing**. v. 60, 2022.

CAO, Q; YU, G; SUN, S; DOU, Y; LI, H; QIAO, Z. Monitoring water quality of the Haihe River based on ground-based hyperspectral remote sensing. **Water**, 14(1), 22, 2022.

CBHSF. Resumo Executivo do Plano de Recursos Hídricos da Bacia Hidrográfica do Rio São Francisco 2016-2025 - Comitê da Bacia Hidrográfica do Rio São Francisco. p. 300, 2016.

CHENG, C.; WEI, Y.; LU, G.; XU, N. Remote sensing estimation of chlorophyll-a concentration in Taihu Lake considering spatial and temporal variations. **Environmental Monitoring and Assessment**, v. 191, n. 2, 2019.

CHONG, W. S.; ZAKI, N. H. M.; HOSSAIN, M. S.; MUSLIN, A. M.; POUR, A. B. Introducing Theil-Sen estimator for sun glint correction of UAV data for coral mapping. **Geocarto International**. v. 37, n. 15, p. 4527-4556, 2021.

COSTA, P. Avaliação da qualidade das águas superficiais de reservatórios e seus tributários por meio de abordagens quali-quantitativas: Estudo de caso dos reservatórios de Emborcação e Nova Ponte - MG. 176 f. Dissertação (Mestrado em Saneamento, Meio Ambiente e Recursos Hídricos) - Escola de Engenharia, Universidade Federal de Minas Gerais, Belo Horizonte, 2021.

COPASA. Plano de ação sistema de esgotamento sanitário da bacia da Lagoa da Pampulha. Belo Horizonte, 2022. p. 133.

COUTINHO, W. Emprego da flotação a ar dissolvido no tratamento de cursos d'água: Avaliação de desempenho da estação de tratamento dos córregos Ressaca e Sarandi afluentes à Represa da Pampulha. Dissertação (Mestrado em Saneamento, Meio Ambiente e Recursos hídricos) – Escola de Engenharia, Universidade Federal de Minas Gerais, Belo Horizonte, 2007.

CUI, Y.; FEI, M.; FU, P.; YANG, X.; ZHANG, Y.; LIU, P. Application of hyperspectral analysis of chlorophyll-a concentration inversion in Nansi Lake. **Ecological Informatics**, v. 64, 2021.

DONG, J.; WANG, G.; YAN, H.; XU, J.; ZHANG, X. A survey of smart water quality monitoring system. **Environmental Science and Pollution Research**. v. 22, p. 4893-4906, 2015.

DORNHOFER, K.; OPPELT N. Remote sensing for lake research and monitoring - Recent advances. **Ecological Indicators**. V. 64, p. 105-122, 2016.

ESTEVEES, F. DE A. Fundamentos de limnologia. Rio de Janeiro. Interciência, P. 790, 2011.

EWING, J.; OOMMEN, T.; JAYAKUMAR, P.; ALGER, R. Utilizing Hyperspectral Remote Sensing for Soil Gradation. **Remote sensing**. v.12, 20, 2020.

FIGUEIREDO, C. C.; GUCKER, B.; PINTO-COELHO, R. M.; GIANI, A. From intermitente to persistente cyanobacterial blooms: Identifying the main drivers in an urban tropical reservoir. **Journal of Limnology**, v. 35, n. 3, 2016.

FONTES, J. A. Abordagens de seleção de variáveis para classificação e regressão em dados espectrais para controle de qualidade. Dissertação (Mestrado em Engenharia de Produção) - Escola de Engenharia, Universidade Federal do Rio Grande do Sul. Porto Alegre. 101 p. 2020.

FURTADO, A. P. F. V.; MONTE-MOR, R. C. DE A.; COUTO, E. DE A. DO. Evaluation of reduction of external load of total phosphorus and total suspended solids for rehabilitation of urban lakes. **Journal of Environmental Management**, v. 296, n. May, 2021.

GAI, Y.; YU, D.; ZHOU, Y.; YANG, L.; CHEN, C.; CHEN, J. An improved model for chlorophyll-a concentration retrieval in coastal waters based on UAV-Borne Hyperspectral imagery: A case study in Qingdao, China. **Water**. v. 12, n. 10, 2020.

GARG, V.; KUMAR, A. S.; AGGARWAL, S. P.; KUMAR, V.; DHOTE, P. R.; THAKUR, P. K.; NIKAM, B. R.; SAMBARE, R. S.; SIKKIQUI, A.; MUDULI, P. R.; RATOGI, G. Spectral similarity approach for mapping turbidity of an inland waterbody. **Journal of Hydrology**, v. 550, p. 527–537, 2017.

GEGE, P.; GROESTSCH, P. A spectral model for correating sun glint and sky glint. *In: Ocean Optics XXIII*, 2016, Victoria, Canada. p. 23-28, 2016.

GHOLIZADEH, M. H.; MELESSE, A. M.; REDDI, L. A comprehensive review on water quality parameters estimation using remote sensing techniques. **Sensors (Switzerland)**, v. 16, n. 8, 2016.

GIANI, A.; TARANU, Z. E.; RUCKERT, G.; GREGORY-EAVES, I. Comparing key drivers of cyanobacteria biomass in temperate and tropical systems. **Harmful Algae**. V. 97, 2020.

GIANNONI, L.; LANGE, F.; TACHTSIDIS, I. Hyperspectral imaging solutions for brain tissue metabolic and hemodynamic monitoring: past, current and future developments. **Journal of Optics**. v. 20, n. 4, 2018.

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

GOWEN, A. A.; FENG, Y.; GASTON, E.; VALDRAMIDIS, V. Recent applications of hyperspectral imaging in microbiology. **Talanta**, v. 137, p. 43-54, 2015.

GROSS, G.; HELDER, D.; BEGEMAN, C.; HEIGH, L.; KAEWMANEE, M.; SHAH, R. Initial cross-calibration of Landsat 8 and Landsat 9 using the simultaneous underfly event. **Remote Sensing**. v. 14, n. 2418, 2022.

GUALBERTO, J. A. Comparação das técnicas Kernel e Krigagem indicativa na predição de valores de variáveis especialmente distribuídas - Estudo de caso. 2020. 97 f. Dissertação (Mestrado em Biometria) - Instituto de Biociências de Botucatu, Universidade Estadual Paulista “Júlio de Mesquita Filho”, Botucatu, 2020.

HE, J.; CHEN, Y.; WU, J.; STOW, D. A.; CRISTAKOS, G. Space-time chlorophyll-a retrieval in optically complex waters that accounts for remote sensing and modeling uncertainties and improves remote estimation accuracy. **Water Research**, v. 171, p. 115403, 2020.

HOOKER, S. B.; HOUSKEEPER, H. F.; KUDELA, R. M.; MATSUOKA, A.; SUZUKI, K.; ISADA, T. Spectral modes of radiometric measurements in optically complex waters. **Continental Shelf Research**, v. 219, 2021.

HOU, Y.; ZHANG, A.; LV, R.; ZHAO, S.; MA, J.; ZHANG, H.; LI, Z. A study on water quality parameters estimation for urban rivers based on ground hyperespectral remote sensing technology. **Environmental Science and Pollution research**. v. 29, p. 63640-63654, 2022.

HUISMAN, J.; CODD, G. A.; PAERL, H. W.; IBELINGS, B. W.; VERSPAGEN, J. M. H.; VISSER, P. M. Cyanobacterial blooms. **Nature Reviews Microbiology**. v. 16, p. 471-483, 2018.

IGAM. Avaliação da qualidade das águas superficiais de Minas Gerais em 2020: Resumo executivo anual. Instituto Mineiro de Gestão das Águas. Belo Horizonte, 2022.

INPE. Introdução ao Sensoriamento Remoto de sistemas aquáticos: princípios e aplicações. São José dos Campos. LabISA\INPE, 2019. 161 p.

JIANG, J.; TANG, S.; HAN, D.; FU, G.; SOLOMATINE, D.; ZHENG, Y. A comprehensive review on the design and optimization of surface water quality monitoring networks. **Environmental Modelling and Software**, v. 132, 2020.

JUST, A. C.; ARFER, K. B.; RUSH, J.; DORMAN, M.; SHTEIN, A.; LYAPUSTIN, A.; KLOOG, I. Advancing methodologies for applying machine learning and evaluating spatiotemporal models of fine particulate matter (PM<sub>2.5</sub>) using satellite data over large regions. **Atmospheric Environment**. v. 239, n. 117649, 2020.

KAY, S.; HEDLEY, J. D.; LAVENDER, S. Sun glint correction of high and low spatial resolution images of aquatic scenes: A review of methods for visible and near-infrared wavelengths. **Remote Sensing**. v.1, n. 4, 2009.

KELLER, S.; MAIER, P. M.; RIESE, F. M.; NORRA, S.; HOLBACH, A.; BÖRSIGH, N.; WILHELMS, A.; MOLDAENKE, C.; ZAAKE, A.; HINZ, S. Hyperspectral data and machine learning for estimating CDOM, chlorophyll a, diatoms, green algae and turbidity. **International Journal of Environmental Research and Public Health**, v. 15, n. 9, p. 1–15, 2018.

KHAN, R. M.; SALEHI, B.; MAHDIANPARI, M.; MOHAMMADIMAMESH, F.; MOUNTRAKIS, G.; QUACKENBUSH, L. J. A meta-analysis on harmful algal bloom (HAB) detection and monitoring: A remote sensing perspective. **Remote Sensing**. v. 13, n. 21, 2021.

KHEZRABAD, M. A.; ZOEJ, M. J. V.; SAFDARINEZHAD, A. A new approach for geometric correction of UAV-based pushbroom images through the processing of simultaneously acquired frame images. **Measurement**. v. 199, 2021.

KIM, S.; JEONG, J.; JAHANDAR, M.; LIM, D. C.; Recent progress in solar cell technology for low-light indoor applications. **Current Alternative Energy**. v. 2, n. 3, p. 1-15, 2018.

KUMAR, V.; SHARMA, A.; KUMAR, R.; GHARDWAJ, R.; THUKRAL, A. K.; RODRIGO-COMINO, J. Assessment of heavy-metal pollution in three different Indian water bodies by combination of multivariate analysis and water pollution indices. **Human and Ecological Risk Assessment**. V 26, p1-16, 2018.

LI, H.; XIE, X.; YANG, X.; CAO, B.; XIA, X. An integrated model of summer and winter of chlorophyll-a retrieval in the Pearl River Estuary based on hyperspectral data. **Remote Sensing**. v. 14, n. 9, 2022.

LIBÂNIO, M. Fundamentos de qualidade e tratamento de água. 3 ed. Campinas, SP. Átomo. 2010.

LOPES, F. A.; DAVIES-COLLEY, R.; PIAZI, J.; SILVEIRA, J. S. ; LEITE, A. C.; LOPES, N. I. A.. Challenges for contact recreation in a tropical urban lake: assessment by a water quality index. *Environment, Development and Sustainability*, v. 22, n. 6, p. 5409–5423, 2020.

LU, X.; SITU, C.; MA, W.; DAI, X.; HE, C.; LI, L.; ZHOU, L. Comparative analysis of four semi-analytical models for estimating chlorophyll-a concentration in case-2 waters using field hyperspectral reflectance. **International Journal of Remote Sensing**, v. 41, n. 2, p. 584–594, 2020.

MAIA, K. P.; SILVA, G. A.; LIBÂNIO, M. Aplicação de análise multivariada no estudo da frequência de amostragem e do número de estações de monitoramento de qualidade da água. **Revista Engenharia Ambiental e Sanitária**. v. 24, n. 5, 2019.

MELLO, N. A. S. T. DE; BRIGHENTI, L. S.; BARBOSA, F. A. R.; STAEHR, P. A.; BEZERRA-NETO, J. F. Spatial variability of methane (CH<sub>4</sub>) ebullition in a tropical

hypereutrophic reservoir: silted areas as a bubble hot spot. **Lake and Reservoir Management**, v. 34, n. 2, p. 105–114, 2018.

MENESES, P. R.; ALMEIDA, T. DE. Introdução ao processamento de imagem de sensoriamento remoto. Cnpq/Unb, p. 266, 2012.

MENON, H. B.; ADHIKARI, A. Remote Sensing of Chlorophyll-A in Case II Waters: A Novel Approach With Improved Accuracy Over Widely Implemented Turbid Water Indices. **Journal of Geophysical Research: Oceans**, v. 123, n. 11, p. 8138–8158, 2018.

MONTEIRO, L. N. Sistema de monitoramento da qualidade da água em tempo real. 109 f. Dissertação (Mestrado em Engenharia Elétrica) – Universidade Estadual Paulista ‘Júlio de Mesquita Filho’, Sorocaba-SP, 2020.

NEHIR, M.; FRANK, C.; ABMANN, S.; ACHTERBERG, E. P. Improving optical measurements: Non-linearity compensation of compact charge-coupled device (CCD) spectrometers. **Sensors**. v. 19, n. 2833, 2019.

NOVIYANTO, A.; ABDULLA, W. H. Segmentation and calibration of hyperspectral imaging for honey analysis. **Computers and Electronics in Agriculture**. v. 159, p. 129-139, 2019.

NOVO, E. M. L. DE M. Sensoriamento Remoto: Princípios e aplicações. 4. ed. São Paulo: Blucher, 2010.

OSANEN, J. Application of an underwater hyperspectral imager (UHI) for the study of sea ice algae. Analysis of sea ice algal chlorophyll a biomass distribution and production using hyperspectral imaging. 71 f. Dissertação (Master's in Biology) - Faculty of Fisheries, Biosciences and Economics, Departament of Arctic and Marine Biology, 2022.

PAIVA, A. C. DA E.; NASCIMENTO, N.; RODRIGUEZ, D. A.. TOMASELLA, J.; CARRIELLO, F.; REZENDE, F. S. Urban expansion and its impact on water security: The case of the Paraíba do Sul River Basin, São Paulo, Brazil. **Science of the Total Environment**, v. 720, 2020.

PATEL, A. K.; ALBARICO, F. P.; PERUMAL, P. K.; VADRALE, A. P.; NIAN, C.; CHAU, H. T. B.; ANWAR, C.; WANI, H. M.; PAL, A.; SAINI, R.; HA, L. H.; SENTHILKUMAR, B.; TSANG, Y.; CHEN, C.; DONG, C.; SINGHANIA, R. R. Algae as an emerging source of bioactive pigments. **Bioresource technology**. v. 351, 126910, 2022.

PETERSON, K. T.; SAGAN, V.; SIDIKE, P.; HASENMUELLER, E. A.; SLOAN, J. J.; KNOUFT, J. H. Machine learning-based ensemble prediction of water-quality variables using feature-level and decision-level fusion with proximal remote sensing. **Photogrammetric Engineering & Remote Sensing**. v. 85, n. 4, p. 269-280, 2019.

PLEC, D. F.; SILVA, T. G. F.; VIÇON-LEITE, B.; NASCIMENTO, N. Thermal functioning of a tropical reservoir assessed through three-dimensional modelling and high-frequency monitoring. **Revista Brasileira de Recursos Hídricos**. v. 26, 2021.

PIZANI, F. M. C.; MAILLARD, P.; AMORIM, C. C. The determination of water quality parameters by remote sensing technologies: 2000 - 2020. **Progress in Physical Geography** v. 20, p. 1–29, 2020.

QU, L.; LEI, T.; NING, D.; CIVCO, D.; YANG, X. A spectral mixing algorithm for quantifying suspended sediment concentration in the Yellow River: a simulation based on a controlled laboratory experiment. **International Journal of Remote Sensing**, v. 37, n. 11, p. 2560–2584, 2016.

QUAN, Q.; HAO, Z.; XIFENG, H.; JINGCHUN, L. Research on water temperature prediction based on improved support vector regression. **Neural computing and applications**. v. 34, p. 8501-8510, 2022.

RASTI, B.; SCHEUNDERS, P.; GHAMISI, P.; LICCIARDI, G.; CHANUSSOT, J. Noise reduction in Hyperspectral imagery: Overview and Application. **Remote Sensing**. v. 3, n. 482, 2018.

RESCK, R.; BEZERRA-NETO, J. F.; COELHO, R. M. Nova batimetria e avaliação de parâmetros morfométricos da Lagoa da Pampulha (Belo Horizonte, Brasil). **Geografias**. v. 3, n. 2, p. 24-37, 2007.

RIBA, J. Application of image sensors to detect and locate electrical discharges: a review. **Sensors**. v. 22, n. 5886, 2022.

RIETZLER, A. C.; BOTTA, C. R.; RIBEIRO, M. M.; ROCHA, O.; FONSECA, A. L. Accelerated eutrophication and toxicity in tropical reservoir water and sediments: na ecotoxicological approach. **Environmental Science and Pollution Research**. v. 25, p. 13292-13311, 2018.

ROSSITER, T; FUREY, T; McCARTHY, T; STENGEL, D. B. UAV-mounted hyperspectral mapping of intertidal macroalgae. Estuarine, **Coastal and Shelf Science**, v. 242, 2020.

ROSTAMI, A. A.; KARIMI, V.; KHATIBI, R.; PRADHAN, V. An investigation into seasonal variations of groundwater nitrate by spatial modelling strategies at two levels by kriging and co-kriging models. **Journal of Environmental Management**. v. 270, 2020.

SAGAN, V.; PETERSON, K. T.; MAIMAITIJIANG, M.; SIDIKE, P.; SLOAN, J.; GREELING, B. A.; MAALOUF, S.; ADAMS, C. Monitoring inland water quality using remote sensing: potential and limitations of spectral indices, bio-optical simulations, machine learning, and cloud computing. **Earth-Science Reviews**, v. 205, 2020.

SALGADO, A.; A. **Modelagem dinâmica de algas e cianobactérias em um reservatório de abastecimento**. 2018. 161 f. Tese (Doutorado em Ciências Ambientais) - Universidade Federal de Goiás, Goiânia, 2018.

SALMI, P.; ESKELINEN, M. A.; LEPPÄNEN, M. T.; PÖLÖNEN, I. Rapid quantification of microalgae growth with hyperspectral camera and vegetation indices. **Plants**, v. 10, n. 2, p. 1–12, 2021.

SANTOS, E. P.; SANTOS, R. A.; CARNEIRO, A. S. R.; JESUS, T. B. Sensoriamento remoto hiperespectral no estudo do corpo hídrico do baixo curso do rio Subaé, Bahia, Brasil. **Revista Brasileira de Geografia Física**. V. 11, n. 06, p. 2018, 2028, 2018.

SOUZA, M. E. V. Uso de sensoriamento remoto para estimar a concentração de componentes opticamente ativos na Lagoa da Pampulha – MG. Dissertação (Mestrado em Análise e Modelagem de Sistemas Ambientais) – Instituto de Geociências, Universidade Federal de Minas Gerais, Belo Horizonte, 2019.

SÃO PAULO, S. Fundamentos para Gestão da Água. São Paulo. CD.G Casa de Soluções e Editora, 2012.

SARAVANAN, K.; ANUSUYA, E.; RAGHAVENDRA, K.; SON, L. H. Real-time water quality monitoring using Internet of Things in SCADA. **Environmental Monitoring and Assessment**. v. 190, n. 556, 2018.

SEIDL, M. et al. Impact of urbanisation (trends) on runoff behaviour of Pampulha watersheds (Brazil). **Environmental Science and Pollution Research**, v. 27, n. 13, p. 14259–14270, 2020.

SHAIKH, M. S.; JAFERZADEH, K.; THORNBERG, B.; CASSELGREN, J. Calibration of a Hiperes-Spectral imaging system using a low-cost reference. **Sensors**. v. 21, n. 3738, 2021.

SILVA, T. F. G. et al. Impact of urban stormwater runoff on cyanobacteria dynamics in a tropical urban lake. **Water (Switzerland)**, v. 11, n. 5, 2019.

SOARES, S. F. C. Um novo método para transferência de modelos de calibração NIR e uma nova estratégia para classificação de sementes de algodão usando imagem hiperespectral NIR. Tese (Doutorado em Química) – Departamento de Química, Universidade Federal da Paraíba, João Pessoa, 2016.

SOUZA, M. E. V. Uso de sensoriamento remoto para estimar a concentração de componentes opticamente ativos na Lagoa da Pampulha – MG. Dissertação (Mestrado em Análise e Modelagem de Sistemas Ambientais) – Instituto de Geociências, Universidade Federal de Minas Gerais, Belo Horizonte, 2019.

SPERLING, M. A.; VERBYLA, M. E.; OLIVEIRA, S. M. C. Assessment of treatment plant performance and water quality data: A guide for students, researchers and practitioners. IWA Publishing, 2020, 609 p.

STENGEL, D. B.; CONNAN, S.; POPPER, Z. A. Algal chemodiversity and bioactivity: Sources of natural variability and implications for comercial application. **Biotechnology Advances**. v. 29, n. 5, p. 483-501, 2011.

SU, T. C. A study of a matching pixel by pixel (MPP) algorithm to establish an empirical model of water quality mapping, based on unmanned aerial vehicle (UAV) images. **International Journal of Applied Earth Observation and Geoinformation**, v. 58, p. 213–224, 2017.

SUN, D.; LI, Y.; WANG, Q.; GAO, J.; LE, C.; HUANG, C.; GONG, S. Hyperspectral remote sensing of the pigment C-Phycocyanin in turbid inland waters, based on optical classification. **IEEE Transactions on Geoscience and Remote Sensing**. v. 51, n. 7, 2013.

TIBSHIRANI, R.; WALTHER, G.; HASTIE, T. Estimating the number of clusters in a data set via the gap statistic. **Journal of the Royal Statistical Society. Series B (Statistical Methodology)**, v. 63, n. 2, p. 411-423, 2001.

TUNDISI, J. G.; MATSUMURA-TUNDISI, T. *Limnologia*. São Paulo: Oficina de Textos, 2013, 331 p.

UDANOR, C. N.; OSSAI, N.I.; NWEKE, E. O.; OGBUOKIRI, B. O.; ENEH, A. H.; UGWUISHIWU, C. H.; ANEKE, S. O.; WZUWGU, A. O., UGWOKE, P. O.; CHRISTIANA, A. An internet of things labelled dataset for aquaponics fish pond water quality monitoring system. **Data in Brief**. v. 43, 2022.

VILMIN, L.; FLIPO, N.; ESCOFFIER, N.; GROULEAU, A. Estimation of the water quality of a large urbanized river as defined by the European WFD: what is the optimal sampling frequency?. **Environ Sci Pollut Res**. v. 25, p. 23485–23501, 2018.

VON SPERLING, M. *Introdução à qualidade das águas e ao tratamento de esgotos*. 3 ed. Belo Horizonte: Departamento de Engenharia Sanitária e Ambiental; Universidade Federal de Minas Gerais; 2005.

WAN, Z.; XU, Y.; SAVIJA, B. On the use of machine learning models for prediction of compressive strength of concrete: Influence of dimensionality reduction on the model performance. **Materials**. v. 14, n. 4, 2021.

WANG, Q.; WU, Z.; JIN, J.; WANG, T.; SHEN, T. Low rank constraint and spatial spectral total variation for hyperspectral image mixed denoising. **Signal Processing**. v. 142, p. 11-26, 2018.

WANG, Z.; KAWAMURA, K.; SAKUNO, Y.; FAN, X.; GONG, Z.; LIM, J. Retrieval of chlorophyll-a and total suspended solids using iterative stepwise elimination partial least squares (ISE-PLS) regression based on field hyperspectral measurements in irrigation ponds in Higashihiroshima, Japan. **Remote Sensing**, v. 9, n. 3, p. 1–14, 2017.

ZHAO, R.; AN, L.; SONG, D.; LI, M.; QIAO, L.; LIU, N.; SUN, H. Detection of chlorophyll fluorescence parameters of potato levels based on continuous wavelet transform and spectral analysis. **Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy**. v. 5, 2021.

ZHOU, Z.; ZHOU, X.; APPLE, M. E.; MIAO, J.; WYSS, G.; SPANGLER, L. Quantification of cell concentration in biofuel-important microalgae using hyperspectral reflectance and hyperspectral extinction coefficient. **International Journal of Remote Sensing**, v. 40, n. 23, p. 8764–8792, 2019.

ZIBORDI, G. Experimental evaluation of theoretical sea surface reflectance factors relevant to above-water radiometry. **Optics Express**, v. 24, n. 6, p. A446, 2016.

YANG, H.; KONG, J.; HU, H.; DU, Y.; GAO, M.; CHEN, F. A review of remote sensing for water quality retrieval: Progress and challenges. **Remote sensing**. V. 14, 1770, 2022.

**APÊNDICE A – Coordenadas dos pontos amostrais monitorados no presente estudo e sua subdivisão em relação aos pontos monitorados pelo IGAM, pela sonda multiparâmetros e pela sonda multiparâmetros/sensor hiperespectral**

<b>Ponto</b>	<b>Latitude</b>	<b>Longitude</b>	<b>Descrição</b>	<b>Ponto</b>	<b>Latitude</b>	<b>Longitude</b>	<b>Descrição</b>
PV110	-19,8611	-43,9966	IGAM	P13	-19,8617	-43,9852	Campo
PV115	-19,8611	-43,9965	IGAM	P14	-19,8516	-43,9862	Campo
PV210	-19,8632	-43,9736	IGAM	P15	-19,8505	-43,989	Campo/ Sensor
PV230	-19,8478	-43,9902	IGAM	P16	-19,8578	-43,9926	Campo/ Sensor
PV235	-19,8538	-43,9791	IGAM	P17	-19,8531	-43,9788	Campo/ Sensor
PV240	-19,8509	-43,9781	IGAM	P18	-19,8497	-43,9923	Campo
P1	-19,8514	-43,9852	Campo/ Sensor	P19	-19,854	-43,9879	Campo/ Sensor
P2	-19,8496	-43,9862	Campo/ Sensor	P20	-19,8506	-43,9847	Campo
P3	-19,8503	-43,989	Campo	P21	-19,8525	-43,9777	Campo/ Sensor
P4	-19,8485	-43,9926	Campo/ Sensor	P22	-19,8505	-43,9842	Campo/ Sensor
P5	-19,8562	-43,9788	Campo/ Sensor	P23	-19,8526	-43,9933	Campo
P6	-19,8459	-43,9923	Campo	P24	-19,8516	-43,9816	Campo/ Sensor
P7	-19,8501	-43,9879	Campo	P25	-19,8513	-43,9758	Campo
P8	-19,8509	-43,9847	Campo	P26	-19,8494	-43,9838	Campo/ Sensor
P9	-19,8588	-43,9777	Campo/ Sensor	P27	-19,8537	-43,9759	Campo
P10	-19,8501	-43,9842	Campo	P28	-19,8475	-43,9802	Campo
P11	-19,8446	-43,9933	Campo	P29	-19,8503	-43,9785	Campo/ Sensor
P12	-19,8524	-43,9816	Campo/ Sensor	P30	-19,8474	-43,9843	Campo/ Sensor

**APÊNDICE B – Estatística descritiva dos dados de monitoramento do reservatório da Lagoa da Pampulha, referentes aos pontos PV230, PV235 e PV240. Subdivididos em período seco e chuvoso**

**Tabela B.1: Estatística descritiva dos dados de monitoramento do reservatório da Lagoa da Pampulha do ponto PV230**

Parâmetros	Período chuvoso						Período seco					
	n	Mediana	Média aritmética	Desvio padrão	Mínimo	Máximo	n	Mediana	Média aritmética	Desvio padrão	Mínimo	Máximo
OD (mg/L)	18	5,70	5,81	2,59	0,90	10,90	17	5,50	6,45	3,46	2,00	16,40
pH	18	7,30	7,31	0,40	6,60	8,10	17	7,30	7,19	0,64	5,80	8,60
Turbidez (NTU)	18	35,40	50,00	48,28	12,7	198,00	17	38,70	60,20	37,22	24,30	138,00
N. Amon. (mg/L N)	18	3,79	3,68	2,18	0,95	8,80	17	6,16	6,77	4,11	2,29	19,60
Nitrato (mg/L N)	18	0,05	0,10	0,13	0,02	0,44	17	0,04	0,08	0,09	0,02	0,35
Nitrito (mg/L N)	18	0,029	0,03	0,02	0,07	0,05	17	0,016	0,03	0,03	0,007	0,15
Fósforo (mg/L P)	18	0,21	0,21	0,09	0,1	0,44	17	0,30	0,42	0,35	0,12	1,55
Clorofila-a (µg/L)	18	77,36	116,44	105,63	5,34	387,00	17	225,35	282,67	204,72	26,61	748,09
Den. cian. (mm <sup>3</sup> /L)	16	59,71	28,66	23,50	0,75	69,03	17	28,66	59,70	65,33	3,59	202,37
Microcistina (µg/L)	16	0,15	1,13	2,08	0,10	7,50	17	0,54	1,28	1,98	0,15	7,69
Saxitoxina (µg/L)	16	0,07	0,23	0,38	0,02	1,57	17	0,10	0,21	0,38	0,02	1,58
Temp. da água (°C)	18	26,35	26,15	1,99	21,9	31,1	17	22,92	22,97	2,11	19,10	27,20

**Tabela B.2: Estatística descritiva dos dados de monitoramento do reservatório da Lagoa da Pampulha do ponto PV235**

Parâmetros	Período chuvoso						Período seco					
	n	Mediana	Média aritmética	Desvio padrão	Mínimo	Máximo	n	Mediana	Média aritmética	Desvio padrão	Mínimo	Máximo
OD (mg/L)	18	9,45	9,24	2,81	5,40	15,20	17	7,3	7,86	3,36	2,40	13,60
pH	18	8,15	8,22	0,68	6,90	9,30	17	7,80	7,95	0,79	6,00	9,30
Turbidez (NTU)	18	31,55	47,53	56,10	6,07	216,00	17	21,90	29,73	21,37	8,87	74,70
N. Amon. (mg/L N)	18	1,90	2,10	1,27	0,10	4,97	17	4,50	4,53	2,63	1,90	13,20
Nitrato (mg/L N)	18	0,06	0,10	0,11	0,02	0,45	17	0,04	0,10	0,22	0,02	0,94
Nitrito (mg/L N)	18	0,02	0,03	0,02	0,01	0,09	17	0,01	0,10	0,18	0,006	0,67
Fósforo (mg/L P)	18	0,16	0,17	0,09	0,04	0,37	17	0,18	0,18	0,09	0,018	0,34
Clorofila-a ( $\mu\text{g/L}$ )	18	81,60	96,53	54,27	7,48	197,28	17	67,89	101,21	96,96	0,006	338,67
Den. cian. ( $\text{mm}^3/\text{L}$ )	18	48,34	42,37	28,14	1,49	96,72	17	20,11	36,18	35,28	0,35	123,22
Microcistina ( $\mu\text{g/L}$ )	16	0,47	3,73	9,11	0,15	34,74	15	0,36	2,49	7,18	0,15	28,34
Saxitoxina ( $\mu\text{g/L}$ )	15	0,53	0,64	0,68	0,02	2,09	16	0,29	0,47	0,56	0,02	1,87
Temp. da água ( $^{\circ}\text{C}$ )	18	26,90	27,73	4,93	22,40		17	22,80	23,17	1,78	20,50	25,80

**Tabela B.3: Estatística descritiva dos dados de monitoramento do reservatório da Lagoa da Pampulha do ponto PV240**

Parâmetros	Período chuvoso						Período seco					
	n	Mediana	Média aritmética	Desvio padrão	Mínimo	Máximo	n	Mediana	Média aritmética	Desvio padrão	Mínimo	Máximo
OD (mg/L)	18	7,05	7,32	3,04	4,20	15,90	17	5,50	6,44	3,48	1,60	12,70
pH	18	7,90	8,06	0,60	6,90	9,40	17	7,50	7,81	0,92	6,10	10,10
Turbidez (NTU)	18	22,25	24,69	14,68	3,41	60,70	17	10,90	20,05	19,19	2,37	65,40
N. Amon. (mg/L N)	18	1,89	1,21	2,09	0,93	5,17	17	4,68	4,35	1,67	1,68	6,78
Nitrato (mg/L N)	18	0,06	0,06	0,05	0,02	0,18	17	0,05	0,20	0,46	1,62	2,25
Nitrito (mg/L N)	18	0,02	0,03	0,03	0,004	1,60	17	0,02	0,03	0,03	0,004	0,16
Fósforo (mg/L P)	18	0,11	0,14	0,10	0,04	0,48	17	0,12	0,13	0,06	0,06	0,25
Clorofila-a (µg/L)	18	60,90	70,60	60,12	4,45	206,59	17	49,56	57,60	38,32	4,45	127,09
Den. cian. (mm <sup>3</sup> /L)	18	27,03	33,25	25,72	0,002	78,55	16	11,20	26,92	45,73	0,16	186,67
Microcistina (µg/L)	15	0,29	1,43	2,72	0,10	9,76	14	0,28	0,67	1,14	0,15	4,56
Saxitoxina (µg/L)	15	0,47	0,65	0,68	0,02	2,11	14	0,31	0,65	0,85	0,02	3,04
Temp. da água (°C)	18	26,30	27,50	3,90	23,80		17	20,00	20,35	15,45		

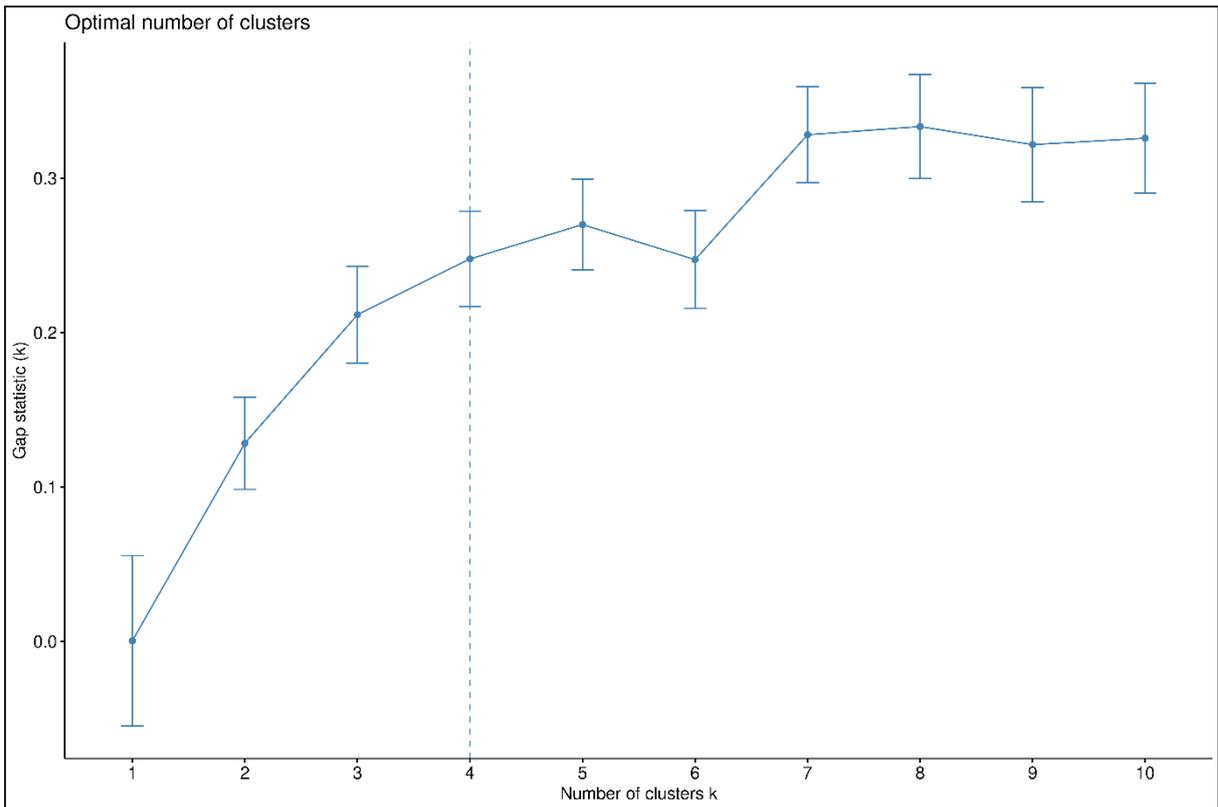
**APÊNDICE C - Resultados do teste de Mann-Whitney para comparação das medianas dos parâmetros analisados entre os períodos seco e chuvoso nos pontos PV230, PV235 e PV240. Valor de  $\alpha$  foi igual a 0,05**

Parâmetro	PV 230		PV 235		PV 240	
	Dif. Significativa	Valor-P	Dif. Significativa	Valor-P	Dif. Significativa	Valor-P
OD	Não	0.85	Não	0.20	Não	0.34
pH	Não	0.64	Não	0.26	Não	0.15
Turbidez	Não	0.17	Não	0.42	Não	0.12
N. amoniacal	Sim	$4 \times 10^{-3}$	Sim	$2 \times 10^{-4}$	Sim	$9 \times 10^{-5}$
Nitrato	Não	0.84	Não	0.30	Não	1.00
Nitrito	Não	0.22	Não	0.39	Não	0.57
Fósforo total	Sim	$7 \times 10^{-3}$	Não	0.72	Não	0.89
Clorofila-a	Sim	0.003	Não	0.54	Não	0.71
Dens. cianobactérias	Não	0.27	Não	0.36	Não	0.12
Microcistina	Não	0.06	Não	0.90	Não	0.46
Saxitoxina	Não	0.95	Não	0.87	Não	0.96
Temp. da água	Sim	$2 \times 10^{-4}$	Sim	$3 \times 10^{-5}$	Sim	$2 \times 10^{-4}$

**APÊNDICE D - Resultados para o teste de Kruskal-Wallis para verificar diferenças significativamente estatísticas entre as medianas dos 3 pontos analisados para os períodos seco e chuvoso, complementado pelo teste *post-hoc* de Dunn (Pontos 1 ,2 e 3 representam os pontos amostrais PV230, PV235 E PV240, respectivamente)**

Parâmetros	Diferença entre as medianas dos 3 pontos amostrais					
	Dif. significativa	p-Valor	Teste de Dunn	Dif. significativa	p-Valor	Teste de Dunn
OD (mg/L)	<b>0,9129</b>	0,0141	<b>0,2800</b>	<b>0,1762</b>	0,0160	<b>0,3692</b>
pH	<b>0,7744</b>	<b>0,2235</b>	<b>0,5981</b>	<b>0,4626</b>	<b>0,5833</b>	<b>0,3505</b>
Turbidez (NTU)	0,0001	0,0033	2,5 x 10 <sup>-6</sup>	0,0073	<b>0,4990</b>	0,0003
N. Amoniacal (mg/L N)	<b>0,2418</b>	0,0048	<b>0,0563</b>	0,0005	0,0004	<b>0,1908</b>
Nitrato (mg/L N)	4,5 x 10 <sup>-5</sup>	0,0001	0,0007	1,6 x 10 <sup>-7</sup>	0,0022	<b>0,4442</b>
Nitrito (mg/L N)	<b>0,1090</b>	1,6 x 10 <sup>-5</sup>	0,0012	1,5 x 10 <sup>-5</sup>	1,1 x 10 <sup>-5</sup>	<b>0,5808</b>
Fósforo total (mg/L P)	0,0456	0,0001	<b>0,1054</b>	<b>0,4893</b>	0,0009	<b>0,0826</b>
Clorofila-a (µg/L)	0,0029	<b>0,0598</b>	<b>0,2911</b>	0,0202	0,0092	<b>0,1492</b>
Den. cian. (mm <sup>3</sup> /L)	<b>0,0628</b>	0,0023	<b>0,4346</b>	0,0078	<b>0,2096</b>	5,2 x 10 <sup>-6</sup>
Microcistina (µg/L)	7,2 x 10 <sup>-6</sup>	3,3 x 10 <sup>-5</sup>	9,4 x 10 <sup>-7</sup>	2,8 x 10 <sup>-7</sup>	1,1 x 10 <sup>-5</sup>	2,2 x 10 <sup>-6</sup>
Saxitoxina (µg/L)	9,8 x 10 <sup>-6</sup>	1,5 x 10 <sup>-6</sup>	0,0209	0,0006	0,0248	0,0006
Temp. da água (°C)	<b>0,1624</b>	<b>0,9396</b>	<b>0,2621</b>	<b>0,1177</b>	<b>0,2548</b>	<b>0,1497</b>

**APÊNDICE E - Resultado gráfico do número ótimo de clusters obtido através do método gap statistic**



**APÊNDICE F - Resultados dos modelos para o teste com as bandas preditivas, expondo  $R^2$  para o treino, teste e validação e RMSE de validação para cada tipo de entrada dos dados utilizada, com destaque para o melhor resultado observado**

**Tabela G.1: Resultados dos modelos (RLM, RFR, SVR e XGB) para o Teste (A) utilizando somente os meses de setembro, outubro e fevereiro, com os valores de R<sup>2</sup> para o treino, teste e validação e RMSE de validação para cada tipo de entrada dos dados utilizada, com destaque para o melhor resultado observado**

Modelo	Tipo utilizado	R <sup>2</sup> Treino	R <sup>2</sup> Teste	R <sup>2</sup> Validação	RMSE Validação (µg/L)
Regressão Linear Multivariada (RLM)	Linear	0,98	-0,10	0,37	13,21
	Polinomial	0,43	0,14	0,65	9,76
	Exponencial	0,46	0,29	0,72	8,84
	Logarítmica	0,98	-0,03	,037	13,21
	Min/Máx	0,98	-0,11	0,37	13,21
Random Forest Regressor (RFR)	Linear	0,90	0,20	0,57	10,88
	Polinomial	0,82	0,17	0,61	10,34
	Exponencial	0,87	0,15	0,71	8,88
	Logarítmica	0,92	0,22	0,63	10,13
	Min/Máx	0,92	0,29	0,76	8,18
Support Vector Regressor (SVR)	Linear	1	0,59	0,86	6,11
	Polinomial	0,91	0,50	0,83	6,76
	Exponencial	-0,25	-0,55	-0,01	16,62
	Logarítmica	1	0,63	0,86	6,15
	Min/Máx	0,99	0,69	0,89	5,56
XGBooster Regressor (XGB)	Linear	0,96	0,26	0,96	3,38
	Polinomial	0,97	0,09	0,95	3,69
	Exponencial	1	0,36	0,92	4,58
	Logarítmica	0,96	0,23	0,96	3,38
	Min/Máx	0,96	0,16	0,96	3,38

**Tabela G.2: Resultados dos modelos (RLM, RFR, SVR e XGB) para o Teste (B) utilizando somente os meses de novembro e janeiro, com os valores de R<sup>2</sup> para o treino, teste e validação e RMSE de validação para cada tipo de entrada dos dados utilizada, com destaque para o melhor resultado observado**

Modelo	Tipo utilizado	R <sup>2</sup> Treino	R <sup>2</sup> Teste	R <sup>2</sup> Validação	RMSE Validação (µg/L)
Regressão Linear Multivariada (RLM)	Linear	1	-0,92	-0,43	11,02
	Polinomial	1	-1,38	0,06	8,96
	Exponencial	0,04	-1,57	0,02	9,11
	Logarítmica	0,13	-2,15	0,06	8,95
	Min/Máx	0,13	-1,72	0,08	8,86
Random Forest Regressor (RFR)	Linear	-	-	-	-
	Polinomial	-	-	-	-
	Exponencial	-	-	-	-
	Logarítmica	-	-	-	-
	Min/Máx	-	-	-	-
Support Vector Regressor (SVR)	Linear	-	-	-	-
	Polinomial	-	-	-	-
	Exponencial	-	-	-	-
	Logarítmica	-	-	-	-
	Min/Máx	-	-	-	-
XGBooster Regressor (XGB)	Linear	0,94	-0,43	0,50	6,54
	Polinomial	1	-0,95	0,45	6,85
	Exponencial	0,94	-0,29	0,50	6,54
	Logarítmica	0,94	-0,43	0,50	6,54
	Min/Máx	0,94	-0,44	0,50	6,54

**Tabela G.3: Resultados dos modelos (RLM, RFR, SVR e XGB) para o Teste (D) utilizando todos os meses e aplicando o método de seleção de bandas *forward inclusion stepwise*, com os valores de  $R^2$  para o treino, teste e validação e RMSE de validação para cada tipo de entrada dos dados utilizada, com destaque para o melhor resultado observado**

Modelo	Tipo utilizado	$R^2$ Treino	$R^2$ Teste	$R^2$ Validação	RMSE Validação ( $\mu\text{g/L}$ )
Regressão Linear Multivariada (RLM)	Linear	0,69	0,31	0,74	8,05
	Polinomial	0,21	0,06	0,07	15,28
	Exponencial	0,66	-0,22	0,63	9,64
	Logarítmica	0,79	0,62	0,68	8,94
	Min/Máx	0,69	0,27	0,74	8,05
Random Forest Regressor (RFR)	Linear	0,91	0,26	0,63	9,59
	Polinomial	0,91	0,27	0,58	10,26
	Exponencial	0,90	0,28	0,58	10,22
	Logarítmica	0,91	0,25	0,58	10,22
	Min/Máx	0,91	0,28	0,60	10,08
Support Vector Regressor (SVR)	Linear	0,88	0,69	0,68	8,98
	Polinomial	0,86	0,62	0,73	8,19
	Exponencial	-0,31	-0,42	-0,12	16,75
	Logarítmica	<b>0,93</b>	<b>0,78</b>	<b>0,81</b>	<b>6,87</b>
	Min/Máx	0,88	0,73	0,57	10,37
XGBooster Regressor (XGB)	Linear	1	0,19	0,55	10,61
	Polinomial	1	0,21	0,59	10,16
	Exponencial	1	0,20	0,58	10,29
	Logarítmica	1	0,20	0,55	10,61
	Min/Máx	1	0,20	0,55	10,61