

UNIVERSIDADE FEDERAL DE VIÇOSA

FELIPE CARVALHO SANTANA

**QUALIDADE E COMPORTAMENTO ESPECTRAL DA ÁGUA DE RIO AFETADO
POR REJEITO DE MINERAÇÃO**

**VIÇOSA - MINAS GERAIS
2021**

FELIPE CARVALHO SANTANA

**QUALIDADE E COMPORTAMENTO ESPECTRAL DA ÁGUA DE RIO AFETADO
POR REJEITO DE MINERAÇÃO**

Tese apresentada à Universidade Federal de Viçosa, como parte das exigências do Programa de Pós-Graduação em Solos e Nutrição de Plantas, para obtenção do título de *Doctor Scientiae*.

Orientador: Márcio Rocha Francelino

Coorientador: Elpídio Inácio Fernandes Filho

**VIÇOSA - MINAS GERAIS
2021**

**Ficha catalográfica elaborada pela Biblioteca Central da Universidade
Federal de Viçosa - Campus Viçosa**

T

S232q
2021 Santana, Felipe Carvalho, 1990-
Qualidade e comportamento espectral da água de rio
afetado por rejeito de mineração / Felipe Carvalho Santana. –
Viçosa, MG, 2021.

96 f. : il. (algumas color.) ; 29 cm.

Orientador: Márcio Rocha Francelino.

Tese (doutorado) - Universidade Federal de Viçosa.

Inclui bibliografia.

1. Água - Análise - Gualaxo do Norte, Rio (MG). 2. Água -
Qualidade - Gualaxo do Norte, Rio (MG). 3. Sensoriamento
remoto. 4. Aprendizado do computador. 5. Turbidez.

I. Universidade Federal de Viçosa. Departamento de Solos.
Programa de Pós-Graduação em Solos e Nutrição de Plantas.

II. Título.

CDD 22. ed. 333.9151

Bibliotecário(a) responsável: Bruna Silva CRB6/2552

FELIPE CARVALHO SANTANA

**QUALIDADE E COMPORTAMENTO ESPECTRAL DA ÁGUA DE RIO AFETADO
POR REJEITO DE MINERAÇÃO**

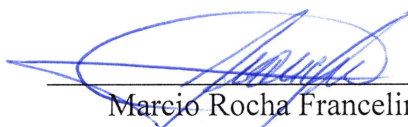
Tese apresentada à Universidade Federal de Viçosa, como parte das exigências do Programa de Pós-Graduação em Solos e Nutrição de Plantas, para obtenção do título de *Doctor Scientiae*.

APROVADA: 30 de abril de 2021

Assentimento:



Felipe Carvalho Santana
Autor



Marcio Rocha Francelino
Orientador

Ao meus avós Mauro e Cleide

DEDICO

AGRADECIMENTOS

Agradeço a Deus pela força, sabedoria e por me permitir estar vivendo esta história, iluminando e me mostrando sempre os melhores caminhos.

A Adalgisa, que se tornou minha esposa, ao longo dessa caminhada, por cuidar e me incentivar.

Aos professores pelas experiências compartilhadas e apoio.

À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES).

Ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPQ).

À Fundação de Amparo à Pesquisa do Estado de Minas Gerais (FAPEMIG) pela concessão da bolsa de doutorado e financiamento do projeto.

RESUMO

SANTANA, Felipe Carvalho Santana, D.Sc., Universidade Federal de Viçosa, abril de 2021. **Qualidade e comportamento espectral da água de rio afetado por rejeito de mineração.** Orientador: Márcio Rocha Francelino. Coorientador: Elpídio Inácio Fernandes Filho.

Com o rompimento da barragem de Fundão foram lançados cerca de 44 milhões de metros cúbicos de rejeito de mineração que atingiu diretamente a bacia hidrográfica do rio Doce. Sendo assim, esse trabalho teve com objetivos: estudar a qualidade da água do rio Gualaxo do Norte no trecho atingido pelo rompimento da barragem de Fundão por meio de diferentes técnicas. Apresentar o cenário histórico da qualidade do rio Gualaxo no que tange o parâmetro de Turbidez, antes (1999), após o acidente (2016) e após o início das atividades de recuperação (2017 a 2019). Modelar a turbidez do rio Doce utilizando técnicas de sensoriamento remoto. Predizer variáveis físicas e químicas a partir de valores de turbidez verificadas em campo. Correlacionar a resposta espectral de imagens de satélite do sensor orbital Sentinel-2 com parâmetros de qualidade da água. Levantar melhores modelos de *machine learning* que possam prever a qualidade da água dos rios. Para compor esse trabalho foram utilizados os dados da qualidade da água dos rios Gualaxo do Norte e Carmo durante o inverno de 1999 e verão de 2000, do inverno de 2016, um ano após o rompimento da barragem de Fundão e de coletas de amostras de água realizadas ao longo dos anos 2017 a 2019, após a consolidação das atividades de recuperação. No total foram realizadas 25 coletas nos trechos dos rios Gualaxo e Carmo. Foram utilizados dados de monitoramento do rio Doce em 13 estações automáticas, no total de 78 amostras de turbidez. A Turbidez no rio Gualaxo do Norte, mesmo com as atividades de recuperação teve picos de 300 e 400 NTU em 2017 e 2019. O modelo gerado para Turbidez para o rio Doce mostra que essa variável pode ser mensurada a partir de dados de sensores remotos com precisão de 0,6, podendo ser apresentada de forma contínua em todo o trecho do rio Doce. A regressão linear do SST e do Fe total foram superiores a 0,80 de R^2 mostrando a correlação linear entre as variáveis. As relações de bandas desenvolvidas neste trabalho tiveram boa correlação com as variáveis Turbidez e Cor. Os modelos de árvore de decisão como Ranger, cubist e *Random Forest*, em geral obtiveram melhores resultados na modelagem de variáveis físicas e químicas em cursos d'água. A modelagem contínua das variáveis pode ser utilizada para avaliar, a qualidade e a quantidade ao longo de todo curso d'água, em especial em áreas onde não possui pontos de monitoramento.

Palavras-chave: Análise de Componentes Principais. Turbidez. Legislação Ambiental. Sensoriamento Remoto. Machine Learning.

ABSTRACT

SANTANA, Felipe Carvalho Santana, D.Sc., Universidade Federal de Viçosa, April, 2021. **Quality and spectral behavior of river water affected by mining tailings**. Adviser: Márcio Rocha Francelino. Co-adviser: Elpídio Inácio Fernandes Filho.

With the collapse of the Fundão dam, approximately 44 million cubic meters of mining waste were launched, which directly hit the Doce River hydrographic basin. This work had the following objectives: to study the water quality of the Gualaxo do Norte river in the stretch reached by the Fundão dam rupture through different techniques. To present the historical scenario of the quality of the Gualaxo River with respect to the parameter of Turbidity, before (1999), after the accident (2016) and after the start of recovery activities (2017 to 2019). Model the turbidity of the Doce River using remote sensing techniques. Predict physical and chemical variables from turbidity values verified in the field. Correlate the spectral response of satellite images from the Sentinel-2 orbital sensor with water quality parameters. To raise better models of machine learning that can predict the water quality of rivers. To compose this work, water quality data from the Gualaxo do Norte and Carmo rivers were used during the winter of 1999 and summer of 2000, of the winter of 2016, one year after the Fundão dam burst and collections of water samples carried out over the years 2017 to 2019, after the consolidation of the recovery activities. In total, 25 collections were made on the stretches of the Gualaxo and Carmo rivers. Monitoring data from the Doce River were used in 13 automatic stations, in a total of 78 turbidity samples. Turbidity in the Gulaxo do Norte River, even with recovery activities, peaked at 300 and 400 NTU in 2017 and 2019. The use of hierarchical grouping and principal component analysis is an analysis that can be used to visualize the quality behavior of river water. The model generated for Turbidity for the Doce River shows that this variable can be measured from remote sensor data with a precision of 0.6, and can be presented continuously throughout the stretch of the Doce River. The linear regression of SST and total Fe were greater than 0.80 of R^2 showing the linear correlation between the variables. The relationship of bands developed in this work had a good correlation with the variables Turbidity and Color. The decision tree models such as Ranger, cubist and Random Forest, in general obtained better results in the modeling of physical and chemical variables in water courses. Continuous modeling of variables can be used to assess quality and quantity throughout the watercourse, especially in areas where there are no monitoring points.

Keywords: Principal Component Analysis. Turbidity. Environmental Legislation. Remote Sensing. Machine Learning.

SUMÁRIO

INTRODUÇÃO GERAL	8
REFERÊNCIAS	10
Capítulo 1 - Water quality of the Gualaxo do Norte and Carmo rivers after the Fundão dam collapse, Mariana, MG (artigo publicado: https://doi.org/10.1007/s11270-021-05113-3)	13
ABSTRACT	14
1 INTRODUCTION.....	15
2 MATERIALS AND METHODS	16
3 RESULTS.....	19
3.1 Descriptive statistics	19
3.2 Multivariate analysis	22
4 DISCUSSION.....	24
5 CONCLUSIONS	31
REFERENCES	32
Capítulo 2 - Sensoriamento Remoto aplicado a modelagem de águas interiores	37
RESUMO	38
1 INTRODUÇÃO.....	39
2 MATERIAL E MÉTODOS.....	40
2.1 Sensoriamento Remoto.....	42
2.2 Covariáveis	43
2.3 Modelagem de qualidade da água	43
2.4 Avaliação dos algoritmos de aprendizado de máquina	47
3 RESULTADOS E DISCUSSÃO	47
4 CONCLUSÕES	57
REFERÊNCIAS	58
Capítulo 3 - Modelagem de parâmetros da qualidade da água de rio afetado por rejeito de mineração por meio do comportamento espectral	61
RESUMO	62
1 INTRODUÇÃO.....	63
2 MATERIAL E MÉTODOS.....	65
2.1 Área de Amostragem.....	65
2.2 Aquisição e espacialização dos dados	66
2.2.1 Parâmetros gerais.....	66
2.2.2 Aquisição dos parâmetros de qualidade da água	68
2.2.3 Covariáveis espectrais	68
2.2.4 Covariáveis não espectrais e precipitação acumulada	71
2.3 Modelagem de qualidade da água	72
2.4 Avaliação dos algoritmos de aprendizado de máquina	73
3 RESULTADOS E DISCUSSÃO	74
4 CONCLUSÕES	88
REFERÊNCIAS	89
CONSIDERAÇÕES FINAIS	96

INTRODUÇÃO GERAL

A extração mineral foi responsável por 3% do PIB brasileiro em 2019 (IBGE, 2020), sendo que entre os recursos minerais o minério de ferro foi o mais extraído, mais de 510 milhões de toneladas, contribuindo com 68% da extração mineral nacional (ANM, 2020). Esse minério é encontrado principalmente em rochas de formações ferríferas bandadas, como itabirito, que contêm de 30 a 35% de Fe em sua composição. Sendo assim, as formações ferríferas bandadas necessitam ser beneficiadas para remoção de outros materiais que as compõe, principalmente a sílica (Carvalho et al., 2014).

No processo de beneficiamento é gerado o rejeito que tem como destino final a disposição em barragens (Ghose and Sen, 2001; Rao et al., 2016; Srivastava et al., 2001). Em virtude do aumento da extração mineral no Brasil, que entre os anos de 1995 e 2012, contribuiu de 0,8% para 4,3% do PIB nacional (Souza Júnior et al., 2018), elevou a quantidade de barragens, sendo registrados 790 em 2018 (ANA, 2018).

Logo, devido o aumento da extração mineral e do alto potencial poluidor das barragens, em menos de dez anos três grandes acidentes foram registrados, retirando vidas e causando danos ambientais. Em 2019, o rompimento da barragem de Córrego do Feijão em Brumadinho, liberou cerca de 11 milhões de toneladas de rejeito na calha do rio Paraopeba, depositando sedimentos (Porsani et al., 2019) e afetando a qualidade das águas (Thompson et al., 2020). O rompimento do mineroduto da Anglo American em Santo Antônio do Grama em 2018, liberou 300 toneladas de rejeito, constituído de 70% de Fe (Anglo American, 2018) que afetou primeiramente o córrego Santo Antônio, e o rio Casca até sua confluência com o rio Doce. Menos recente, o rompimento da barragem de Fundão em Mariana em 2015, que lançou 43,8 Mm³ de rejeito de mineração atingindo toda a bacia do rio Doce até sua foz no oceano atlântico em Linhares/ES (Marta-Almeida et al., 2016). Boa parte desse rejeito ficou depositado nas calhas do rio Gualaxo do Norte e Carmo, sendo a fração grosseira, depositada na Pequena Central Hidrelétrica (PCH) de Risoleta Neves (Schaefer et al. 2016).

Em virtude do dano ambiental, estudos foram realizados para entender a dinâmica do rejeito nos cursos d'água e terraços, e suas implicações na qualidade do meio físico ambiental (Davila et al., 2020; Guerra et al., 2017; Silva et al., 2016), além de identificar o grau de contaminação das águas após esse acidente (Foesch et al., 2020; Santana et al., 2021; Silva et al., 2018), embora esses estudos tratem os danos de forma pontual são partes importantes para o entendimento espacial dos danos.

Como avanço de sensores com a disponibilização de dados com maior resolução espacial e menor resolução temporal, estudos vêm sendo realizados com intenção de verificar a qualidade das águas por meio de sensores remotos em áreas atingidas pelo rompimento da barragem de Fundão (Coimbra et al., 2019; Francini-Filho et al., 2019; Marta-Almeida et al., 2016; Oliveira et al., 2012; Rudorff et al., 2018), no entanto, não avaliam as águas interiores. A utilização de sensoriamento remoto para o monitoramento em águas interiores também pode ser aplicado, informando valores de parâmetros de qualidade de água em locais onde não possui dados (Barbosa et al., 2019; Isidro et al., 2018; Panepinto and Genon, 2010), no entanto, essas técnicas necessitam de maior desenvolvimento, como na escolha de índices espectrais que se relacionam com os parâmetros de qualidade, e de algoritmos que possam ser utilizados para prever valores em locais não amostrados.

Sendo assim, a tese está dividida em três capítulos. O primeiro, intitulado “Water Quality of the Gualaxo do Norte and Carmo Rivers After the Fundão Dam Collapse, Mariana, MG”. Tratou de estudar as consequências ambientais do lançamento dos rejeitos de mineração de ferro após o rompimento da barragem de Fundão na qualidade da água do rio Gualaxo do Norte, primeiro rio que recebeu a carga de rejeito proveniente do rompimento da barragem de Fundão.

O segundo capítulo intitulado “Sensoriamento Remoto aplicado a modelagem de águas interiores” abordou a dinâmica de variáveis físicas e químicas da qualidade da água dos rios afetados e de cursos d’água interiores por meio de técnicas de sensoriamento remoto.

O terceiro, intitulado: “Modelagem de parâmetros da qualidade da água de rio afetado por rejeito de mineração por meio do comportamento espectral”, traz alternativa de monitoramento de águas interiores por meio de imagens orbitais gratuitas.

REFERÊNCIAS

- ANA. (2018). Relatório de segurança de Barragens 2017.
- Anglo American. (2018). Atualização do evento do mineroduto do Minas-Rio.
<https://brasil.angloamerican.com/pt-pt/imprensa/noticias/2018/13-03-2018>
- ANM, A. N. de M. (2020). Anuário Mineral Brasileiro: principais substâncias metálicas. Brasília. https://www.gov.br/anm/pt-br/centrais-de-conteudo/publicacoes/serie-estatisticas-e-economia-mineral/anuario-mineral/anuario-mineral-brasileiro/amb_2020_ano_base_2019_revisada2_28_09.pdf
- Barbosa, C. C. F., Novo, E. M. L. M., & Martins, V. S. (2019). Introdução ao sensoriamento remoto de sistemas aquáticos.
- Carvalho, P. S. L. de, Silva, M. M. da, Rocio, M. A. R., & Moszkowicz, J. (2014). Minério de Ferro. BNDES Setorial, 39, 197–234. <http://www.bndes.gov.br/bibliotecadigital>
- Coimbra, K. T. O., Alcântara, E., & de Souza Filho, C. R. (2019). An assessment of natural and manmade hazard effects on the underwater light field of the Doce River continental shelf. *Science of the Total Environment*, 685, 1087–1096.
<https://doi.org/10.1016/j.scitotenv.2019.06.127>
- Davila, R. B., Fontes, M. P. F., Pacheco, A. A., & Ferreira, M. da S. (2020). Heavy metals in iron ore tailings and floodplain soils affected by the Samarco dam collapse in Brazil. *Science of the Total Environment*, 709, 136151.
<https://doi.org/10.1016/j.scitotenv.2019.136151>
- Foesch, M. D. S., Francelino, M. R., Rocha, P. A., & Gomes, A. R. L. (2020). River Water Contamination Resulting from the Mariana Disaster, Brazil. *Floresta e Ambiente*, 27(4), 2–10. <https://doi.org/10.1590/2179-8087.013218>
- Francini-Filho, R. B., Cordeiro, M. C., Omachi, C. Y., Rocha, A. M., Bahiense, L., Garcia, G. D., et al. (2019). Remote sensing, isotopic composition and metagenomics analyses revealed Doce River ore plume reached the southern Abrolhos Bank Reefs. *Science of The Total Environment*, 697, 134038. <https://doi.org/10.1016/j.scitotenv.2019.134038>
- Ghose, M. K., & Sen, P. K. (2001). Characteristics of Iron Ore Tailing Slime in India and. *Environmental Monitoring and Assessment*, 68, 51–61.
- Guerra, M. B. B., Teaney, B. T., Mount, B. J., Asunskis, D. J., Jordan, B. T., Barker, R. J., et al. (2017). Post-catastrophe Analysis of the Fundão Tailings Dam Failure in the Doce River System, Southeast Brazil: Potentially Toxic Elements in Affected Soils. *Water, Air, and Soil Pollution*, 228(7). <https://doi.org/10.1007/s11270-017-3430-5>

- IBGE, I. B. de G. e E. (2020). Indicadores IBGE - Contas nacionais trimestrais. Contas Nacionais Trimestrais: Indicadores de Volume e Valores Correntes (Vol. 8). <http://www.ibge.gov.br/%0Ahttp://repositorio.unan.edu.ni/2986/1/5624.pdf%0Ahttps://www.unhcr.org/publications/manuals/4d9352319/unhcr-protection-training-manual-european-border-entry-officials-2-legal.html?query=excom> 1989
- Isidro, C. M., McIntyre, N., Lechner, A. M., & Callow, I. (2018). Quantifying suspended solids in small rivers using satellite data. *Science of the Total Environment*, 634, 1554–1562. <https://doi.org/10.1016/j.scitotenv.2018.04.006>
- Marta-Almeida, M., Mendes, R., Amorim, F. N., Cirano, M., & Dias, J. M. (2016). Fundão Dam collapse: Oceanic dispersion of River Doce after the greatest Brazilian environmental accident. *Marine Pollution Bulletin*, 112(1–2), 359–364. <https://doi.org/10.1016/j.marpolbul.2016.07.039>
- Oliveira, E. N. de, Knoppers, B. A., Lorenzetti, J. A., Medeiros, P. R. P., Carneiro, M. E., & Souza, W. F. L. de. (2012). A satellite view of riverine turbidity plumes on the NE-E Brazilian coastal zone. *Brazilian Journal of Oceanography*, 60(3), 283–298. <https://doi.org/10.1590/S1679-87592012000300002>
- Panepinto, D., & Genon, G. (2010). Modeling of Po River Water Quality in Torino (Italy). *Water Resources Management*, 24(12), 2937–2958. <https://doi.org/10.1007/s11269-010-9588-x>
- Porsani, J. L., de Jesus, F. A. N., & Stangari, M. C. (2019). GPR survey on an iron mining area after the collapse of the tailings Dam I at the Córrego do Feijão mine in Brumadinho-MG, Brazil. *Remote Sensing*, 11(7), 1–13. <https://doi.org/10.3390/RS11070860>
- Rao, G. V., Markandeya, R., & Sharma, S. K. (2016). Recovery of Iron Values from Iron Ore Slimes of Donimalai Tailing Dam. *Transactions of the Indian Institute of Metals*, 69(1), 143–150. <https://doi.org/10.1007/s12666-015-0809-0>
- Rudorff, N., Rudorff, C. M., Kampel, M., & Ortiz, G. (2018). Remote sensing monitoring of the impact of a major mining wastewater disaster on the turbidity of the Doce River plume off the eastern Brazilian coast. *ISPRS Journal of Photogrammetry and Remote Sensing*, 145, 349–361. <https://doi.org/10.1016/j.isprsjprs.2018.02.013>
- Santana, F. C., Francelino, M. R., Schaefer, C. E. G. R., Veloso, G. V., Fernandes-Filho, E. I., Santana, A. de J. P., et al. (2021). Water Quality of the Gualaxo do Norte and Carmo Rivers After the Fundão Dam Collapse, Mariana, MG. *Water, Air, & Soil Pollution*, 232(4). <https://doi.org/10.1007/s11270-021-05113-3>

- Schaefer, C. E. G. R., Santos, E. E. dos, Fernandes Filho, E. I., & Assis, I. R. de. (2016). Paisagens de Lama: Os Tecnosolos para recuperação ambiental de áreas afetadas pelo desastre da barragem do Fundão, em Mariana. *Boletim informativo da SBCS*, 1(1), 18–23.
- Silva, A. C., Cavalcante, L. C. D., Fabris, J. D., Júnior, R. F., Barral, U. M., Farnezi, M. M. de M., et al. (2016). Chemical, mineralogical and physical characteristics of a material accumulated on the river margin from mud flowing from the collapse of the iron ore tailings dam in Bento Rodrigues, Minas Gerais, Brazil. *Revista Espinhaço*, 5(2), 44–53.
- Silva, D. de C., Bellato, C. R., Marques Neto, J. de O., & Fontes, M. P. F. (2018). Trace elements in river waters and sediments before and after a mining dam breach (Bento Rodrigues, Brazil). *Quimica Nova*, 41(8), 857–866. <https://doi.org/10.21577/0100-4042.20170252>
- Souza Júnior, T. F. de, Moreira, E. B., & Heineck, K. S. (2018). Barragens De Contenção De Rejeitos De Mineração No Brasil. *Holos*, 5, 2–39. <https://doi.org/10.15628/holos.2018.7423>
- Srivastava, M. P., Pan, S. K., Prasad, N., & Mishra, B. K. (2001). Characterization and processing of iron ore fines of Kiruburu deposit of India. *International Journal of Mineral Processing*, 61(2), 93–107. [https://doi.org/10.1016/S0301-7516\(00\)00030-2](https://doi.org/10.1016/S0301-7516(00)00030-2)
- Thompson, F., de Oliveira, B. C., Cordeiro, M. C., Masi, B. P., Rangel, T. P., Paz, P., et al. (2020). Severe impacts of the Brumadinho dam failure (Minas Gerais, Brazil) on the water quality of the Paraopeba River. *Science of the Total Environment*, 705, 1–6. <https://doi.org/10.1016/j.scitotenv.2019.135914>

Capítulo 1 - Water quality of the Gualaxo do Norte and Carmo rivers after the Fundão dam collapse, Mariana, MG (artigo publicado: <https://doi.org/10.1007/s11270-021-05113-3>)

ABSTRACT

We investigated the effects of Fe-mining tailings of the Fundão dam on water quality, concerning physical and chemical parameters, aiming to infer the recovery process of the affected areas, and prospects of water quality in the coming years. Water quality data from the Gualaxo do Norte and Carmo rivers were evaluated in three moments: (i) before the dam rupture (1999–2000), (ii) 1 year after the rupture of the Fundão dam (2016), and (iii) after the implementation of recover mitigation activities (2017–2019). Concerning the variables evaluated, the mean concentrations of Fe and Mn were higher in the dry period, and increasing turbidity was detected in the rainy season. The turbidity, even after the recovery measures, peaked at 300 and 400 NTU in 2017 and 2019, respectively. The concentrations of dissolved Fe, Mn, and Cr, closely related to the tailings, decreased with rainfall by a dilution mechanism. Conversely, the concentrations of Cd, Al, Zn, and Cu increased in the rainy season, probably by greater erosion. The implementation of the recovery actions at the Gualaxo do Norte headwaters are helping to reduce the Mn and Fe concentrations and turbidity downriver, all associated with the Fe-mining tailings.

Keywords: Principal component analysis; Turbidity; Color water; Dissolved iron; Dissolved manganese; Environmental legislation

1 INTRODUCTION

The Minas Gerais state, in southeastern Brazil, extracted more than 406 million ton of iron ore in 2017, which represented 68% of the Brazilian Fe production (DNPM 2019), besides having 81.6% of the country's iron reserves (ANM 2019). However, the volume of ore extracted results in a high environmental liability, due to the amount of tailings generated in the ore processing (Mechi and Sanches 2010; Schaefer et al. 2016).

In the 'Ore Treatment Units'; OTU-(place that allocates the residue from the plants to the tailing's dams); consists of successive steps to concentrate iron ore, operate soon after extraction, with ore fragmentation, concentration for flotation, and drainage, in the wet procedure, as in the case of Fundão dam (Quaresma, 2009). In this unit, tailings dams are the final destination of the silica rich material without market value (Ghose and Sen 2001; Srivastava et al 2001; Rao et al 2016).

Most mining operations in Minas Gerais state (MG) are categorized as large because the amount of iron ore mined per year exceeds one and a half million tons (ANM 2019). However, mining activities are considered by the environmental agencies as having medium pollution potential, depending on the scale. However, OTU with wet treatment of tailings are considered to have a large pollution potential, due to the presence of large volumes of disposed mot ends in dams.

The Fundão dam, located in the municipality of Mariana, MG, was characterized as a low-risk enterprise with a high associated pollution potential, according to the DNPM's Mineral Dams Registry (National Department of National Production). It was originally designed to receive more than 100 million tons of ore tailings, including fine and sandy tailings (SUPRAM 2008).

In November 2015, after successive adaptations to increase the containment volume, occur the Fundão tailing dam collapse in zone releasing 43.8 million cubic meters of mining waste, directly at Doce River watershed, reaching more than 600 km downstream to the mouth of the Doce River, and the Atlantic Ocean (Marta-Almeida et al. 2016; Golder Associates 2017). The mud plume had high turbulence and, dragged a mix of riverbed materials that had previously lain inert at the river bottom, causing changes in water quality parameters at different levels (Theis et al. 1988; Coimbra et al. 2019). The main characteristics of the plume were low amounts of clay and high silt and sand contents, in addition to low values of heavy metals, with the exception of manganese (Schaefer et al. 2016; Silva et al. 2016).

Since the Fundão tailing dam collapsed, the monitoring of water quality in the Doce River Basin has become a relevant factor in ensuring potability and restoring life in the affected areas. Such monitoring not only refers to the material pertaining to the Fundão dam tailing but also to the riverbed material that has been disturbed and transported in the affected rivers.

Several monitoring and recovery programs have been implemented in the Doce River Basin. The systematic qualitative and quantitative monitoring program (PMQQS); Springs program aiming out recovering more than five thousand springs; Program for the recovery of permanent preservation areas (APP) and water recharge aiming out recovering 40,000 hectares of degraded permanent preservation areas; and Revegetation, and backfill of the floodplains affected, to rehabilitate 2,000 hectares of floodplain and terraces directly affected by of tailings depositions at the Gualaxo do Norte, Carmo and Doce rivers.

It is known that some inert materials can affect water quality, depending on the volume suspended or dissolved in water, affecting riparian communities. In the present work we evaluated the water quality of the Gualaxo do Norte and Carmo rivers in the first zone, just below the Fundão dam burst. The effects the tailings on water quality with, regard to physical and chemical parameters, their interplay and to draw inferences about the recovery process of the affected areas.

2 MATERIALS AND METHODS

Previous water quality data obtained in 1999 by Costa (2001) of the Gualaxo do Norte and Carmo rivers from was used on pre-collapse reference. For water quality data after the collapse, we used physical and chemical characteristics and spectral behavior of the contaminated waters in the Gualaxo do Norte and Carmo rivers obtained in 2016 by Foesch et al. (2020). It must be emphasized that the upper Carmo river sector before the Gualaxo do Norte mouth was not affected by any tailing deposition after the Fundão dam collapse, representing a reference for local background.

In addition, we collected water samples, monthly, out over the years 2017 to 2019 after the implementation of different recovery activities along the affected areas. Samples were collected at a maximum depth of 30 cm with the aid of a 2 m driller at the channel margin. In situ analyses of physical-chemical parameters (pH, temperature) were performed using a multiparameter probe (PCE Instruments pH Meter PCE-PHD) and collections of simple water samples were carried out according to APHA (2005) at 13 sites distributed in easily accessible

places with low disturbances, along the Gualaxo do Norte and Carmo rivers (Figure 1). In total, 325 collections were made in the 25 field campaigns in the study area.

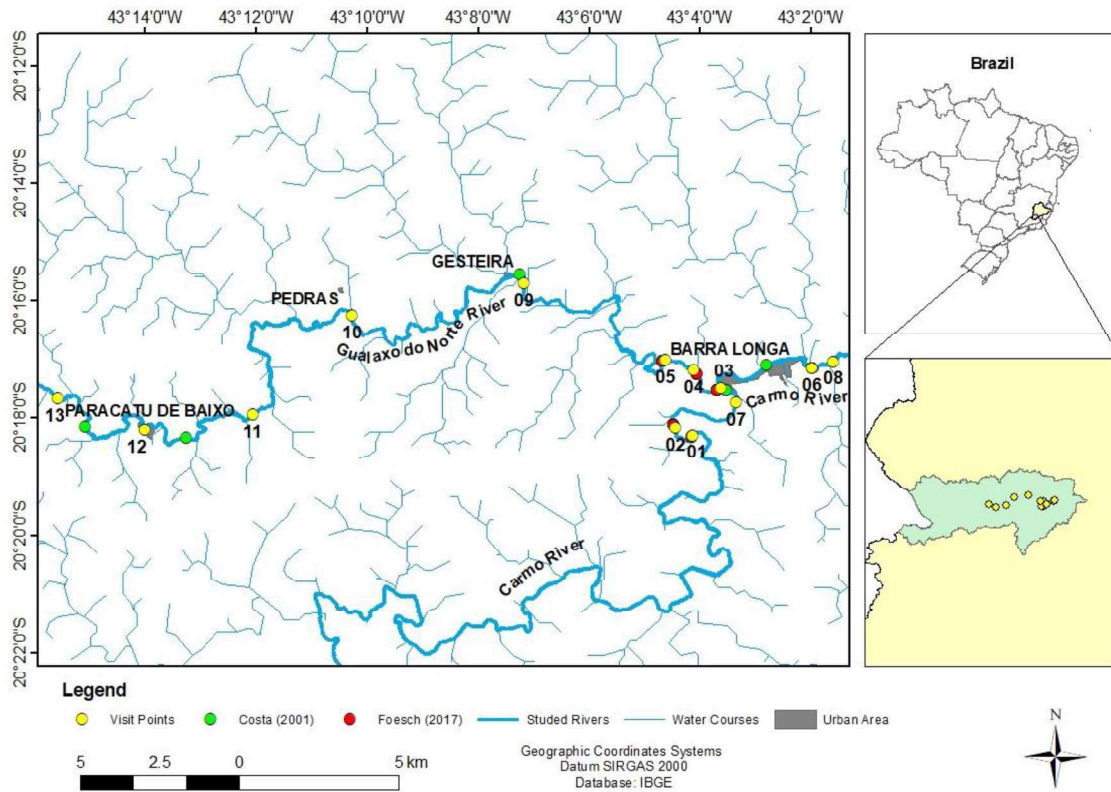


Figure 1: Location of water sample collection points evaluated between 1999 and 2000 before dam collapse by Costa (2001; green), in 2016 by Foesch (2017; red), and between 2017 and 2019 (yellow).

In the laboratory, of the following physical and chemical characteristics were determined: total suspended solids (TSS), turbidity, and color; and water-soluble: chromium (Cr), copper (Cu), manganese (Mn), iron (Fe), cadmium (Cd), zinc (Zn), and aluminum (Al).

Turbidity was determined using a turbidimeter (model 2100 AN Turbidimeter-Hach; unit: NTU) and the apparent color was determined using an Aqua Color Cor-Policontrol instrument (unit: uC). The TSS analyses were performed using the gravimetric method (unit: mg L^{-1}), in glass fiber membrane (Whatman, 47 mm in diameter, particle retention greater than $1.2 \mu\text{m}$, thickness of $260 \mu\text{m}$). Water samples for the determination of dissolved metals were stored in sterile container. After collection, the samples were acidified with HNO_3 and refrigerated before analysis.

For analysis of dissolved metals, the water samples were filtered through a cellulose acetate membrane (particle retention $0.45 \mu\text{m}$), determine the dissolved metal contents according to EPA 3005A (EPA 1992) by Atomic Spectrometry Laboratory at UFV; using Agilent Technologies 200 Series AA Model 240 FS the contents Cr, Cu, Mn, Fe, Cd, Zn, and

Al dissolved in water were determined. Before reading the samples, the spectrometer calibration curve was used. The detection limit was performed according to González and Herrador (2007), were, determined based RDC n° 166 (Brasil 2017) on the inclination of the calibration curve and the standard deviation of the standard deviation of the intercept with the Y axis of three calibration curves. For each evaluated parameters, and all analyses were performed on a single sample.

In the study area there is less rainfall in winter when compared to summer, in the months of October to March the precipitation is more than 100 mm per month, already in the dry period that goes from April to September less than 100 mm per month, and the average annual precipitation is 1804 mm, for the history data (Pedreira and Souza 2011). For the study, period rainfall in 2017 was 1048,3 mm, 283,2 mm in dry and 765,1 mm in wet season; in 2018 was 1538,9 mm, 219,7 in dry and 1319,2 in wet season; in 2019 was 1049,65 mm, 206,05 mm in dry and 843,6 mm in wet season. Data analysis was performed considering the dry season (April to September) and rainy season (October to March). The 13 sampling sets were grouped into three sectors: the Gualaxo do Norte River (eight points); the Carmo River upstream (unaffected), the Barra Longa city (three points); and Carmo River downstream (affected) (two points).

Descriptive statistics (mean, standard deviation, and coefficient of variation) of the variables sampled in the period from 2017 to 2019 were obtained, considering the grouping by river sectors, and the averages were assessed by the Kruskal-Wallis non-parametric test at a probability of 0.001, since for this test assumptions such as the normality of the variable, homogeneity of variances between treatments are not required. To identify the impact of recovery activities, principal component analysis (PCA) was performed to select the quality parameters, grouping the collection sites according to their characteristics and seasonal period. In the PCA, 50 samples were used for the Carmo River downstream (affected), 34 in the dry period and 16 in the wet period; 200 samples for the Gualaxo do Norte, 136 in the dry period and 64 in the wet period; and 75 samples for the Carmo River upstream (unaffected), 51 in the dry period and 24 in the wet period. These analyses were carried out separately for the two periods evaluated: shortly after the collapse (Foesch 2020) and after the recovery actions.

The recovery activities implemented were the impediment of access to animals and people in the areas of springs and Permanent Preservation Areas (PPA), allowing natural regeneration, along with revegetation with native species, these actions took place throughout the area of the Doce river basin. The species used in revegetation started with the planting of legumes, such as *Crotalaria sp.*, *Mucuna aterrima*, *Cajanus cajan*; grasses: *Brachiaria* and

Pennisetum purpureum, after this phase of biomass insertion, native tree species such as *Acacia mangium* were planted. The mud deposit areas, which comprise the areas of riverbed and banks of the Gualaxo do Norte, Carmo and Doce rivers, including its trainers and tributaries, in the municipalities of Mariana, Barra Longa, Rio Doce and Santa Cruz do Escalvado, not only the revegetation of the PPA was carried out, but also the revegetation in the entire deposit area of the dam sediment, also installed rocking of stones in the intricacies, drainages of the terraces with installation of bio blankets to reduce the speed of water and surface transport of sediment to the river (Renova 2019).

The evaluated parameters were compared with the Maximum Allowable Value (MAV) of the current legislation, which classifies water studied as class 2 (CONAMA 2005).

3 RESULTS

3.1 Descriptive statistics

Among the physical and pH parameters evaluated, the temperature values had the lowest values of the coefficient of variation (COV), unlike turbidity, apparent color, and TSS, in sectors that received the tailing loads from the dam. At Fundão, Gualaxo do Norte, and downstream Carmo, COV values greater than 100 were recorded (Table 1). The mean water temperature in the rainy season was 28 °C the three sectors evaluated, while in the dry period it varied among the different sectors.

In general, the chemical parameters evaluated showed great variability, and differed between the dry and rainy seasons; in the rainy season, Cu, Mn, Zn, and Al had greater variability in the three sectors evaluated. The values of dissolved Cr were those that showed the least variability, with close values of COV in the dry and rainy seasons.

The mean concentrations of Fe and Mn were higher in the dry period, and decreased in the rainy period probably by dilution. The dissolved amounts of Cr, Fe, and Cd showed the least variability in both periods. However, in the most affected Gualaxo do Norte River and the downstream Carmo River sectors, the variability of Fe increased in the dry period.

The concentrations of Cr and Cu were higher in the dry period than in rainy period. Maximum dissolved Cr concentrations of 0.093, 0.137, and 0.130 mg L⁻¹ and minimum dissolved Cr concentrations of 0.011, 0.001, and 0.004 mg L⁻¹, respectively, were found in the upstream Carmo River sectors (unaffected) in the rainy season and in the Gualaxo do Norte River and the downstream Carmo sector in the dry season.

Table 1: Mean, standard deviation (S.D.); coefficient of variation (COV); maximum (max.) and minimum (min.) values of the variables monitored from April 2017 to November 2019 in the upstream stretch of the Carmo River, the Gualaxo do Norte stretch, and the downstream stretch of the Carmo River after the confluence with the Gualaxo do Norte River

River	Season	Statistic	Temp. °C	pH	Turb. NTU	Color mg P L ⁻¹	TSS	Cr	Cu	Mn	Fe	Cd	Zn	Al
Carmo upstream	Rainy	Mean	28.37*	7.66	55.53*	86.08*	47.69*	0.038	0.029*	0.093	0.197*	0.057	0.096	0.105
		S.D.	2.96	0.70	94.25	126.73	89.12	0.016	0.074	0.171	0.167	0.049	0.120	0.177
		COV	10.42	9.09	169.73	147.23	186.86	41.74	254.50	185.12	84.69	86.43	125.14	169.60
		Max.	35.00	8.78	330.50	402.50	341.20	0.093	0.359	0.594	0.579	0.109	0.504	0.630
		Min.	25.00	6.62	2.26	6.40	0.100	0.011	0.001	0.001	0.015	0.001	0.004	0.005
	Dry	Mean	24.32*	7.22	6.67*	17.10*	5.80*	0.045	0.037*	0.058	0.089*	0.054	0.091	0.051
		S.D.	2.78	1.57	7.81	13.45	8.90	0.014	0.035	0.040	0.060	0.049	0.149	0.082
		COV	11.43	21.76	117.19	78.67	153.42	31.40	94.43	68.69	67.13	89.85	164.24	160.26
		Max.	30.40	9.74	38.50	65.50	43.87	0.088	0.081	0.091	0.246	0.109	0.635	0.332
		Min.	19.10	3.30	1.60	4.05	<0.001	0.007	0.001	0.001	0.001	0.001	0.001	0.001
Gualaxo do Norte	Rainy	Mean	28.20*	7.55	120.90*	127.18*	72.16*	0.040	0.027*	0.124	0.303*	0.054	0.096	0.081
		S.D.	2.42	0.74	157.58	139.60	131.59	0.018	0.030	0.254	0.221	0.048	0.091	0.188
		COV	8.58	9.82	130.33	109.76	182.35	45.05	113.05	204.28	72.97	89.50	94.60	232.33
		Max.	33.70	8.86	659.00	420.50	646.75	0.082	0.081	0.989	0.786	0.109	0.296	1.388
		Min.	24.10	5.71	2.44	4.27	<0.001	0.002	0.001	0.001	0.018	0.001	0.002	0.003
	Dry	Mean	23.25*	7.23	34.88*	44.27*	12.82*	0.043	0.039*	0.052	0.372*	0.059	0.090	0.098
		S.D.	2.90	1.66	47.96	46.69	19.02	0.021	0.035	0.038	0.528	0.048	0.152	0.175
		COV	12.46	23.03	137.49	105.47	148.28	48.11	89.35	73.09	141.94	80.70	169.47	179.64
		Max.	29.60	10.67	349.00	363.00	131.80	0.137	0.081	0.091	3.013	0.109	0.703	0.999
		Min.	17.60	3.08	2.80	5.81	0.625	0.001	0.001	0.001	0.004	0.001	0.001	0.002
Carmo downstream	Rainy	Mean	28.84*	7.53	61.60*	100.15*	44.34*	0.042	0.027*	0.116	0.211*	0.041	0.102	0.073
		S.D.	2.59	0.77	78.25	132.43	69.47	0.016	0.033	0.245	0.187	0.048	0.108	0.121
		COV	8.98	10.17	127.04	132.24	156.69	38.14	123.13	211.59	88.79	118.24	105.73	166.09
		Max.	35.00	8.78	330.50	402.50	341.20	0.093	0.359	0.594	0.579	0.109	0.504	0.630
		Min.	25.00	6.62	2.26	6.40	0.100	0.011	0.001	0.001	0.015	0.001	0.004	0.005
	Dry	Mean	25.29*	7.23	16.55*	25.40*	9.01*	0.046	0.039*	0.056	0.161*	0.056	0.099	0.055
		S.D.	3.16	1.82	15.25	16.92	8.50	0.021	0.035	0.041	0.199	0.049	0.165	0.102
		COV	12.50	25.23	92.14	66.60	94.27	46.11	89.68	72.66	123.38	86.82	166.45	186.20
		Max.	31.40	10.59	54.00	71.80	37.80	0.130	0.081	0.091	0.765	0.109	0.710	0.447
		Min.	18.90	3.15	2.035	3.60	1.625	0.004	0.004	0.001	0.004	0.002	0.002	0.001

Temp. – temperature (°C); Turb. – turbidity (NTU, Nephelometric Turbidity Units); color – apparent color; TSS – total suspended solids. *Means differ statistically according to the Kruskal-Wallis test at a probability of 0.001. Limits established for Class 2 freshwater by CONAMA 357/2005: pH from 6.0 to 9.0; turbidity up to 100 NTU; dissolved Cu of 0.009 mg L⁻¹; dissolved Fe of 0.3 mg L⁻¹; dissolved Al of 0.1 mg L⁻¹.

Although manganese has the highest concentration in the dam tailings, its concentration in water was not much increased in the affected sectors, compared with the local background the (upstream Carmo River). There, the maximum mean value for Mn was 0.09 mg L^{-1} , while affected sectors Mn reached 0.12 mg L^{-1} . The maximum concentration was found in the rainy season (0.989 mg L^{-1}), and the concentrations of this element generally increased in the rainy season.

The maximum value of dissolved Cd was 0.109 mg L^{-1} and the minimum was 0.002 mg L^{-1} , and did not differ between the dry and rainy seasons in all three sectors. However, a mean of 0.041 mg L^{-1} was determined in the downstream Carmo River sector in the rainy season.

The content of dissolved Zn in water was above 0.05 mg L^{-1} on average, with higher concentrations in the rainy period with a maximum of 0.710 mg L^{-1} detected in the downstream Carmo River sector in the dry period. Zn was positively correlated with the rainfall 10 and 20 days before the collection of the samples; that is, the increasing precipitation directly increased the concentration of dissolved Zn, either by sedimentary load or remobilization of riverbed sediments.

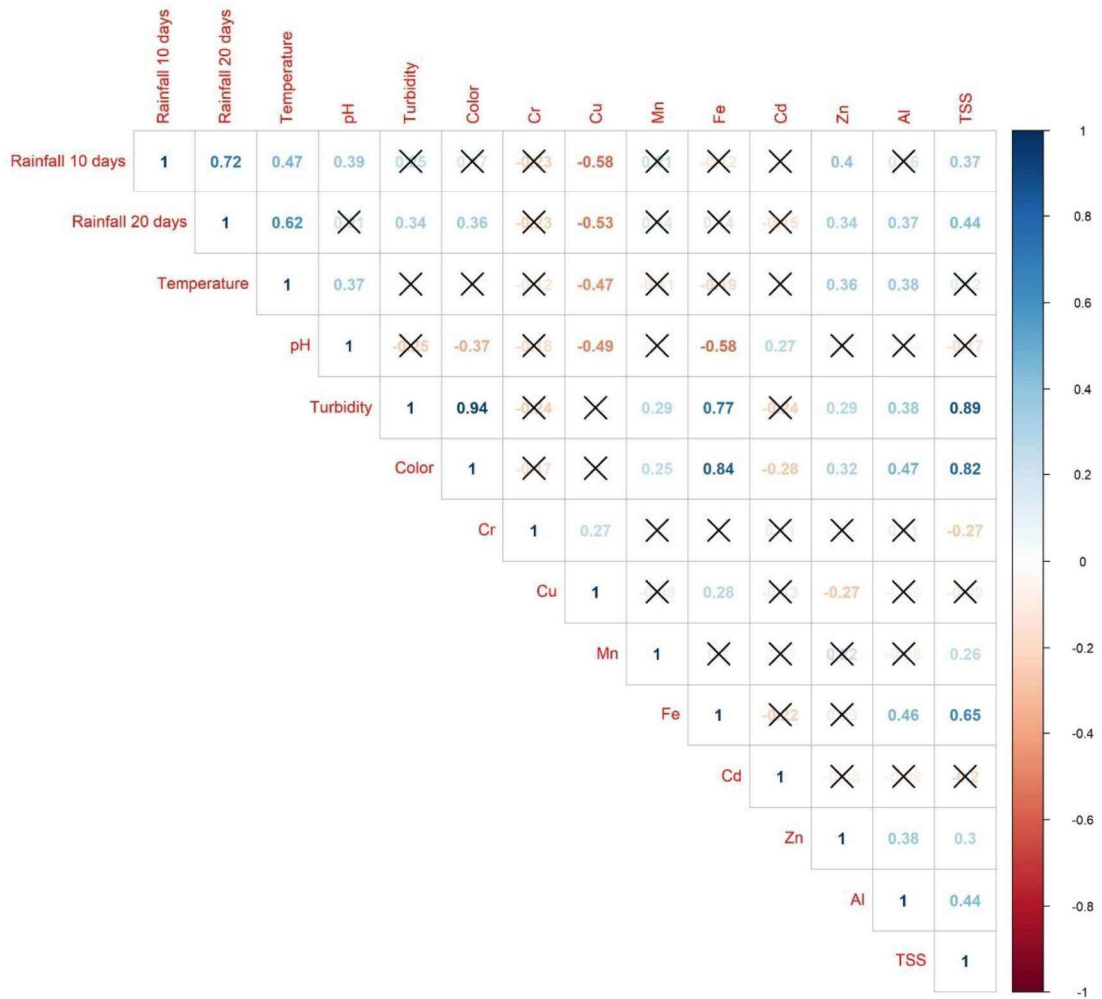


Figure 2: Spearman correlation of the analyzed variables. The X symbol distinguishes correlations that were not significant at 0.005 probability. Al – dissolved aluminum; Mn – dissolved manganese; Cd – dissolved cadmium; Zn – dissolved zinc; Cr – dissolved chromium; Fe – dissolved iron; pH – hydrogen ionic potential; Sample turbidity is measured in Nephelometric Turbidity Units; TSS – Total Suspended Solids; COR – apparent color of the sample; Sample temperature. The scale color represents the correlation between the variables, indicating that the closer to blue the more positively correlated the variables are, the closer to red, the more negatively correlated the variables are.

3.2 Multivariate analysis

Grouping all variables by the rivers sectors, and classifying them into the dry and rainy periods, it appears that in the rainy period 53.5% of the variance is explained by the sum of the first two dimensions, confirming the correlation of the variables Zn, Al, Mn, Fe, color, and turbidity in this cluster (Figure 3). For TSS, rainfall 10 days, rainfall 20 days (accumulated

precipitation 10 and 20 days prior to collection), temperature, Cu, Cd, and Cr, explain 34% of the variance in dimension 1, showing that the effect of the accumulated precipitation values 10 and 20 days prior to sample collection affects the temperature and dissolved amounts of Cu, Cd, and Cr. In this case, the ellipses are very similar, which indicates that the three sectors evaluated are similar from the point of view of the evaluated parameters, consistently with the descriptive analysis of the data.

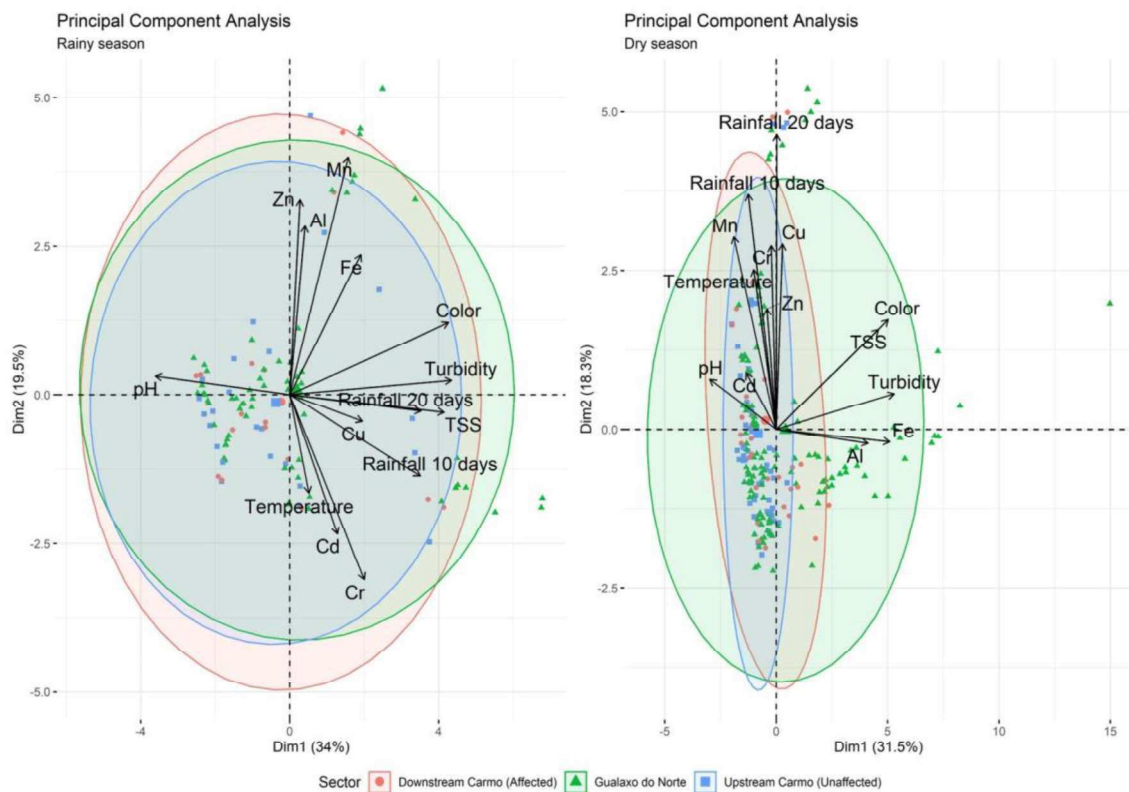


Figure 3: PCA for the rainy and dry season data evaluated from April 2017 to January 2019 grouped by River sectors.

In the dry season (Figure 3) the behavior of the upstream and downstream Carmo River sectors are similar, and distinguished the Gualaxo do Norte River. The variations of variables TSS, turbidity, Color, rainfall 10 days, and Cu are explained by dimensions 1 and 2. Fe and Al are explained by dimension 1, which distinguishes the Gualaxo do Norte River from the other sectors evaluated in the dry period. The other variables are explained by dimension 2, and for Cd and Zn, the vectors are smaller in both dimensions, showing lower load values. Two dimensions explain 49.8% of the variance.

When performing the PCA for the period from April to September 2016 (Figure 4), for the same rivers evaluated, but at different points, it was possible to verify the relationship

of the variables Cd and Cr, with their variances being better explained by dimension 2 than by dimension 1. We also observed high ratio values of some chemical parameters such as Fe, Cu, Al, and Mn with the physical variables of turbidity and color, explained by dimension 2.

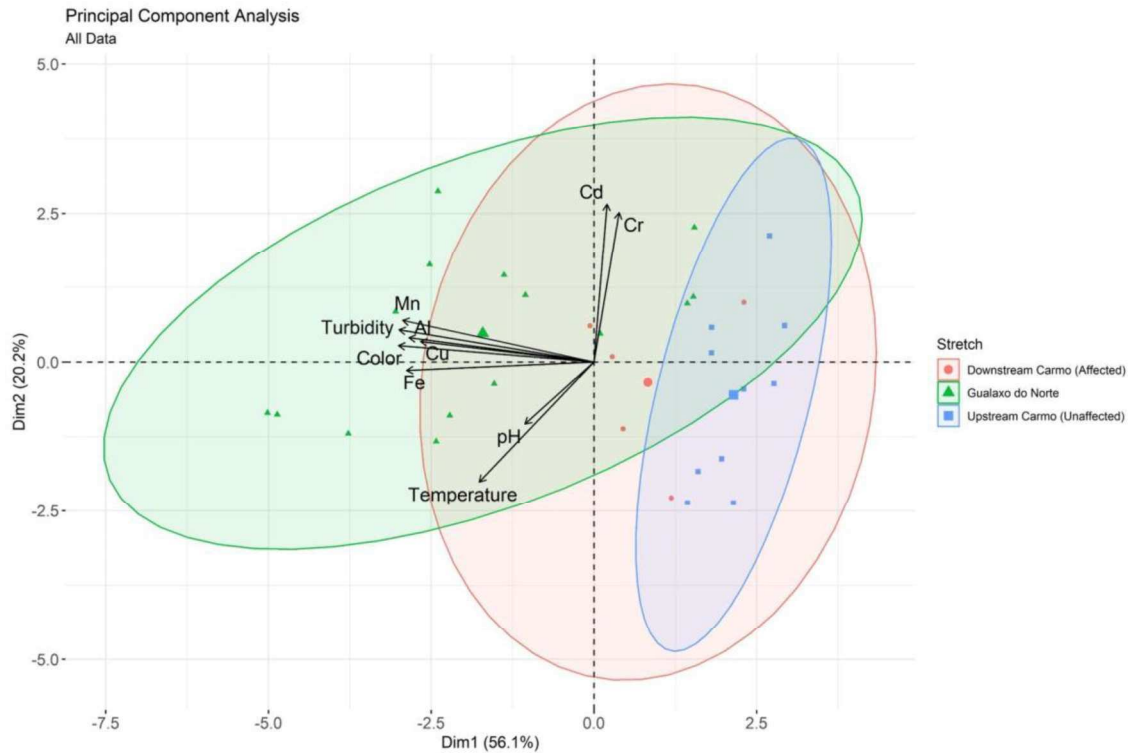


Figure 4: PCA for data collected in the dry period of the year 2016 by Foesch (2020) in the Gualaxo do Norte and Carmo Rivers.

4 DISCUSSION

The reduction in the concentration of Fe and Mn in the dry period is attributed to lower metal dilution effect, increasing the concentration of metal dissolved in water (Dornfeld 2002), and can be influenced by mud tailings from the Fundão dam, since it did not affect the upstream Carmo River sector (Golder Associates 2017). The dissolved Cd, on the other hand, had a high variability in the rainy period in the downstream Carmo River sector, which may be associated with the greater release of domestic effluents by the Barra Longa city (Ulmgren 2000; Appiah-Effah et al. 2015).

Similar temperature values were reported by Marques and Barbosa (2001), who evaluated the water temperature in the Doce River, and Silva et al. (2018), who evaluated trace elements in water and sediment before and after the Fundão dam failure in the periods of April 2015 and April 2016.

High water temperatures can influence the aquatic plant and animal life. Abowei (2010) reports that the temperature of tropical waters generally varies between 25 and 35 °C. Thus, it can be said that this parameter was not directly affected by the rupture and deposition of tailings. The pH values also varied with the season, but it is worth mentioning that pH is an important factor in dissolved metal concentration in water (Costa 2001).

Other physical variables, such as turbidity, TSS, and apparent color, showed high seasonal variability. Turbidity values above 120 NTU were recorded for the rainy season in the Gualaxo do Norte River, above the limit (100 NTU) established by CONAMA 357/2005 for class 2 waters. However, a peak of 659 NTU was observed in the Gualaxo do Norte River sector, consistent with Fernandes (2017). In both upstream and downstream Carmo River sectors, the average values were below this limit in both seasons.

Regarding TSS, a value of 646.75 mg L⁻¹ was recorded in the Gualaxo do Norte River in the rainy season, the maximum value found in the three sectors. CONAMA 357/2005 does not consider this parameter for water quality. Although there is a linear relation between TSS and turbidity, we observed lesser association in the dry season (Oliveira et al. 2018). Turbidity expresses, the apparent number of particles present in a water body, as determined by a beam of light that passes through the sample, quantifying the density (Toniolo 2016). Four classes are recognized: fairly turbid from 15 to 25 NTU, rather turbid from 25 to 35 NTU, turbid from 35 to 50 NTU, and very turbid above 50 NTU (Azis et al. 2015).

In recent study of Doce River, turbidity values above 400 NTU in the rainy period, and 250 NTU in the dry period were recorded by Oliveira et al. (2018), along with TSS values of 300 mg L⁻¹ in the rainy period and 200 mg L⁻¹ in the dry period. This highlights the natural high turbidity and TSS of the Doce River basin regardless of the Fundão dam contribution.

The three stretches have similar changes in physical parameters when comparing the measured mean values, however, they have different intensities. While in the Carmo Upstream stretch (unaffected), for example, in the rainy season Turbidity and TSS are 9 times higher than the dry season, in the affected stretches the intensity is lower, on average 3 times higher in the wet period when compared to the dry period.

In general, after precipitation variables such as Color, Turbidity and TSS are changed by increasing their values, this is because these variables are causally related to fine particles. Therefore, when precipitation occurs in the area increases the flow of the river allowing the entry of new sediments and the revolving of the bottom of the river. Even in tropical rivers that high physical variables (Color, Turbidity and TSS) are normal for a few months, changes in the concentration of these variables can occur due to rainfall, allowing the tailings deposited in the

margin to be taken to the river trough, favoring resuspension processes (Hatje et al 2017). In the dry period, due to the lower velocity of the rivers, the suspended particles tend to sediment, with this the variables Cor, Turbidity and TSS reduce (Gong et al 2016).

Historically, the water courses of the Doce River basin have high turbidity, as in the case of the Santarém stream, located closer to the Fundão dam, where values of 35.65 NTU were recorded in the rainy season. In the Gualaxo do Norte River, turbidity values were not higher than 43 NTU in the rainy period of 2000 (Costa 2001). However, at the automatic station located in Governador Valadares, an extreme peak of 140,000 NTU was recorded on 11 November of 2015 when the mud plume from the Fundão dam passed through that sector (Maia and Pereira 2017). This confirms the influence of the fine colloidal tailings dam Fundão released, downstream river, below the Risoleta Neves Hydroelectric Plant and dam (Milanez and Losekann 2016; Schaefer et al. 2016).

In the rainy season, there is a natural increase in turbidity, which is a direct contribution of eroded and transported resulting from severe soil loss and reworking of tailings deposited on the floodplain affected by mud (Theis et al. 1988). With the deposition of mud on floodplain, forming technosols (Schaefer et al. 2015). These areas became prone to erosion soil exposure, surface aluvic and little structure of this sediment. This contributed to several increases in turbidity and constant modifications of the fluvial channel by the lag deposition (Schaefer et al. 2015; Santos et al. 2017).

Apparent color values above 127.18 mg Pt L⁻¹ on were observed in the Gualaxo do Norte River sector, during the rainy season. In both downstream and upstream Carmo River sectors, the mean values were 110.15 and 86.08 mg Pt L⁻¹, respectively. These values are well above the maximum allowable value for the water potability standard, which is 15 mg Pt L⁻¹ according to consolidated Ordinance N°. 5 of the Health Ministry (Saúde 2017). The increasing in value of apparent color is associated high concentration of suspended particles suspended (> 1,2 µm), and closely associated with turbidity, although humic acids and the presence of biofilm can also change the water color.

The dissolved Cr showed values higher than those established in the CONAMA Resolution 357/2005 (Cr to 0.05 mg L⁻¹). Cr may be present in water as trivalent Cr or as hexavalent Cr, the later form being extremely toxic (Kaczynski and Kieber 1993), but there is no speciation study that would allow conclusions on toxicity to be drawn.

With refence to Cu, the mean concentrations of dissolved Cu for the two periods and in the three sectors evaluated were higher than those established for class 2 waters (0.009 mg

L⁻¹); an outlier of a dissolved Cu of 0.359 mg L⁻¹ was found in the upstream Carmo River sector (Table 1).

With regard to Fe, which is the element with the second highest concentration in the tailings, the waters of the Gualaxo do Norte River showed mean values above the limit imposed by the CONAMA resolution (0.3 mg L⁻¹) for class 2. The maximum value found was 3.013 mg L⁻¹, 10 times above this limit. Overall, in the sectors most directly affected by the mud tailings, the Fe concentration was nearly four times higher than the concentration of unaffected sector. Also, higher concentrations of Fe were observed in the dry season, probably because, in the dry period, the waters have lower pH values. According to Nordstrom et al. (2009), the Fe solubility increases with lower pH values, and correlation analysis (Figure 2) showed that the pH was negatively related to the values of the Fe metal, suggesting that acidity promoted Fe hydrolysis and release (Guedes et al. 2004), besides the diluting effect.

High Cd concentration in waters can be harmful to plants (Barceló et al. 1986) or human health (Silva et al. 2019). Cd remains in the environment for 10 to 30 years, and possesses slow excretion by humans (Nordberg 1996). On average, the three sectors had dissolved Cd, concentrations above the limit for the total Cd (0.001 mg L⁻¹) (CONAMA 2005) for class 2. In the dry season, the average concentration of dissolved Cd was lower in the downstream Carmo River sector.

Silva et al. (2018) found values of total Cd in the range of 0.65 µg L⁻¹ at in the Gualaxo do Norte River in 2016, shortly after the Fundão dam collapse, whereas Costa (2001) found values below 0.001 mg L⁻¹ in the Gualaxo do Norte River sector in the dry and rainy periods of 2000 (before the dam burst).

Average values of dissolved Zn were below the limit established by CONAMA 357/2005 for total Zn, which is 0.18 mg L⁻¹. Maximum values of total Zn of 0.077 mg L⁻¹ were observed in the same Carmo River sector (Silva et al. 2018). Abnormal Zn concentrations were considered natural anomalies by Rodrigues et al. (2013), who detected concentrations between 28.7 and 85.6 mg kg⁻¹ in sediments from the Gualaxo do Norte River, before the dam failure.

Regarding dissolved aluminum, the average concentrations in the three sectors and periods were within the allowable range for class 2 freshwater, according to CONAMA 357/2005 (Table 1). However, in the rainy season, an outlier 1.388 mg L⁻¹ was recorded in the Gualaxo do Norte River sector. Al is generally present in water in the form of Al³⁺ (aq) under acid conditions, and as Al(OH) (aq.) in neutral or alkaline conditions. In acidic waters, Al is toxic to fish from 0.1 mg L⁻¹ (Gensemer and Playle 1999; Ferreira et al. 2020). In all sectors and periods, the maximum values were above this limit.

Turbidity, color, and TSS are directly and highly correlated with each other and with the Fe content, which can be explained by the fact that Fe-oxides forms gives a reddish color to the suspended sediments present in waters (Moruzzi and Reali 2012), and it is the main element present in the tailings of Fundão.

The pH behavior in monitored period is consistent with that found by Costa (2001), and is poorly explained by both dimensions 1 and 2. In the period evaluated about eight months after the rupture of the Fundão dam, Mariana, MG, it is possible to distinguish three River sectors: Gualaxo do Norte, downstream and upstream Carmo River sectors.

In the dry period of 2017, two years after the accident, the Gualaxo do Norte River sector, closer to the source of tailings from dam, had turbidity, color, TSS, and Fe contents with vectors in the same direction, indicating a close relationship between these parameters in the dry period, unlike the upstream Carmo River sector which did not receive the same load from the dam rupture. The downstream Carmo River, after the confluence with the Gualaxo do Norte River, showed from the influence of tailings, but lower than the Gualaxo do Norte itself (Figure 5).

When separating the rivers by dry and rainy periods over the three-year period from 2017 to 2019, the physical and chemical parameters with greatest correlation (Figure 2), the change grouping of rivers and the dimensions of explanatory variables (Figure 5).

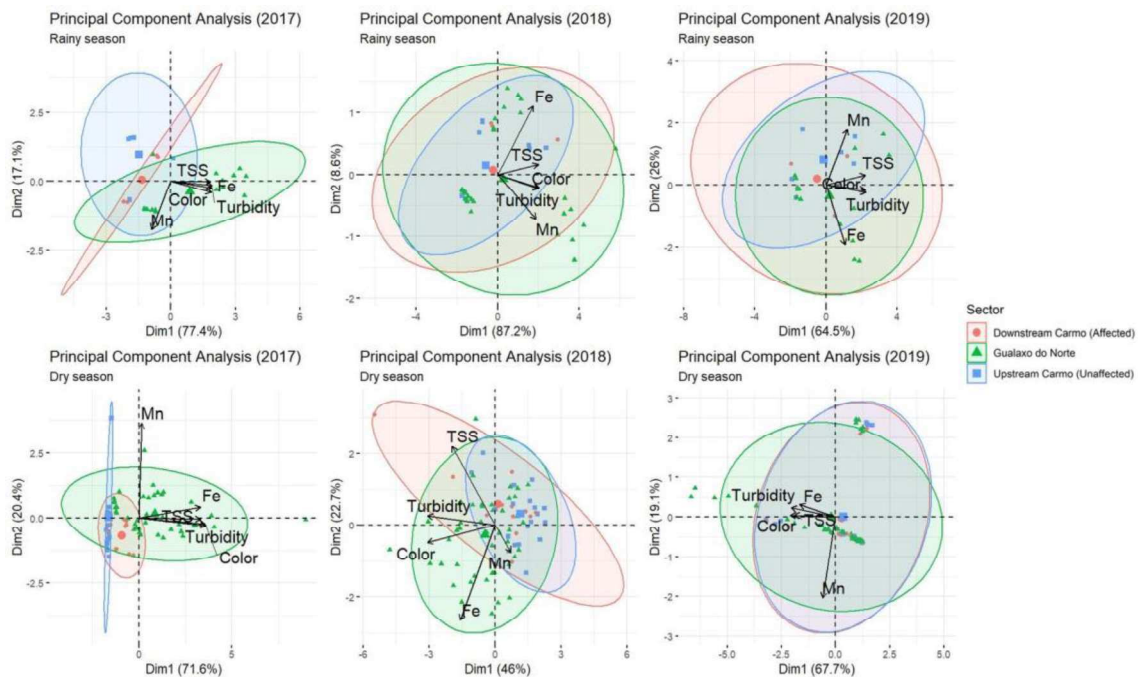


Figure 5: PCA of data collected in the rainy and dry period of the years 2017 to 2019 for only the variables evaluated throughout the period.

Three years after the dam burst, improvement of water quality at unlike the Carmo River sector can be observed, the behavior of the Gualaxo do Norte, which is distorted affected by Fe, color, and to a lesser extent, turbidity (Figure 5).

In the dry period in 2019, four years after the dam burst, it is possible to see a similar quality of waters upstream and downstream the Carmo River (Figure 5). Closer to the dam, the Gualaxo do Norte River still shows abnormal Fe contents in the dry season.

In the rainy season in 2017, the Gualaxo do Norte River was the most affected by the TSS, turbidity, color, and Fe variables, explaining 77.4% of the variance. One year after (rainy period 2018), the discrepancy was smoothed, with the turbidity, color and Mn explaining 87.2% in dimension 1. In the following rainy period (2019), differences remained smooth but, the Fe content was associated with turbidity and color. Here, four years after the rupture, the influence of Mn on the parameters of turbidity and color was reduced.

The improvement in physical and chemical parameters highlights the ability of tropical rivers affected by mining tailings to naturally recover. After implementation of recovery activities at the affected areas, the levels of some parameters were reduced.

On the other hand, the variation on some parameters is strongly influenced by the seasonal accumulated precipitation (Figure 6). For example, the mean concentration of dissolved Fe was higher in Gualaxo do Norte River at the beginning of the monitoring, while in the Carmo River sectors peaks of dissolved Fe and Mn occurred just after precipitation events, at the end of the first year of monitoring period.

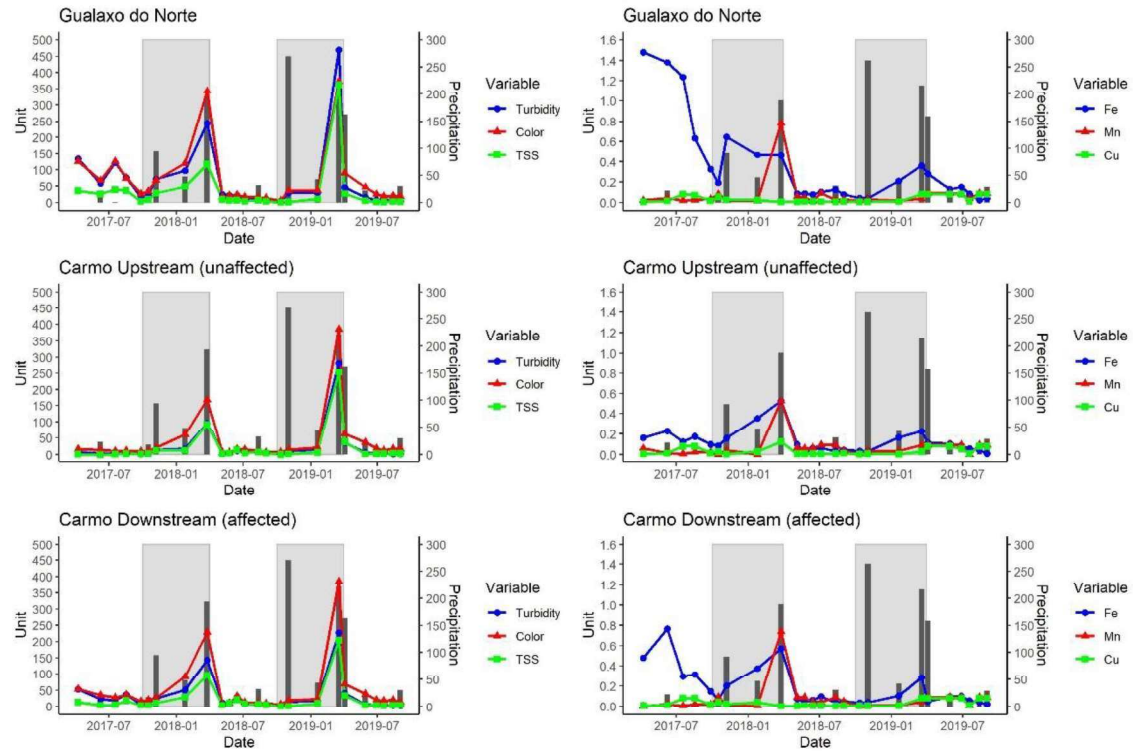


Figure 6: Time series for averages of the physical and chemical variables analyzed six months after consolidation of the recovery services. Precipitation – precipitation accumulated 20 days before sampling (mm); Fe – dissolved iron; Mn – dissolved manganese; Cu – dissolved copper (mg L^{-1}); NTU – turbidity (Nephelometric Turbidity Units); color – apparent color (mg Pl L^{-1}); TSS – Total Suspended Solids (mg L^{-1}). Gray stripes distinguish the rainy season from the dry season. The first column shows the physical variables in relation to precipitation and the second column shows the chemical variables in relation to precipitation.

Rainfall influence pH variability (Carvalho et al. 2000; Silva et al. 2008; Girardi et al. 2016), and on average the value increased by 0.5 points in the rainy period when compared to the dry period. However, in a limestone mining area, where CaCO_3 contribute to pH changes increasing pH by the dissolution of carbonates occurs in the dry season, when dilution is reduced (Fritzsons et al. 2009). The pH variation influences other physic-chemical variables (Yabe and Oliveira 1998; Girardi et al. 2016), and greater acidity leads to metal solubility, especially Cu, Mn and Fe (Yabe and Oliveira 1998). However, with greater rainfall, increasing concentrations Pb, Cd, Ni, Cr, and Cu were detected, attributed to soil erosion. Although Fe and Al showed reduced concentration after the implementation recovery activities, peaks in turbidity above 100 NTU continued recorded after the measures.

High, abnormal values of dissolved Fe were recorded before 2015 as contents above 2 mg L^{-1} were detected in 1999. In this stretch, even with the rupture of the Fundão dam, such values were not observed. After the rupture, the highest Fe content was 0.75 mg L^{-1} , and Fe decreased significantly after the recovery activities. Overall, the concentration of Fe increase in

dry season with lower water volume (Correll et al. 1982), and occurred either before and after the Fundão dam burst.

The values of dissolved Al peaked in 2016, but decreased after recovery activities, reaching contents similar to pre-rupture values.

Turbidity also decreased after recovery activities, with seasonal peaks observed in early 2018 and early 2019. After the recovery activities, the turbidity values followed the same seasonal trend comparing the three sectors, with an overall reduction due to the implementation of vegetation at the recovery the margins and backfill.

5 CONCLUSIONS

The presence a sedimentary lag of tailings on the riverbed of the Gualaxo-Carmo rivers, affected by the Fundão dam burst, has no significant effect on the concentration of dissolved metals in water.

The physical parameters monitored (turbidity, color, and TSS) showed great seasonal variability, closely associated with rainfall.

After the Fundão dam burst, turbidity in the Gualaxo do Norte-Carmo River reached values unrecorded previously, and the Gualaxo do Norte had, turbidity values five times higher than that required for class 2 rivers. On average, the turbidity values were generally higher than those established for River classes 1 to 3 in the Gualaxo do Norte River, and reduced in the Carmo River sector, downstream.

REFERENCES

- ANM, A. N. de M. (2017). Sumário mineral 2017 (ano base 2016). *Agência Nacional de Mineração*, 37, 201 p. http://www.anm.gov.br/dnpm/publicacoes/serie-estatisticas-e-economia-mineral/sumario-mineral/sumario-brasileiro-mineral-2017/aco_sm_2017
- APHA, A. P. H. A. (2005). *Standard Methods for the Examination of Water and Wastewater*. (L. S. Clesceri, A. E. Greenberg, & A. D. Eaton, Eds.) (21st ed.). Washington, DC.
- Appiah-Effah, E., Nyarko, K. B., Antwi, E. O., & Awuah, E. (2015). Heavy metals and microbial loads in raw fecal sludge from low income areas of ashanti region of ghana. *Water Practice and Technology*, 10(1), 124–132. <https://doi.org/10.2166/wpt.2015.014>
- Barceló, J., Poschenrieder, C., Andreu, I., & Gunsé, B. (1986). Cadmium-Induced Decrease of Water Stress Resistance in Bush Bean Plants (*Phaseolus vulgaris* L. cv. Contender) I. Effects of Cd on Water Potential, Relative Water Content, and Cell Wall Elasticity. *Journal of Plant Physiology*, 125(1–2), 17–25. [https://doi.org/10.1016/S0176-1617\(86\)80239-5](https://doi.org/10.1016/S0176-1617(86)80239-5)
- Brasil (2017). *Resolução RDC nº 166, de 25 de julho de 2017*.
- Carvalho, A. R., Schlittler, F. H. M., & Tornisielo, V. L. (2000). Relações da atividade agropecuária com parâmetros físicos químicos da água. *Química Nova*, 23(5), 618–622. <https://doi.org/10.1590/s0100-40422000000500009>
- Coimbra, K. T. O., Alcântara, E., & de Souza Filho, C. R. (2019). An assessment of natural and manmade hazard effects on the underwater light field of the Doce River continental shelf. *Science of the Total Environment*, 685, 1087–1096. <https://doi.org/10.1016/j.scitotenv.2019.06.127>
- CONAMA (2005). *Resolução nº 357, 18 de março de 2005*.
- Correll, D. L., Goff, N. M., & Peterjohn, W. T. (1982). Ion Balances Between Precipitation Inputs and Rhode River Watershed Discharges. In *Geological aspects of acis deposition* (7th ed., pp. 77–111). Las Vegas, NV.
- Costa, A. T. (2001). *Geoquímica das águas e dos sedimentos da bacia do rio Gualaxo, Leste-Sudeste do Quadrilátero Ferrífero (MG): Estudo de uma área afetada por atividades de extração mineral*. Universidade Federal de Ouro Preto.
- DNPM, D. N. de P. M. (2019). *Anuário Mineral Brasileiro: Principais Substâncias Metálicas - 2018. Anuário Mineral Brasileiro*. Brasília, DF. http://www.anm.gov.br/dnpm/publicacoes/serie-estatisticas-e-economia-mineral/anuario-mineral/anuario-mineral-brasileiro/amb_2018.pdf

- Dornfeld, C. B. (2002). *Utilização de análises limnológicas, bioensaios de toxicidade e macroinvertebrados bentônicos para o diagnóstico ambiental do reservatório de Salto Grande (Americana, SP)*. USP.
- EPA. (1992). *Method 3005A - Acid digestion of waters for total recoverable or dissolved metals for analysis by FLAA or ICP spectroscopy*.
- Fernandes, K. N. (2017). *Qualidade das águas nos rios Gualaxo do Norte, Gualaxo do Sul e do Carmo, afluentes do alto rio Doce (Watu)*. Universidade Federal de Ouro Preto.
- Ferreira, F. F., de Freitas, M. B. D., Szinwelski, N., Vicente, N., Medeiros, L. C. C., Schaefer, C. E. G. R., et al. (2020). Impacts of the Samarco Tailing Dam Collapse on Metals and Arsenic Concentration in Freshwater Fish Muscle from Doce River, Southeastern Brazil. *Integrated Environmental Assessment and Management*, 16(5), 622–630. <https://doi.org/10.1002/ieam.4289>
- Foesch, M. D. S., Francelino, M. R., Rocha, P. A., & Gomes, A. R. L. (2020). River Water Contamination Resulting from the Mariana Disaster, Brazil. *Floresta e Ambiente*, 27(4), 2–10. <https://doi.org/10.1590/2179-8087.013218>
- Fritzsons, E., Mantovani, L. E., Chaves Neto, A., & Hindi, E. C. (2009). A influência das atividades mineradoras na alteração do pH e da alcalinidade em águas fluviais: o exemplo do rio Capivari, região do carste paranaense. *Engenharia Sanitaria e Ambiental*, 14(3), 381–390. <https://doi.org/10.1590/s1413-41522009000300012>
- Gensemer, R. W., & Playle, R. C. (1999). The bioavailability and toxicity of aluminum in aquatic environments. *Critical Reviews in Environmental Science and Technology*, 29(4), 315–450. <https://doi.org/10.1080/10643389991259245>
- Ghose, M. K., & Sen, P. K. (2001). Characteristics of Iron Ore Tailing Slime in India and. *Environmental Monitoring and Assessment*, 68, 51–61.
- Girardi, R., Pinheiro, A., Garbossa, L. H. P., & Torres, É. (2016). Water quality change of rivers during rainy events in a watershed with different land uses in Southern Brazil. *RBRH*, 21(3), 514–524. <https://doi.org/10.1590/2318-0331.011615179>
- Golder Associates. (2017). *Programa de Monitoramento Quali-Quantitativo Sistema de Água e Sedimentos*.
- Gong, Y., Liang, X., Li, X., Li, J., Fang, X., & Song, R. (2016). Influence of rainfall characteristics on total suspended solids in urban runoff: A case study in Beijing, China. *Water (Switzerland)*, 8(7). <https://doi.org/10.3390/w8070278>

- González, A. G., & Herrador, M. Á. (2007). A practical guide to analytical method validation, including measurement uncertainty and accuracy profiles. *TrAC - Trends in Analytical Chemistry*, 26(3), 227–238. <https://doi.org/10.1016/j.trac.2007.01.009>
- Guedes, C. D., Pereira, J. G., Lena, J. C. de, Paiva, J. F. de, & Lima, R. M. F. (2004). Coagulação/floculação de suspensões ricas em óxidos de ferro por sulfato de alumínio. *Química Nova*, 27(5), 715–719. <https://doi.org/10.1590/s0100-40422004000500007>
- Hatje, V., Pedreira, R. M. A., De Rezende, C. E., Schettini, C. A. F., De Souza, G. C., Marin, D. C., & Hackspacher, P. C. (2017). The environmental impacts of one of the largest tailing dam failures worldwide. *Scientific Reports*, 7(1), 1–13. <https://doi.org/10.1038/s41598-017-11143-x>
- Kaczynski, S. E., & Kieber, R. J. (1993). Aqueous trivalent chromium photoproduction in natural waters. *Environmental Science & Technology*, 27(8), 1572–1576. <https://doi.org/10.1021/es00045a011>
- Maia, L. C., & Pereira, A. R. (2017). Impactos no abastecimento de água decorrentes do rompimento da barragem de fundão: estudo de caso de Governador Valadares. In *Congresso ABES* (pp. 1–6). <https://www.tratamentodeagua.com.br/artigo/impactos-no-abastecimento-rompimento-barragem/>
- Marques, M. M., & Barbosa, F. (2001). Biological quality of waters from an impacted tropical watershed (middle Rio Doce basin, southeast Brazil), using benthic macroinvertebrate communities as an indicator. *Hydrobiologia*, 457(1997), 69–76. <https://doi.org/10.1023/A:1012297915323>
- Marta-Almeida, M., Mendes, R., Amorim, F. N., Cirano, M., & Dias, J. M. (2016). Fundão Dam collapse: Oceanic dispersion of River Doce after the greatest Brazilian environmental accident. *Marine Pollution Bulletin*, 112(1–2), 359–364. <https://doi.org/10.1016/j.marpolbul.2016.07.039>
- Mechi, A., & Sanches, D. L. (2010). Impactos ambientais da mineração no estado de São Paulo. *Estudos Avançados*, 24(68), 209–220.
- Milanez, B., & Losekann, C. (2016). Desastre no Vale do Rio Doce: Antecedentes, impactos e ações sobre a destruição. *Folio Digital*.
- Moruzzi, R. B., & Reali, M. A. P. (2012). Oxidação E Remoção De Ferro E Manganês Em Águas Para Fins De Abastecimento Público Ou Industrial– Uma Abordagem Geral. *Revista De Engenharia E Tecnologia*, 4(1), 29–43.
- Nordberg, G. F. (1996). Current issues in low-dose cadmium toxicology: nephrotoxicity and carcinogenicity. *Environ. Sci*, 4, 133–147.

- Nordstrom, D. K. A., Jenne, E. W., & Ball, J. (2009). Redox Equilibria of Iron in Acid Mine Waters, In: ACS Symposium Series, 51–79. <https://doi.org/10.1021/bk-1979-0093.ch003>
- Oliveira, A. R. M. De, Borges, A. C., Matos, A. T., & Nascimento, M. (2018). Estimation on the concentration of suspended solids from turbidity in the water of two sub-basins in the Doce river basin. *Engenharia Agrícola*, 38(5), 751–759.
- Pedreira, G., & Sousa, H. C. de. (2011). Comunidade arbórea de uma mancha florestal permanentemente alagada e de sua vegetação adjacente em Ouro Preto-MG, Brasil. *Ciencia Florestal*, 21(4), 663–675. <https://doi.org/10.5902/198050984511>
- Quaresma, L. F. (2009). Projeto de assistência técnica ao setor de energia - Perfil da Mineração de Ferro. Ministério de Minas - MME, 4–56.
- Rao, G. V., Markandeya, R., & Sharma, S. K. (2016). Recovery of Iron Values from Iron Ore Slimes of Donimalai Tailing Dam. *Transactions of the Indian Institute of Metals*, 69(1), 143–150. <https://doi.org/10.1007/s12666-015-0809-0>
- Rodrigues, A. S. de L., Malafaia, G., Costa, A. T., & Nalini Júnior, H. A. (2013). Background Values for Chemical Elements in Sediments of the Gualaxo Do Norte River Basin, Mg, Brazil. *Revista de Ciências Ambientais*, 7(2), 15–32. <https://doi.org/10.18316/1142>
- Renova, F. (2019). Relatório Anual de Atividades. Mariana, MG.
- Santos, J. P. dos, Cota, G. E. M., Limoeiro, B. F., Pedras, K. C., Costa, A. M. da, & Viana, J. H. M. (2017). Susceptibilidade à erosão no Rio Gualaxo do Norte-MG. *Caminhos de Geografia*, 18(63), 286–307. <https://doi.org/10.14393/RCG186313>
- Saúde, M. de E. da. (2017). Portaria de Consolidação no 5 de 28 de setembro de 2017.
- Schaefer, C. E. G. R., Santos, E. E. dos, Fernandes Filho, E. I., & Assis, I. R. de. (2016). Paisagens de Lama: Os Tecnoossolos para recuperação ambiental de áreas afetadas pelo desastre da barragem do Fundão, em Mariana. *Boletim informativo da SBCS*, 1(1), 18–23.
- Schaefer, C. E. G. R., Santos, E. E. dos, Souza, C. M. de, Neto, J. D., Filho, E. I. F., & Delpupo, C. (2015). Cenário histórico, quadro fisiográfico e estratégias para recuperação ambiental de Tecnoossolos nas áreas afetadas pelo rompimento da barragem do Fundão, Mariana, MG. *Arquivos do Museu de História Natural e Jardim Botânico - UFMG*, 24(1/2), 104–135.
- Silva, A. C., Cavalcante, L. C. D., Fabris, J. D., Júnior, R. F., Barral, U. M., Farnezi, M. M. de M., et al. (2016). Chemical, mineralogical and physical characteristics of a material accumulated on the river margin from mud flowing from the collapse of the iron ore tailings dam in Bento Rodrigues, Minas Gerais, Brazil. *Revista Espinhaço*, 5(2), 44–53.

- Silva, A. E. P., Angelis, C. F., Machado, L. A. T., & Waichaman, A. V. (2008). Influência da precipitação na qualidade da água do Rio Purus. *Acta Amazonica*, 38(4), 733–742. <https://doi.org/10.1590/S0044-59672008000400017>
- Silva, D. de C., Bellato, C. R., Marques Neto, J. de O., & Fontes, M. P. F. (2018). Trace elements in river waters and sediments before and after a mining dam breach (Bento Rodrigues, Brazil). *Química Nova*, 41(8), 857–866. <https://doi.org/10.21577/0100-4042.20170252>
- Silva, S. K. D. B. da, Santos, J. A. dos, Santos, F. M. dos, & Flores, C. M. (2019). Intoxicação por metais pesados. In *Jornada Científica dos Campos Gerais* (p. 4). Ponta Grossa, PR.
- SUPRAM, S. R. de M. A. e D. S. (2008). Parecer Único SUPRAM - ZM Processo(s) No: 00015/1984/066/2008, 18.
- Theis, T. L., Young, T. C., & DePinto, J. V. (1988). Factors affecting metal partitioning during resuspension of sediments from the Detroit River. *Journal of Great Lakes Research*, 14(2), 216–226. [https://doi.org/10.1016/S0380-1330\(88\)71550-6](https://doi.org/10.1016/S0380-1330(88)71550-6)
- Toniolo, G. R. (2016). Identificação de constituintes opticamente ativos na água do lago Guaíba, a partir de dados de sensores orbitais e espectrorradiometria de campo. Universidade Federal do Rio Grande do Sul.
- Ulmgren, L. (2000). Measures taken in smaller industries to avoid hazardous substances entering domestic wastewater systems. In Stockholm Water Company.
- Yabe, M. J. S., & Oliveira, E. de. (1998). Metais pesados em águas superficiais como estratégia de caracterização de bacias hidrográficas. *Química Nova*, 21(5), 551–556. <https://doi.org/10.1590/s0100-40421998000500003>

Capítulo 2 - Sensoriamento Remoto aplicado a modelagem de águas interiores

RESUMO

Águas interiores são um desafio para a aplicação do sensoriamento remoto e obtenção de resultados promissores, muito devido a resolução espacial das imagens disponibilizadas gratuitamente, impedindo ter a informação correta do pixel que representa o leito dos rios. Entretanto, como o avanço do sensoriamento remoto, facilitou a sua utilização em diversas áreas do conhecimento, sendo cada vez mais comum programas espaciais que disponibilizam imagens orbitais gratuitamente, mesmo aquelas com alta resolução espacial, como as imagens do sensor WPM (*Wide scan Panchromatic and Multispectral camera*) do satélite CBERS-4A, com 2 m, como aquelas com alta resolução temporal, como as imagens do sensor MSI (*Multispectral Instrument*) acoplado ao satélite *Sentinel 2*. Logo objetivo geral foi modelar a turbidez do rio Doce utilizando técnicas de sensoriamento remoto. Predizer variáveis físicas e químicas a partir de valores de turbidez verificadas em campo, bem como modelar as variáveis previstas pela turbidez de forma contínua. Para tanto foram utilizados dados de sensor MSI e de turbidez de 13 estações de monitoramento ao longo da bacia do rio Doce, entre agosto de 2017 e dezembro de 2020. Os dados foram tratados antes da modelagem, realizando a remoção de dados espúrios e outliers, obtendo um total de 78 observações. O modelo gerado para Turbidez mostra que essa variável pode ser mensurada a partir de dados de sensores remotos com precisão de 0,6, podendo ser apresentada de forma contínua em todo o trecho do rio Doce e do rio Carmo de Barra Longa/MG até sua foz em Rio Doce/MG. A regressão linear do SST e do Fe total foram superiores a 0,80 de R^2 mostrando a correlação linear entre as variáveis. A cor verdadeira e o Al total não tiveram bom ajuste linear com a turbidez, mesmo assim na modelagem tiveram resultados superiores a 0,5 de precisão. A modelagem contínua das variáveis avaliadas no trabalho pode ser utilizada para avaliar, a qualidade e a quantidade dessas variáveis ao longo de todo curso d'água, em especial em áreas onde não possui pontos de monitoramento.

Palavras-chave: Rio Doce; Sentinel-2; *machine learning*; Barragem de Fundão.

1 INTRODUÇÃO

Na tentativa de reconstruir digitalmente determinado parâmetro da superfície se faz uso de diferentes formas de representação, a qual será definida pela continuidade ou não desse parâmetro no meio a qual se encontra. Estudos relacionados à qualidade de água geralmente são determinados em coleta pontual, entretanto, o dado, coletado representa o *status quo* naquele momento e no local da coleta. Quando se trata de um rio, a diluição e a turbulência afetam a aleatoriedade de parâmetros avaliados. Ou seja, a representação contínua dos dados nesse tipo de situação é a melhor das opções. Dessa forma várias tentativas vem sendo realizado para estabelecer a relação entre dados de sensoriamento remoto com dados de qualidade da água (Isidro et al., 2018; Kilham et al., 2012).

O avanço do sensoriamento remoto, facilitou a sua utilização em diversas áreas do conhecimento (Asiken Alias and Mansor, 2018; Bousbih et al., 2019; Gomes et al., 2019; Rosa, 2018), e é cada vez mais comum programas espaciais que disponibilizam imagens orbitais gratuitamente, mesmo aquelas com alta resolução espacial, como as imagens do sensor WPM (*Wide scan Panchromatic and Multispectral camera*) do satélite CBERS-4A, com 2 m, como aquelas com alta resolução temporal, como as imagens do sensor MSI (*Multispectral Instrument*) acoplado ao satélite *Sentinel 2*. Esse sistema é dividido em dois satélites faseados entre si a 180°, que permitem tempo de revisita de cinco dias.

Águas interiores representam um desafio para a aplicação do sensoriamento remoto e obtenção de resultados promissores, muito devido a resolução espacial das imagens disponibilizadas gratuitamente, impedindo a aquisição da informação correta do pixel que representa o leito dos rios (Barbosa et al., 2019). Isso porque, muitas das vezes a largura dos rios é menor do que a resolução do sensor, fazendo com que não existam pixels puros de água.

As imagens do *Sentinel 2* podem ser limitadas na aplicação direta em cursos d'água com largura inferior a resolução espacial da imagem, pois parte do pixel poderá não corresponder fielmente a resposta espectral da água, sendo influenciada pelas suas margens. Porém, podem ser interessantes na aplicação cursos d'água com largura maior que 10 m, em lagoas e regiões costeiras (Elhag et al., 2019; Oliveira et al., 2012).

Outra situação se diz respeito a interação da radiação eletromagnética com a coluna d'água e seus constituintes, absorvendo e refletindo-a, gerando então os componentes opticamente ativos responsáveis pelas diferentes assinaturas espectrais da água (Watanabe et al., 2010). Com o rompimento da barragem de Fundão, cerca de 44 Mm³ de rejeitos de mineração de ferro foram liberados nos rios Gualaxo do Norte, Carmo e Doce, sendo que,

segundo Santos et al. (2021), aproximadamente 35% destes foram depositados nos terraços do rio Gualaxo do Norte, sendo essa a área mais afetada. Outros 10 Mm³ ficaram retidos na barragem da Usina hidrelétrica Risoleta Neves (Saadi and Campos, 2015), sendo que o material mais fino, foi transportado até a foz do rio Doce com o Oceano Atlântico (Coimbra et al., 2019).

Mesmo após cinco anos do acidente, eventos de chuvas mais intensas resuspendem os sedimentos e provocam também o carreamento do material depositado no fundo dos leitos do rio Carmo e Gualaxo do Norte, a montante da Barragem Risoleta Neves, alterando as características dos rios sazonalmente (Nascimento et al., 2019).

Com isso um modelo capaz de prever o comportamento da qualidade da água de forma contínua e de baixo custo seria ideal para a manutenção do monitoramento da bacia hidrográfica do rio Doce. Sendo assim, o objetivo desse capítulo é de modelar a turbidez do rio Doce utilizando técnicas de sensoriamento remoto. Como objetivos específicos analisar as variáveis Cor, SST, Fe e Al a partir de valores de turbidez verificadas em campo, bem como modelá-las de forma contínua.

2 MATERIAL E MÉTODOS

A Fundação Renova responsável pela recuperação das áreas afetadas pelo rompimento da barragem de Fundão, instalou diversas estações automáticas de coleta de dados de qualidade de água ao longo do trecho afetado pelo, com liberados a partir de agosto de 2017. Com isso, foi possível utilizar dados de 13 estações automáticas que possuem sensores de turbidez (Figura 7).

realizados entre agosto de 2017 e dezembro de 2020. Os dados foram tratados antes da modelagem, realizando a remoção de dados espúrios e outliers, análise dos valores fora da média, obtendo um total de 225 observações do total de 302 para a regressão dos dados.

Na modelagem dos dados a partir da Turbidez foram utilizadas 78 observações do período de agosto de 2017 a agosto de 2019, em que a as imagens não apresentavam nuvens em toda a área da bacia. Foram realizados ensaios para verificar a interferência de resolução espacial no processo de modelagem, para tal foram removidos pontos amostrais nos quais a largura do rio fosse menor que três vezes a resolução da imagem. Dessa forma foram realizados ensaios utilizando 48 amostras para treinamento e 6 para validação, em um exemplo em que removidos pontos pertencentes a foz entre o rio Gualaxo do Norte e o rio Carmo. Na Tabela 2 pode ser verificado a localização do ponto e a largura do rio onde foi verificado os valores de Turbidez.

Tabela 2: Coordenadas e largura da cessão no ponto onde estão localizados os turbidimetro do Programa de Monitoramento Quali-Quantitativo

Ponto	Largura (m)	Coordenada		Ponto	Largura (m)	Coordenada	
		Latitude	Longitude			Latitude	Longitude
RCA-01	27	-20,3471	-43,1127	RDO-08	294	-18,8833	-41,9526
RCA-02	34	-20,2824	-43,0449	RDO-12	222	-19,4993	-40,7587
RDO-01	53	-20,2560	-42,9125	RDO-16	984	-19,6464	-39,8232
RDO-02	40	-20,2087	-42,8509	RGN-01	15	-20,2850	-43,4772
RDO-03	40	-20,0144	-42,7446	RGN-06	19	-20,3037	-43,2495
RDO-04	41	-19,5542	-42,5214	RGN-08	16	-20,2861	-43,0658
RDO-05	183	-19,3207	-42,3646	Média	151		

Coordenadas DATUM SIRGAS 2000. RCA - pontos localizados no rio do Carmo, RGN - pontos localizados no rio Gualaxo do Norte, e RDO - pontos localizados no rio Doce.

As imagens obtidas pelo sensor MSI do sistema orbital *Sentinel 2*, foram escolhidas em datas nas quais a imagem sobre o ponto amostral não tivesse coberta por nuvens. Sendo assim foram utilizadas 42 imagens do sensor MSI, no período de agosto de 2017 a agosto de 2019.

2.1 Sensoriamento Remoto

Foram utilizadas imagens Nível 1 C e Nível 2 A do sensor MSI (*Multi Spectral Instrument*) acoplado ao sistema orbital *Sentinel 2*. As imagens de Nível 1 C foram corrigidas para reflectância de superfície no software SNAP utilizando o algoritmo Sen2Cor produzido

pela European Space Agency (ESA). Já as imagens Nível 2 A, possuem reflectância corrigida abaixo da atmosférica (BOA) derivada do produto Nível 1 C. Posteriormente, a coleta dos valores de reflectância nos pontos de coleta de amostra de água foram realizadas no software R. Foram também coletados dados nos dois pixels vizinhos ao ponto, elevando o número de amostras. Relações de bandas foram estabelecidas para os valores de reflectância nas bandas de resolução de 10 m do instrumento MSI. As relações e os valores das bandas foram utilizados como covariáveis. Foram também utilizados dados de declividade, altitude da calha do rio e valor da precipitação acumulada 10 e 20 dias antes da coleta da amostra.

Foi utilizado a máscara de nuvem para eliminar pontos que apresentavam na imagem 100% de cobertura de nuvens ou sombra de nuvens. Foram utilizados algoritmos de aprendizado de máquina que utilizam árvores de decisão como Cubist, Ranger e Random Forest e modelo Boost (GBM) na predição dos valores dos parâmetros de qualidade do rio ao longo do trecho avaliado utilizando o software R pacote Caret.

2.2 Covariáveis

Foi utilizada como covariável a precipitação de quatro estações pluviométricas instaladas na área de estudo pela Fundação Renova, com isso obtiveram-se as covariáveis precipitação efetiva no ponto, acumulada dois e cinco dias antes a coleta de amostra de Turbidez. Os cursos d'água foram obtidos da base Ottocodificada da Agência Nacional das Águas (ANA) (<https://dadosabertos.ana.gov.br/>).

2.3 Modelagem de qualidade da água

O processo de treinamento, validação e espacialização dos parâmetros de qualidade de água foram feitos de forma especial, visto o número de dias de coleta pequeno (6 dias) é também o número de pontos de coletas (13 pontos) (Hu et al., 2020; Vergara et al., 2018; Zhang et al., 2020). Dessa forma foi aplicado um método de validação denominado *Nested LOOCV* com adaptações para este trabalho (Figura 8). Esse método considera processo de duplo *loop*, onde no primeiro *loop* é realizado o treinamento do modelo com conjunto de dados de tamanho n-1, e a validação é feita no segundo *loop* com amostra removida, este processo é repetido pelo número de amostras do banco de dados. A aplicação do método *Nested LOOCV* com adaptações foi realizada considerando o segundo *loop* as datas amostradas, dessa forma o modelo rodou seis vezes, onde em cada uma delas, um dos dias, foi colocado no segundo loop de validação.

O resultado final da performance do algoritmo de *Machine Learning* foi a média dos seis resultados (Treinamento/Validação). Este método robusto avalia a performance do algoritmo e detecta possíveis dados com problemas nas coletas ou de covariáveis. Em cada rodada também foi calculada a espacialização do mapa dos parâmetros de qualidade de água para este dia que foi retirado para validação.

A métrica para escolha do melhor subconjunto de cada modelo foi o R^2 . Os algoritmos testados neste trabalho foram Cubist, Random Forest, Ranger, Generalized Boosted Regression Models (GBM) e Suport Vector Machine (SVM) usando o pacote *Caret* (Khun, 2020).

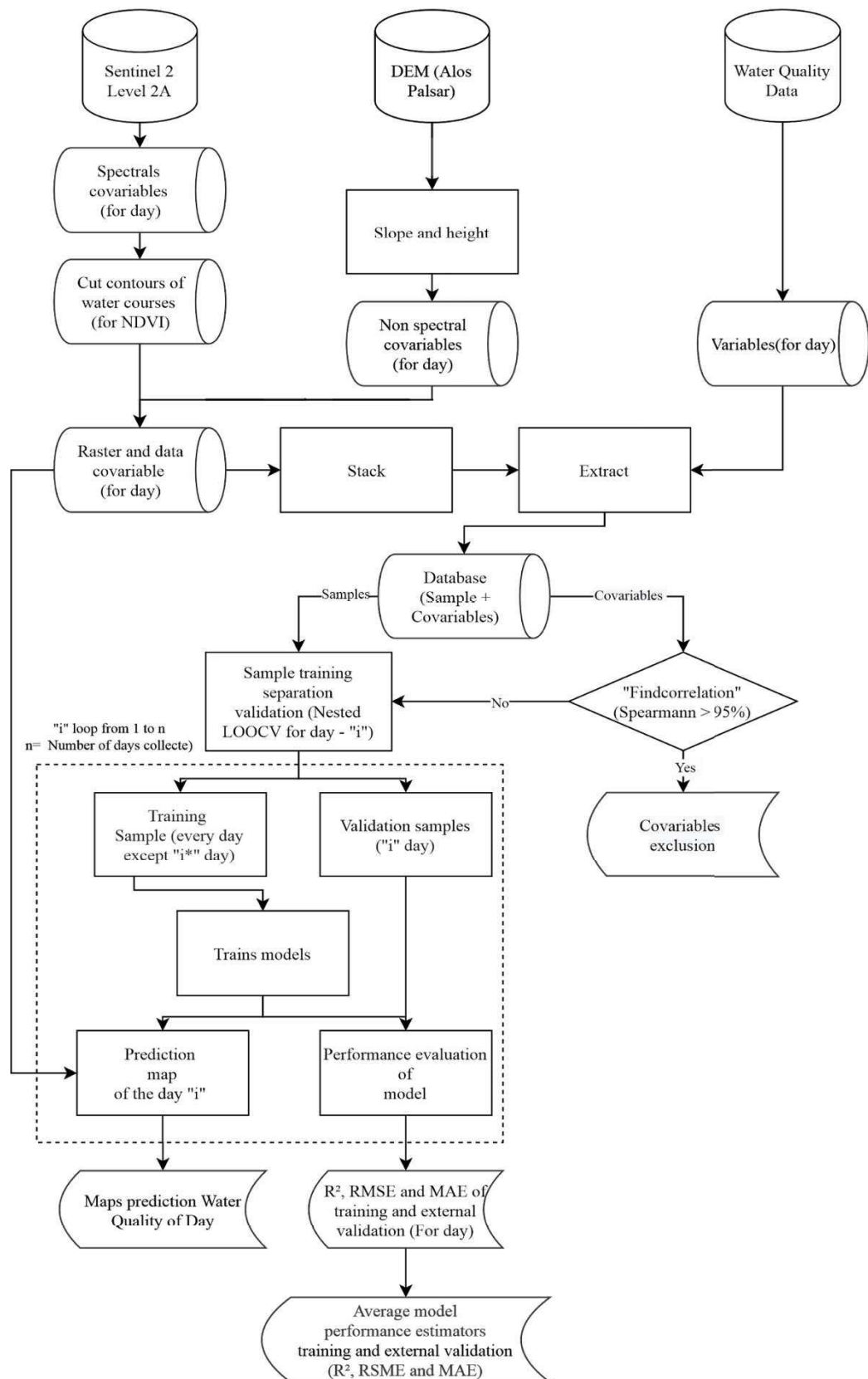


Figura 8: Fluxograma metodológico de modelagem e espacialização dos parâmetros de qualidade da água.

O modelo Cubist é uma função que se baseia no método M5 de Quinlan (1992). Nessa função são construídos modelos baseados em árvores. Regras são criadas e o modelo percorre a árvore de determinada regra, sendo adequado para cada regra com base no subconjunto de dados, quando isso ocorre o conjunto de regras é removido ou combinado. Por fim são criados modelos de regressão no final de cada árvore tendo como base os preditores que foram utilizados para definir as regras removidas (Kuhn, 2020).

O Random Forest, assim como o Cubist, é um modelo que também utiliza árvores de decisão, seu diferencial está no estabelecimento de florestas, para isso cada árvore é utilizada como base o valor de um vetor amostrado independente e com mesma distribuição para todas as árvores da floresta, neste modelo estimativas internas são usadas para mensurar a variável de importância (Breiman, 2001).

Foi utilizado o modelo de regressão por rede neural avNNet, esse modelo é baseado no modelo NNet (Taghizadeh-Mehrjardi et al., 2020), entretanto nesse modelo os resultados são baseados na média, sendo assim esse constrói números aleatórios, denominados de sementes, e o resultado global é calculado a partir da média das saídas de cada um dos números aleatórios criados. Com isso o modelo resultante é usado na previsão (Kuhn, 2020).

O SVM é um algoritmo linear binário não probabilístico, desenvolvido para classificação binária (Cortes and Vapnik, 1995) consiste em primeiramente separar as classes procurando o hiperplano de separação ideal entre duas classes. Para isso é buscado a margem entre pontos mais próximos da classe, os pontos situados nessas margens são chamados vetores de suporte e o meio entre os pontos de suporte o hiperplano de separação ideal (Meyer, 2009). Quando um separador linear não é encontrado, os pontos são projetados em um espaço de maior dimensão, para que eles sejam efetivamente separados linearmente, para isso é requerido projeção por meio de técnicas de Kernel (Meyer, 2009). Esta técnica ajusta parâmetros de custo e o parâmetro Sigma do Kernel. Sendo avaliado, no máximo, seis valores do parâmetro Kernel. Isso permite uma pesquisa ampla sobre o parâmetro de custo e uma pesquisa relativamente estreita sobre Sigma (Amari and Wu, 1999).

Outro algoritmo baseado em Kernel utilizado foi o kkn que tem como base o vizinho mais próximo (kernel) ao estabelecer dados de teste, por um de treinamento, por meio da distância de Minkowski, forma geral da distância Euclidiana e de Manhattan. Sendo assim, esse modelo pode ser utilizado não só na classificação, como também na regressão. Com isso esse algoritmo é definido como não paramétrico, ou seja, que não faz nenhuma suposição explícita, evitando modelar distribuições subjacente em relação aos dados (Abdulrahman et al., 2020).

O modelo linear (lm), também foi aplicado, esse modelo produz um vetor de previsões, para prever o valor da variável resposta com base em uma ou mais variáveis preditoras, com objetivo de estabelecer relação linear entre a variável preditora e a variável resposta, sendo possível estimar o valor da resposta com apenas os valores preditores conhecidos.

2.4 Avaliação dos algoritmos de aprendizado de máquina

Os algoritmos de aprendizado de máquina foram avaliados quanto sua performance preditiva para cada um dos parâmetros de qualidade de água. Foram escolhidos os algoritmos que apresentassem a melhores performances de acordo com coeficiente de correlação (R^2) no conjunto de validação (Equação 1). A raiz do erro quadrático médio – RMSE (Equação 2), erro absoluto médio – MAE (Equação 3), também foram avaliados.

$$\text{Equação 1:} \quad R^2 = 1 - \frac{\sum_{i=1}^n (P_i - P_m)^2}{\sum_{i=1}^n (P_i - O_m)^2}$$

Onde: P_i o valor predito, P_m o valor predito médio, P_i o valor observado, o O_m valor observado médio e o n o número de observação.

$$\text{Equação 2:} \quad RMSE = \left[\frac{1}{n} \sum_{i=1}^n (P_i - O_i)^2 \right]^{\frac{1}{2}}$$

Onde: P_i o valor predito, O_i o valor observado e o n o número de observação.

$$\text{Equação 3:} \quad MAE = \frac{\sum_{i=1}^n |P_i - O_i|}{n}$$

Onde: P_i o valor predito, O_i o valor observado e o n o número de observação.

3 RESULTADOS E DISCUSSÃO

Os testes realizados para verificar a interferência da resolução espacial da imagem em função da modelagem podem ser verificados nas Tabelas 3 e 4. Com um ponto do trecho em que a largura do rio é inferior a resolução da imagem, verifica-se que os pontos localizados na cabeceira da bacia (RCA01, RGN01, RGN06 RGN08) interferem no resultado, em virtude da resolução espacial da imagem. O teste utilizando o ponto mais próximo a foz do rio Gualaxo do Norte (RGN08) tem os menores valores de R^2 quando comparado com os demais pontos.

Tabela 3: Resultado das simulações dos modelos utilizando apenas um, dos quatro pontos localizados onde a largura do rio é inferior a resolução da imagem

Modelo	Com o RCA01			Com o RGN01			Com o RGN06			Com o RGN08		
	R ²	RMSE	MAE	R ²	RMSE	MAE	R ²	RMSE	MAE	R ²	RMSE	MAE
avNNet	0,31	47,84	34,84	0,33	48,03	35,18	0,26	54,08	40,62	0,17	77,77	51,13
cubist	0,61	17,55	13,24	0,61	18,93	14,52	0,66	21,87	17,37	0,25	70,83	37,76
kknn	0,59	20,46	13,84	0,57	20,59	13,82	0,65	23,82	17,49	0,30	71,24	36,46
lm	0,48	20,77	16,61	0,49	22,67	18,25	0,43	26,85	20,24	0,29	70,87	47,57
rf	0,60	20,81	14,60	0,55	20,14	13,86	0,73	21,64	16,05	0,32	65,77	35,24
svm	0,55	27,88	16,77	0,58	25,69	16,71	0,57	27,03	18,98	0,21	61,08	34,21

Onde: RCA é o ponto localizado no rio Carmo; RGN são os pontos localizados no rio Gualaxo do Norte; R² o coeficiente de determinação ajustado; RMSE o erro quadrático médio; e MAE o erro absoluto médio.

Quando o teste foi realizado removendo um dos quatro pontos pode-se verificar que as simulações sem os pontos como o RCA01, RGN01 e RGN06 o R² de ajuste de modelo são inferiores a 0,40 para a maioria dos modelos. No entanto, quando remove o RGN08 todos os preditores são influenciados positivamente, com aumento do valor do R² e redução do MAE em todos os modelos testados.

Tabela 4: Resultado das simulações dos modelos com a remoção de um, dos quatro pontos localizados onde a largura do rio é inferior a resolução da imagem

Modelo	Sem o RCA01			Sem o RGN01			Sem o RGN06			Sem o RGN08		
	R ²	RMSE	MAE	R ²	RMSE	MAE	R ²	RMSE	MAE	R ²	RMSE	MAE
avNNet	0,16	75,88	50,40	0,17	75,80	50,13	0,26	71,66	45,59	0,24	50,13	36,84
cubist	0,23	79,62	42,92	0,27	75,37	38,39	0,34	49,05	27,40	0,70	17,10	12,61
kknn	0,38	70,69	37,39	0,42	66,83	33,84	0,27	64,48	31,66	0,48	23,31	15,58
lm	0,26	70,91	49,46	0,27	68,79	47,06	0,30	65,54	43,89	0,40	23,46	18,57
rf	0,31	61,39	34,64	0,39	61,57	30,92	0,25	61,50	32,36	0,54	22,23	15,83
svm	0,12	58,99	32,45	0,28	55,32	30,55	0,18	56,86	27,93	0,47	28,71	18,42

Onde: RCA é o ponto localizado no rio Carmo; RGN são os pontos localizados no rio Gualaxo do Norte; R² o coeficiente de determinação ajustado; RMSE o erro quadrático médio; e MAE o erro absoluto médio.

Foi encontrado o valor de 0,70 de coeficiente de determinação para o cubist, quando foi utilizado a modelagem removendo o ponto RGN08 e mantendo os outros 12 pontos, entretanto o rf, outro modelo baseado em árvore de decisão, não teve valor próximo como acontece no caso da modelagem para a bacia 2 (Tabela 5).

Na Figura 9 é possível observar a diferença entre os pontos localizados na região onde a largura do curso d'água é inferior a resolução da imagem do sensor MSI (RCA01, RGN01,

RGN06 e RGN 08), comparado com o ponto localizado no rio Doce (RDO08), onde a largura do rio é superior a resolução da imagem do sensor MSI.

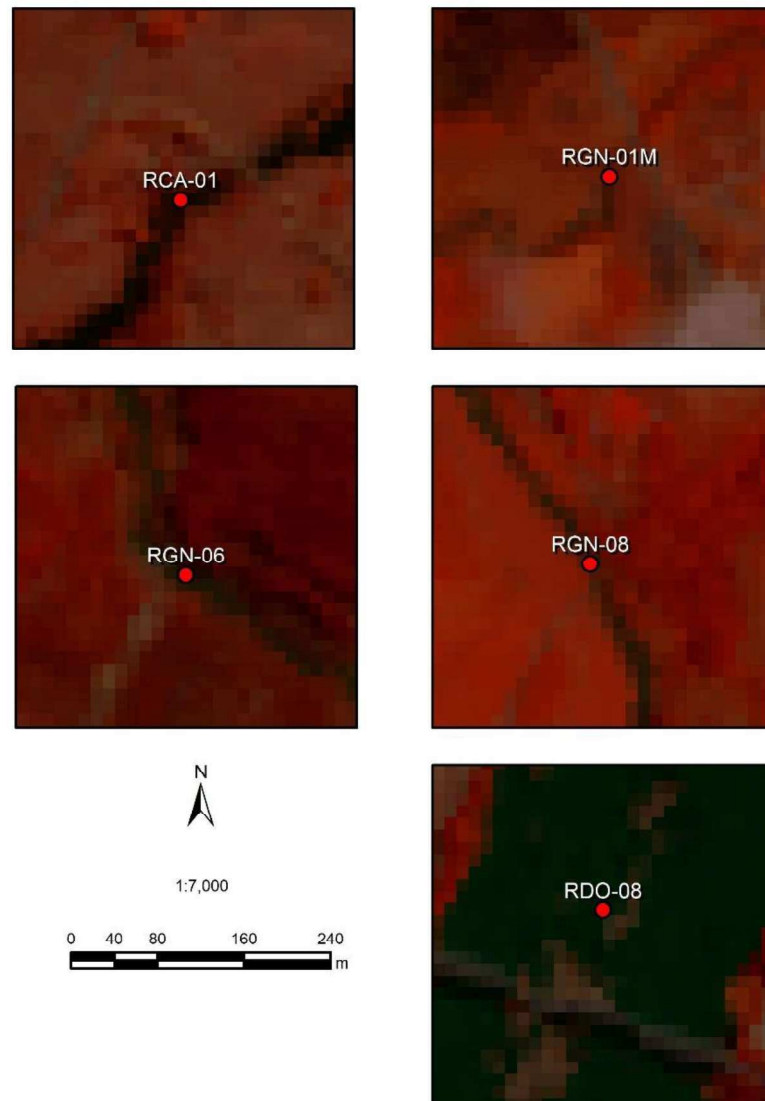


Figura 9: Composição falsa cor (Infravermelho, Verde e Azul) mostrando a localização dos pontos onde a resolução espacial é superior a largura do rio (RGN01, RGN06 e RGN 08), comparado com ponto em que a resolução da imagem é inferior que a largura do rio (RDO08).

A coloração dos cursos d'água na composição falsa cor tem variação com a quantidade de sólidos em suspensão, em cursos d'água limpo a reflectância decai entre 550 e 750 nm absorvendo todo o comprimento de onda próximo a 800 nm. Já com o aumento de sólidos na água a absorção reduz, não refletindo mais que 60 % do espectro (Lillesand et al., 2015). Logo pode ser verificado que nos pontos RGN01, 06, 08 e RCA01 tem maior reflectância da banda Nir (842 nm), ou seja, maior influência de bancos de areia, rochas e vegetação, próximo ao ponto de avaliação, quando comparado com o ponto RDO08.

O esquema da Figura 10, ilustra a influência da largura na qualidade da informação da resposta espectral em função da largura do rio, indicando que, para melhor levantamento dos dados ao utilizar as bandas de 10 m do sensor MSI, a coleta de informação deve ser em pontos onde a largura do rio é superior a 20 m. Contudo, em virtude da cheia e seca dos rios, uma margem de segurança deve ser utilizada devendo ser utilizado a largura de 30 m.

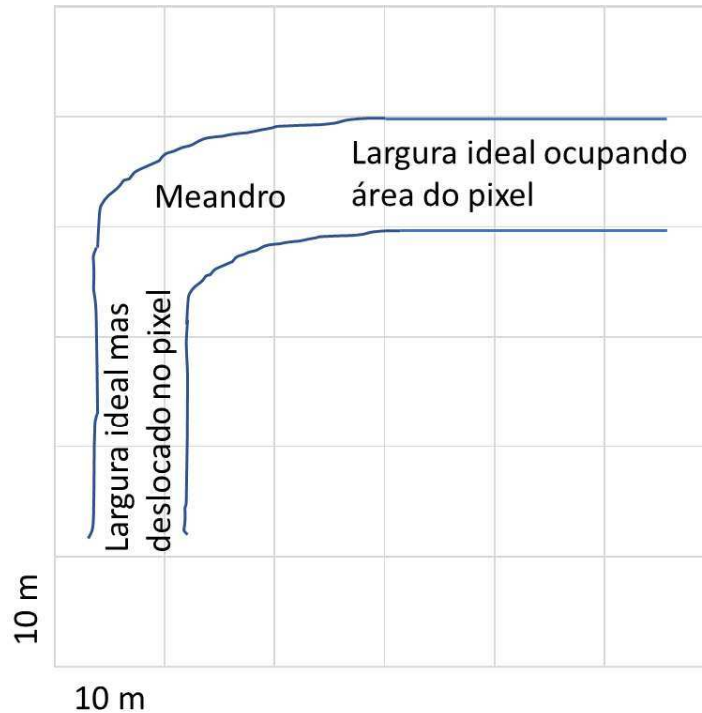


Figura 10: Esquema gráfico da interferência da largura do rio em função da resolução da imagem

Verifica-se que os resultados dos modelos utilizando os dados para toda a bacia do rio Doce (RD) (Bacia 1) em comparação com os dados da bacia RD excluindo os pontos a montante da foz entre o rio Gualaxo do Norte e o rio Carmo (Bacia 2), são inferiores (Tabela 5). Quatro modelos possuem R^2 maior que 0,5, sendo o cubist o modelo que apresentou menor MAE (13,97), que para turbidez é um valor baixo em virtude da amplitude dessa variável (Foesch et al., 2020).

Tabela 5: Resultado dos modelos comparando os dados das duas bacias, rio Doce e com os dados da bacia RD excluindo a bacia do rio Gualaxo do Norte

Modelo	Bacia 1			Bacia 2		
	R ²	RMSE	MAE	R ²	RMSE	MAE
avNNet	0,16	73,13	47,78	0,22	49,96	36,91
cubist	0,23	64,52	31,74	0,62	17,64	13,97
kknn	0,38	60,72	32,63	0,59	24,46	16,94
lm	0,28	64,32	44,05	0,46	23,15	18,21
rf	0,32	60,59	33,44	0,54	22,45	16,08
svm	0,12	58,12	30,72	0,59	25,75	18,02

O modelo kknn foi o que obteve valores de R² mais altos para as duas bacias concomitantemente, mesmo sendo um modelo mais simples, indicando que modelos não paramétricos, podem ser utilizados na predição, quando comparados com modelos de árvores de decisão. Esse resultado mostra também que não houve extrapolação entre os dados de treinamento e validação (Abdulrahman et al., 2020). O modelo linear (lm), teve valores de R² mais alto que o modelo de redes neurais (avNNet), em ambas as bacias. Para a bacia 2, todos os modelos tiveram valores mais altos que o modelo não avNNet, mostrando que modelos paramétricos são menos indicados na predição, como observado também por Goh et al. (2021). Ao contrário do avNNet, o resultado modelado pelo svm foi inferior ao modelar a Bacia 1, mas obteve resultados acima de 0,50 para a Bacia 2, isso porque na Bacia 1 o svm não conseguiu separar os conjuntos de dados efetivamente (Meyer, 2009)

O resultado para as regressões lineares das variáveis, Cor verdadeira, SST, Ferro total e Alumínio total, utilizando os dados da Fundação Renova são apresentados nas Figuras 11 a 14, sendo que as variáveis SST e Fe total obterem R² superior a 0,80.

A regressão da Cor verdadeira teve 0,18 de R², como pode ser verificado na Figura 8, houve dispersão dos pontos com cerca de 400 mg Pt/L e 100 de turbidez, e valores de 300 mg Pt/L e mais de 1400 de turbidez, o que na regressão das variáveis. Esse fato está associado a análise de cor realizada, em que a leitura é realizada após a remoção de partículas $\leq 1,2 \mu\text{m}$, removendo as partículas que dão turbidez a amostra.

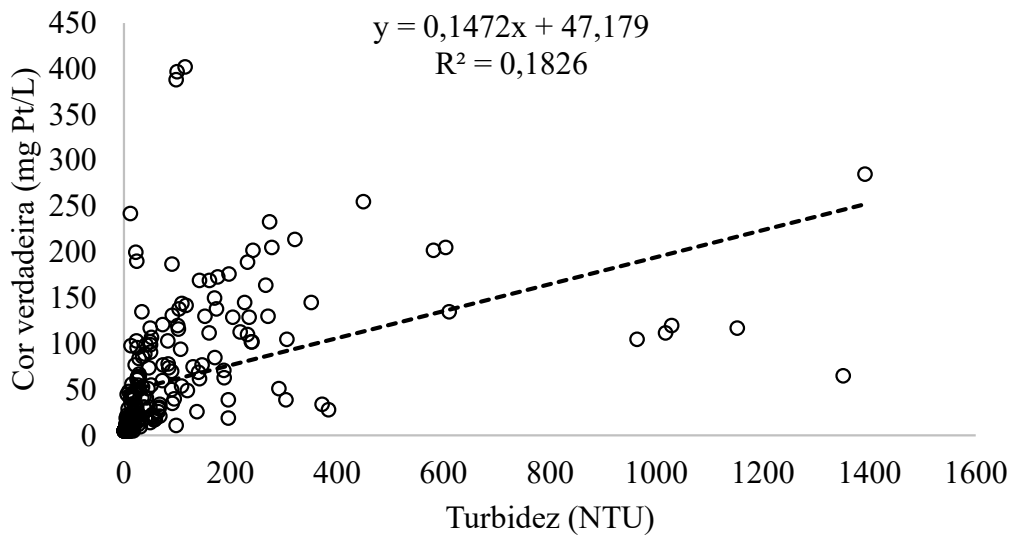


Figura 11: Regressão linear da Turbidez em função da Cor verdadeira (mg Pt/L).

Para os SST o R^2 foi de 0,90, sendo entre as variáveis modeladas a que obteve melhor resultado. Nota-se que a distribuição dos pontos teve baixa dispersão, com pontos próximos a linha de tendência. Esse resultado está muito relacionado a metodologia de determinação dessas variáveis, já que SST considera partículas superiores a 1,2 μm de diâmetro em determinado volume de água.

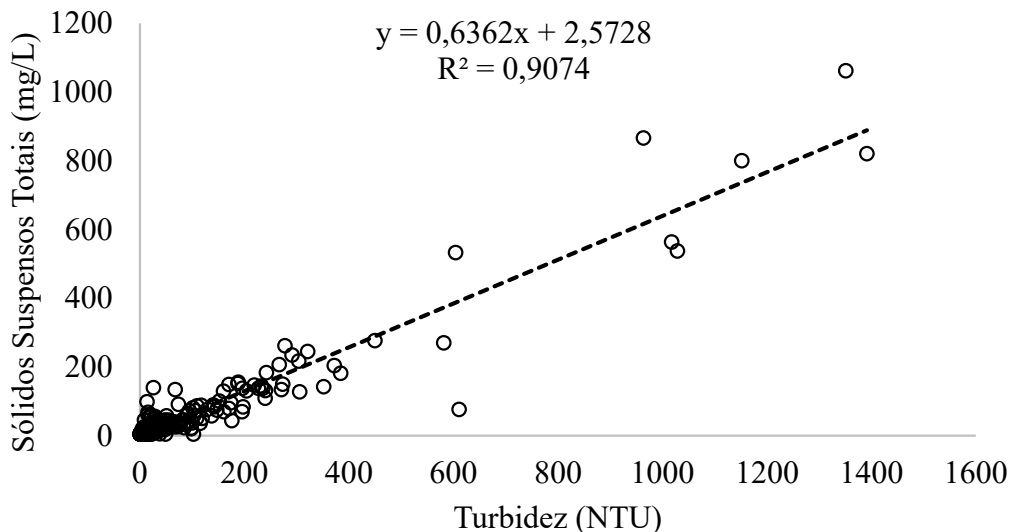


Figura 12: Regressão linear da Turbidez em função dos Sólidos Suspensos Totais (mg/L).

O Alumínio total teve R^2 de 0,58, isso porque um dos pontos utilizados na regressão está muito distante da linha de tendência. A variável Alumínio total considera tanto os valores de metais dissolvidos como em partículas em suspensão, visto que os solos tropicais apresentam

grande quantidade de Al nos constituintes de argila, a tendência seria da alta correlação entre a turbidez e o Al total.

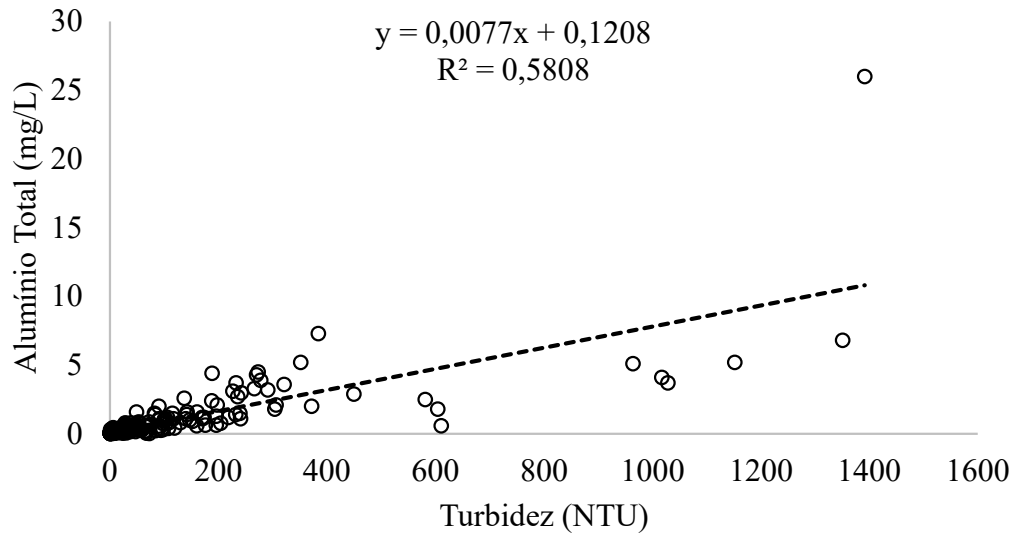


Figura 13: Regressão linear da Turbidez em função do Alumínio Total (mg/L).

Já o Fe total teve R^2 próximo ao do SST, variável física, com pontos da regressão próximo a linha de tendência. Essa relação era esperada, devido a mineralogia dos rejeitos ser rica neste elemento (Coward et al., 2017).

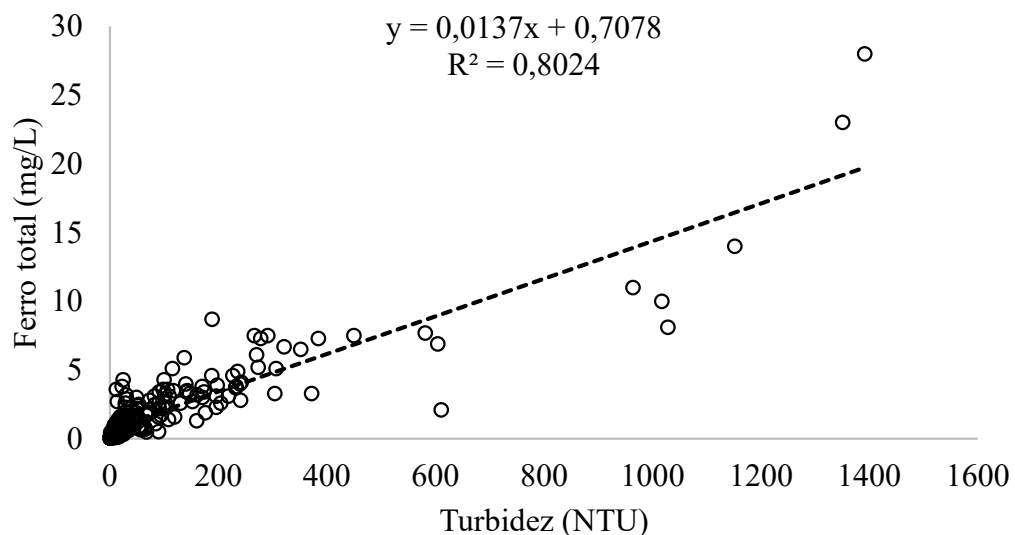


Figura 14: Regressão linear da Turbidez em função do Ferro total (mg/L).

De posse da regressão linear foi realizado a modelagem das variáveis Cor verdadeira, Sólidos Suspensos Totais, Ferro total e Alumínio total. O resultado dos modelos para as variáveis Cor verdadeira, Sólidos suspensos totais, Ferro total e Alumínio total podem ser visualizados nas Tabela 6 e 7.

Tabela 6: Métrica de desempenho dos modelos para predição da Cor Verdadeira e SST na área de estudo

Modelo	Cor			SST		
	R ²	RMSE	MAE	R ²	RMSE	MAE
avNNet	0,27	52,15	51,76	0,35	33,60	25,69
cubist	0,60	3,33	2,55	0,63	12,17	9,71
kknn	0,58	3,49	2,45	0,59	15,03	10,62
lm	0,46	3,41	2,68	0,46	14,73	11,59
rf	0,60	3,39	2,35	0,60	14,06	9,88
svm	0,55	4,28	2,71	0,57	18,30	12,12

Modelos de árvore de decisão como o cubist e o rf foram os que apresentaram melhores resultados, com R² superior a 0,60 e baixos valores de MAE, que estão relacionados ao baixo desvio dos resultados de forma absoluta, ou seja, na mesma dimensão dos valores preditos e observados.

Os metais também tiveram valores de R² correlacionados positivamente e próximos de 0,60, como os modelos seguindo a mesma tendência de quando comparados com as variáveis físicas. O avNNet foi semelhante aos demais modelos para o Al total, diferente das outras variáveis em que esse classificador ficou com o R² abaixo de 0,40.

Tabela 7: Métrica de desempenho dos modelos para predição da (Fe e Al) na área de estudo

Modelo	Fe			Al		
	R ²	RMSE	MAE	R ²	RMSE	MAE
avNNet	0,26	0,54	0,36	0,59	0,20	0,14
cubist	0,55	0,29	0,23	0,59	0,15	0,12
kknn	0,57	0,34	0,23	0,58	0,18	0,13
lm	0,46	0,32	0,25	0,46	0,18	0,14
rf	0,58	0,29	0,20	0,59	0,17	0,12
svm	0,58	0,40	0,26	0,59	0,22	0,14

Os resultados demonstram que em geral a utilização dos dados de turbidez das estações automáticas pode ser utilizada de forma contínua utilizando dados de sensoriamento remoto, e que esses dados podem ser utilizados também para estimar sólidos suspensos e Fe e Al total.

Na Figura 15 pode ser observado a representação dos valores de Turbidez e Al total de forma contínua, em ponto localizado no rio Carmo (RCA02), em Barra Longa/MG e do rio Doce (RDO08) na cidade de Governador Valadares/MG. Para Turbidez os valores são entre 40 a 50 NTU no período seco e de até 90 NTU no período chuvoso no ponto RCA02, e de 10 a 20 NTU no período seco e de até 50 NTU no período chuvoso no ponto RDO08. Na estação

chuvosa naturalmente a turbidez dos cursos d'água são elevados, em virtude da erosão e transporte de sedimentos pela perda de solos (Girardi et al., 2016; Santana et al., 2021; Theis et al., 1988).

No período seco é possível verificar altos valores de Turbidez, que podem estar associados a bancos de areia ou rocha próximo ao local da estação automática, como pode ser observado na Figura 16A, no período seco que se relacionam aos valores altos de Turbidez e Al total, já na Figura 16B, período chuvoso, devido ao maior volume de água boa parte do banco de areia ou rocha ficam submersos. Logo, também é importante analisar os resultados do ponto de vista da altura da lâmina d'água, que podem interferir em falsos valores de Turbidez elevada em períodos secos. Contudo a profundidade da coluna d'água não interfere somente no espalhamento da radiação eletromagnética pela coluna d'água (Barbosa et al., 2019), como também na exposição direta de objetos como bancos de areia e rochas.

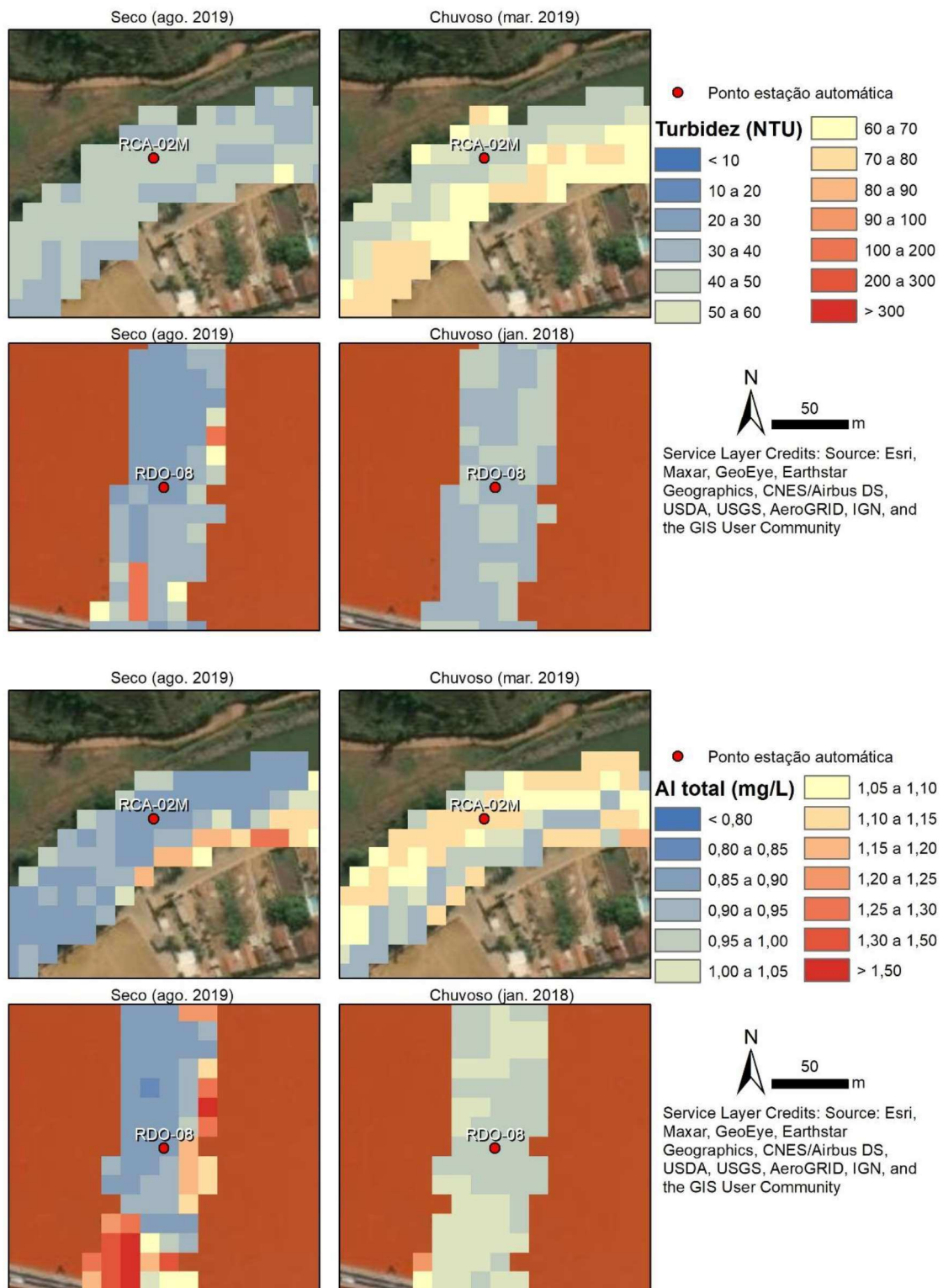


Figura 15: Espacialização dos resultados de Turbidez (NTU) e Al total (mg/L) após a modelagem, sendo representado para a Turbidez os resultados do modelo cubist e para o Al Total o modelo rf.



Figura 16: Imagens do Google Earth mostrando a localidade do ponto da estação automática RDO08, próximo a cidade de Governador Valadares. (A) imagem representando o período seco (julho/2019); (B) imagem representando o período seco (dezembro/2018).

4 CONCLUSÕES

O modelo gerado para Turbidez mostra que essa variável pode ser mensurada a partir de dados de sensores remotos com precisão de 60%, podendo ser apresentada de forma contínua em todo o trecho do rio Doce e do rio Carmo de Barra Longa/MG até sua foz em Rio Doce/MG.

A regressão linear do SST e do Fe total foram superiores a 0,80 de R^2 mostrando a correlação linear entre as variáveis. A cor verdadeira e o AI total não tiveram bom ajuste linear com a turbidez, mesmo assim na modelagem tiveram resultados superiores a 50% de precisão.

A modelagem contínua das variáveis avaliadas no trabalho pode ser utilizada para avaliar, a qualidade e a quantidade dessas variáveis ao longo de todo curso d'água, em especial em áreas onde não possui pontos de monitoramento, consistindo-se numa forma mais eficaz de representar essas variáveis.

REFERENCIAS

- Abdulrahman, S. A., Khalifa, W., Roushdy, M., & Salem, A. B. M. (2020). Comparative study for 8 computational intelligence algorithms for human identification. *Computer Science Review*, 36, 100237. <https://doi.org/10.1016/j.cosrev.2020.100237>
- Amari, S., & Wu, S. (1999). Improving support vector machine classifiers by modifying kernel functions. *Neural Networks*, 12(6), 783–789. [https://doi.org/10.1016/S0893-6080\(99\)00032-5](https://doi.org/10.1016/S0893-6080(99)00032-5)
- Asiken Alias, N., & Mansor, S. (2018). Estimation of dam sedimentation using remote sensing. *Proceedings - 39th Asian Conference on Remote Sensing: Remote Sensing Enabling Prosperity, ACRS 2018*, 2(1), 997–1008.
- Barbosa, C. C. F., Novo, E. M. L. M., & Martins, V. S. (2019). Introdução ao sensoriamento remoto de sistemas aquáticos.
- Bousbih, S., Zribi, M., Pelletier, C., Gorraab, A., Lili-Chabaane, Z., Baghdadi, N., et al. (2019). Soil texture estimation using radar and optical data from Sentinel-1 and Sentinel-2. *Remote Sensing*, 11(13). <https://doi.org/10.3390/rs11131520>
- Breiman, L. (2001). Random forests. *Machine Learning*, 45, 5–32. <https://doi.org/10.1201/9780429469275-8>
- Coimbra, K. T. O., Alcântara, E., & de Souza Filho, C. R. (2019). An assessment of natural and manmade hazard effects on the underwater light field of the Doce River continental shelf. *Science of the Total Environment*, 685, 1087–1096. <https://doi.org/10.1016/j.scitotenv.2019.06.127>
- Cortes, C., & Vapnik, V. (1995). Support-Vector Networks. *Machine Learning*, 20, 273–297. <https://doi.org/10.1109/64.163674>
- Coward, E. K., Thompson, A. T., & Plante, A. F. (2017). Iron-mediated mineralogical control of organic matter accumulation in tropical soils. *Geoderma*, 306, 206–216. <https://doi.org/10.1016/j.geoderma.2017.07.026>
- Elhag, M., Gitas, I., Othman, A., Bahrawi, J., & Gikas, P. (2019). Assessment of water quality parameters using temporal remote sensing spectral reflectance in arid environments, Saudi Arabia. *Water (Switzerland)*, 11(3). <https://doi.org/10.3390/w11030556>
- Foesch, M. D. S., Francelino, M. R., Rocha, P. A., & Gomes, A. R. L. (2020). River Water Contamination Resulting from the Mariana Disaster, Brazil. *Floresta e Ambiente*, 27(4), 2–10. <https://doi.org/10.1590/2179-8087.013218>

- Girardi, R., Pinheiro, A., Garbossa, L. H. P., & Torres, É. (2016). Water quality change of rivers during rainy events in a watershed with different land uses in Southern Brazil. *RBRH*, 21(3), 514–524. <https://doi.org/10.1590/2318-0331.011615179>
- Goh, K. M., Maulidiani, M., Rudiyanto, R., Abas, F., Lai, O. M., Nyam, K. L., et al. (2021). The detection of glycidyl ester in edible palm-based cooking oil using FTIR-chemometrics and ¹H NMR analysis. *Food Control*, 125(January), 108018. <https://doi.org/10.1016/j.foodcont.2021.108018>
- Gomes, L. C., Faria, R. M., de Souza, E., Veloso, G. V., Schaefer, C. E. G. R., & Filho, E. I. F. (2019). Modelling and mapping soil organic carbon stocks in Brazil. *Geoderma*, 340(November 2018), 337–350. <https://doi.org/10.1016/j.geoderma.2019.01.007>
- Hu, J., Zhao, Y., Li, M., Liu, J., Wang, F., Weng, Q., et al. (2020). Machine learning-based radiomics analysis in predicting the meningioma grade using multiparametric MRI. *European Journal of Radiology*, 131, 109251. <https://doi.org/10.1016/j.ejrad.2020.109251>
- Isidro, C. M., McIntyre, N., Lechner, A. M., & Callow, I. (2018). Quantifying suspended solids in small rivers using satellite data. *Science of the Total Environment*, 634, 1554–1562. <https://doi.org/10.1016/j.scitotenv.2018.04.006>
- Kilham, N. E., Roberts, D., & Singer, M. B. (2012). Remote sensing of suspended sediment concentration during turbid flood conditions on the Feather River, California A modeling approach. *Water Resources Research*, 48(1), 1–18. <https://doi.org/10.1029/2011WR010391>
- Kuhn, M. (2020). Package ‘Cubist.’ CRAN Repository, 14.
- Lillesand, T., Kiefer, R. W., & Chipman, J. (2015). *Remote Sensing and Image Interpretation* (7th ed.). Wiley.
- Meyer, D. (2009). *Support Vector Machines*. FH Technikum Wien (Vol. 25).
- Nascimento, L.P. do, Reis, D.A., Roeser, H.M.P., Santiago, A.D.F., 2019. Relationship between land use and water quality in a watershed impacted by iron ore tailings and domestic sewage. *Ambient. e Agua - An Interdiscip. J. Appl. Sci.* 14, 1. <https://doi.org/10.4136/ambi-agua.2383>
- Oliveira, E. N. de, Knoppers, B. A., Lorenzetti, J. A., Medeiros, P. R. P., Carneiro, M. E., & Souza, W. F. L. de. (2012). A satellite view of riverine turbidity plumes on the NE-E Brazilian coastal zone. *Brazilian Journal of Oceanography*, 60(3), 283–298. <https://doi.org/10.1590/S1679-87592012000300002>

- Rosa, M. R. (2018). Classificação do Padrão de Ocupação Urbana de São Paulo Utilizando Aprendizagem de Máquina e Sentinel 2. *Geography Department University of Sao Paulo*, 15–21. <https://doi.org/10.11606/rdg.v0ispe.145784>
- Saadi, A., & Campos, J. C. F. (2015). Geomorfologia do caminho da lama: contexto e consequências da ruptura da Barragem do Fundão (novembro 2015, Mariana–MG). *Arquivos do Museu de História Natural e Jardim Botânico - UFMG*, 24(1–2), 63–103.
- Santana, F. C., Francelino, M. R., Schaefer, C. E. G. R., Veloso, G. V., Fernandes-Filho, E. I., Santana, A. de J. P., et al. (2021). Water Quality of the Gualaxo do Norte and Carmo Rivers After the Fundão Dam Collapse, Mariana, MG. *Water, Air, & Soil Pollution*, 232(4). <https://doi.org/10.1007/s11270-021-05113-3>
- Taghizadeh-Mehrjardi, R., Schmidt, K., Amirian-Chakan, A., Rentschler, T., Zeraatpisheh, M., Sarmadian, F., et al. (2020). Improving the spatial prediction of soil organic carbon content in two contrasting climatic regions by stacking machine learning models and rescanning covariate space. *Remote Sensing*, 12(7). <https://doi.org/10.3390/rs12071095>
- Theis, T. L., Young, T. C., & DePinto, J. V. (1988). Factors affecting metal partitioning during resuspension of sediments from the Detroit River. *Journal of Great Lakes Research*, 14(2), 216–226. [https://doi.org/10.1016/S0380-1330\(88\)71550-6](https://doi.org/10.1016/S0380-1330(88)71550-6)
- Vergara, V. M., Mayer, A. R., Kiehl, K. A., & Calhoun, V. D. (2018). Dynamic functional network connectivity discriminates mild traumatic brain injury through machine learning. *NeuroImage: Clinical*, 19(July 2017), 30–37. <https://doi.org/10.1016/j.nicl.2018.03.017>
- Watanabe, F. S. Y., Imai, N. N., Barbosa, C. C. F., & Araujo, A. P. de. (2010). Avaliação do comportamento espectral da água: Estudo de caso do alto curso do rio Paraná (Presidente Epitácio – SP). III Simpósio Brasileiro de Ciências Geodésicas e Tecnologias da Geoinformação, 27–30.
- Zhang, Y., Liu, S., & Yu, X. (2020). Longitudinal structural MRI analysis and classification in Alzheimer's disease and mild cognitive impairment. *International Journal of Imaging Systems and Technology*, 30(2), 421–433. <https://doi.org/10.1002/ima.22390>

Capítulo 3 - Modelagem de parâmetros da qualidade da água de rio afetado por rejeito de mineração por meio do comportamento espectral

RESUMO

Atividades minerárias são de grande importância econômica no mundo, contudo, são associadas ao elevado potencial poluidor, não somente na área de exploração de minério, mas ao passivo ambiental que os estêreis e rejeitos possuem quando destinados às barragens. O rompimento da barragem de Fundão, Mariana/MG, em novembro de 2015, considerado um dos maiores acidentes ambientais do Brasil liberou cerca de 44 milhões de m³ atingindo o rio Doce até sua foz, elevando valores de parâmetros físicos e químicos desse rio. Em virtude da extensão das bacias hidrográficas monitorar os parâmetros de qualidade da água em campo pode ser difícil do ponto de vista técnico, devido ao dispêndio de pessoas e equipamentos. Com isso, o monitoramento da qualidade da água via sensoriamento remoto torna-se fundamental, à recuperação ou mitigação dos danos causados nos cursos d'água atingidos por rejeitos. Diante do exposto, objetivou-se correlacionar a resposta espectral de imagens de satélite do sensor orbital Sentinel-2 com parâmetros de qualidade da água. Levantar melhores modelos capazes de prever alguns parâmetros de qualidade da água do rio por meio de sensoriamento remoto orbital. Para tanto, nos anos hidrológicos de 2017 a 2019, a Turbidez foi determinada utilizando amostra de água in natura utilizando turbidímetro; a Cor aparente. O pH e a Temperatura determinadas in loco, com uso da sonda multiparâmetro. As análises de Sólidos Suspensos Totais (SST). Os teores de metais dissolvidos (Al, Zn, Cd, Fe, Mn, Cu, Cr) determinados segundo método EPA 3005A. Foram utilizadas imagens de Nível 2A do sensor MSI (Multi Spectral Instrument) acoplado ao sistema orbital Sentinel-2, extraído as bandas 2, 3, 4, 8, 8a, 11 e 12 e realizando relações com essas bandas. A altitude, a declividade, a precipitação pontual e contínua acumulada, ângulos Zenital e Azimutal juntamente com as bandas e suas relações foram utilizadas como covariáveis. Algoritmos de aprendizado de máquina que utilizam árvores de decisão e modelo Boost e Suport Vector Machine (SVM) foram utilizados na predição dos valores das variáveis. Dentre os parâmetros físicos, SST, Turbidez e Cor aparente estão altamente correlacionadas com os parâmetros químicos Mn e Fe dissolvidos em água. Índices espectrais estão inversamente correlacionados com os parâmetros físicos (SST, Turbidez e Cor) que utilizam a razão da banda 3 e banda 8 e correlacionados diretamente com a reflectância das bandas puras 3, 4 e 8. As variáveis com melhores resultados na modelagem foram Turbidez, Cor, Temperatura e Fe, por terem maior número de covariáveis correlacionadas. As relações de bandas desenvolvidas neste trabalho tiveram boa correlação com as variáveis Turbidez e Cor. As variáveis Cor e Turbidez obtiveram melhores resultados em quatro entre os cinco modelos testados. Das variáveis químicas avaliadas, a que obteve melhor resultado na modelagem foi o Fe. Os baixos valores encontrados na validação da modelagem são decorrentes da resolução espacial das imagens Sentinel-2, originalmente com 10 m e a largura dos cursos d'água, inferiores a essa resolução.

Palavras-chave: Random Forest; Sentinel-2; *machine learning*; Barragem de Fundão.

1 INTRODUÇÃO

A mineração, apesar de sua importância econômica, possui elevado potencial poluidor. Nas lavras a céu aberto, no processo de extração do mineral, o material sem valor econômico (estéril), é depositado em barragem de lavagem do minério (Mechi and Sanches, 2010). No Brasil a construção das barragens é realizada preferencialmente pelo método a montante (Cardozo et al., 2017). Em geral a classificação do potencial poluidor dessas barragens é grande segundo normativas do órgão ambiental controlador. A barragem de Fundão, localizada no município de Mariana/MG, e que sofreu rompimento em novembro de 2015, foi construída por esse método (Carmo et al., 2017; Morgenstern et al., 2016).

O evento foi considerado um dos maiores acidentes ambientais do Brasil devido não só pelo volume de rejeitado liberado (cerca de 44 milhões de m³), mas também pela distância que o rejeito percorreu (mais de 600 km), até a foz do rio Doce no oceano Atlântico (Golder Associates, 2017; Marta-Almeida et al., 2016), além de ter causado 19 mortes (AllaouaSaadi and Campos, 2015; Milanez and Losekann, 2016).

O rio Gualaxo do Norte foi um dos primeiros rios a ser atingido pela onda de rejeito e o que sofreu as maiores alterações e impactos ambientais. Os solos dos terraços fluviais desse rio foram totalmente arrastados pela onda de rejeito, sendo substituídos por sedimentos oriundos da barragem, que também cobriu com camada de até 200 cm o leito menor e toda planície fluvial, gerando um novo ambiente dominado por Tecnosolos (Schaefer et al., 2016). Além disso, as águas foram modificadas química e fisicamente, elevando valores de Cor, Turbidez, Sólidos Suspensos Totais (SST), e principalmente metais como Ferro e Manganês (Silva et al., 2016).

As variações nos valores de Turbidez podem estar relacionadas também à composição do assoalho do canal, que por aumento da turbulência por motivo de precipitações intensas pode gerar maior quantidade de sólidos em suspensão (Maia and Pereira, 2017), principalmente naqueles que possuem composição mais argilosa (Gong et al., 2018).

Parâmetros da qualidade da água variam por diversas razões como a sazonalidade climática, alterações no uso e cobertura do solo na área de entorno e outras atividades que são realizadas na bacia hidrográfica. Monitorar os parâmetros de qualidade da água em campo pode ser difícil do ponto de vista técnico devido ao dispêndio de pessoas, equipamentos e custo (Queiroz et al., 2012). No entanto, se faz incontestável a sua importância frente ao valor informacional que esse dado de trabalho gera e que podem colaborar para evitar problemas de

saúde pública, desabastecimento, monitoramento, manejo e sustentabilidade do recurso (Freitas e Freitas, 2006).

Uma alternativa para o monitoramento da qualidade da água é o uso conjunto de ferramentas de geoprocessamento e sensoriamento remoto (SR), que também podem auxiliar no processo de tomada de decisões no âmbito do monitoramento da qualidade hídrica (Donha et al., 2006). Trata-se de método que tem o seu diferencial na rapidez das avaliações e registros, com a possibilidade de manipular um volume maior de dados e informações. Porém, o uso destas ferramentas aplicadas ao monitoramento de sistemas aquáticos continentais é lacuna que ainda necessita de maiores investigações, no que envolve coleta de informações qualitativas e quantitativas de forma acurada para dados tróficos físicos, químicos e bióticos (INPE, 2019).

Com isso, o monitoramento da qualidade da água via ferramentas SR, possibilita a análises e controle de questões ambientais de amplas áreas e com um menor custo, principalmente se for possível o uso de imagens com acesso gratuito. Sendo fundamental no que se refere ao monitoramento dos danos causados nos cursos d'água atingidos por rejeitos oriundos de rompimento de barragens, como os trabalhos desenvolvidos por Bustamante et al. (2009); Martinez et al. (2015); Polidorio et al. (2004); Wang et al. (2018); Kilham et al. (2012) que avaliaram a modificação de cursos d'água de águas interiores.

Assim, três hipóteses são levantadas neste trabalho: a) Existe correlação entre a resposta espectral de imagens de sensores orbitais e parâmetros químicos e físicos em cursos d'água; b) É possível modelar parâmetros físicos e químicos de qualidade da água utilizando o sensoriamento remoto; c) Com uso de imagem orbital com resolução espacial de 10 m é possível modelar a qualidade de águas interiores?

Diante do exposto, objetivou-se correlacionar a resposta espectral de imagens de satélite do sistema orbital Sentinel-2 com parâmetros de qualidade da água, bem como ajustar modelos utilizando técnicas de *machine learning*, capazes de prever a qualidade da água do rio por meio de sensoriamento remoto orbital.

2 MATERIAL E MÉTODOS

2.1 Área de Amostragem

Os rios Gualaxo do Norte e Carmo, considerados cursos d'água de cabeceira, pertencem à bacia hidrográfica do rio Doce, localizados, boa parte, no município de Mariana e Barra Longa, ambos no estado de Minas Gerais. Ambas as bacias hidrográficas são semelhantes quanto às classes de litologia e solo, onde dominam as rochas básicas intrusivas dos Grupos Maquiné, Caraça, Itabira, Piracicaba e Itabira, além da Formação Cauê, com concentração de minério de ferro. Na bacia do rio Carmo ocorre a mineração aurífera de depósitos mesotermiais em rochas *Greenstone belts*, contato entre rochas do grupo Nova Lima e do Supergrupo Minas (Costa et al., 2003).

A montante da bacia dominam os Cambissolos Háplico Tb Distrófico e Distroférico (UFV et al., 2010), enquanto nas áreas de relevo ondulado a suave-ondulado ocorrem domínio de Argissolos Vermelho-Amarelos Eutróficos. A terceira classe de solo de importância são os Latossolos Vermelho-Amarelos Distróficos, que ocupam boa parte da porção a jusante das bacias (Costa et al., 2003; Guerra et al., 2017; Schaefer et al., 2015).

Os climas nas sub-bacias são classificados (Köppen-Geiger) como subtropical úmido com invernos secos (Cwa) nas regiões de altitude elevada, nas áreas com altitude menos elevada o clima é caracterizado como subtropical de altitude, com inverno seco e verão ameno (Cwb), com pluviosidade média anual de 1.250 a 1.450 mm.

As coletas de amostras de água foram realizadas conforme metodologia de coleta de Clesceri et al. (2005), em datas coincidentes com a passagem do sistema orbital Sentinel-2 na região, ao longo do biênio (2017-2018) totalizando 10 dias de coletas, em 13 pontos de amostragem, distribuídos espaçadamente ao longo dos rios Gualaxo do Norte e Carmo (Figura 17).

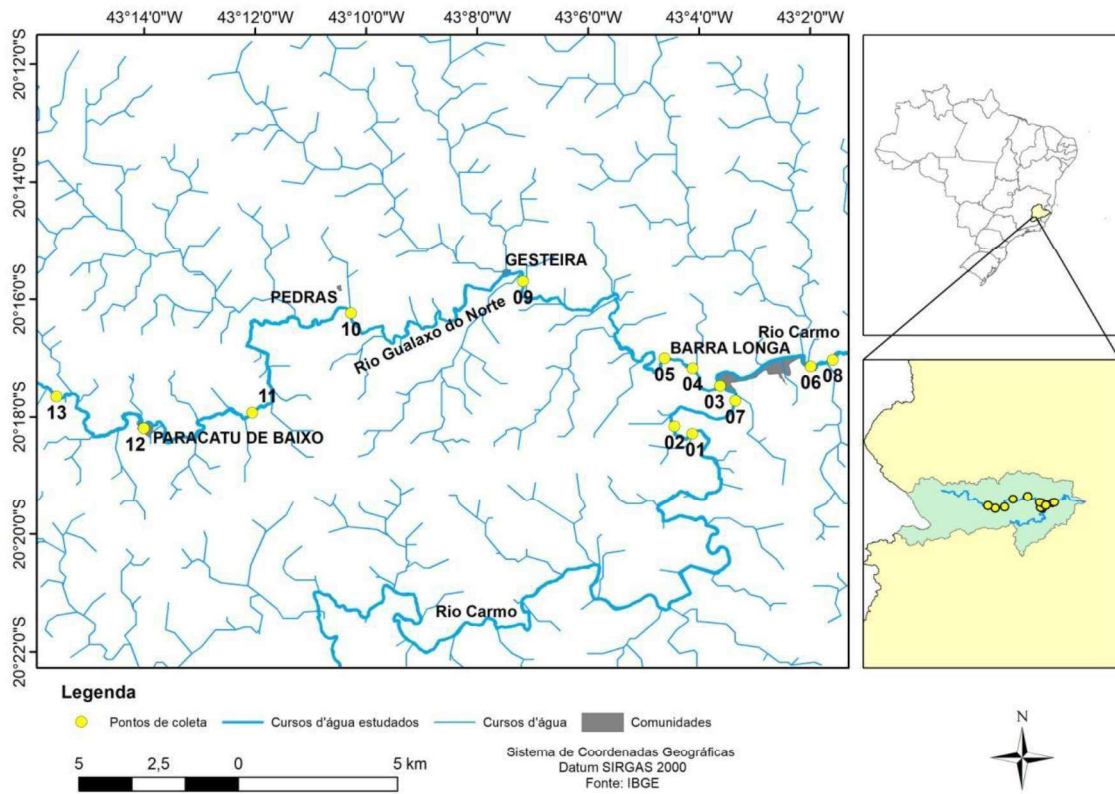


Figura 17: Pontos de amostragem nos trechos dos rios Gualaxo do Norte e Carmo.

2.2 Aquisição e espacialização dos dados

2.2.1 Parâmetros gerais

O processo de modelagem e espacialização de parâmetros de qualidade da água segue a estrutura definida pelo fluxograma abaixo (Figura 18), definindo de forma simples toda a metodologia que será descrita.

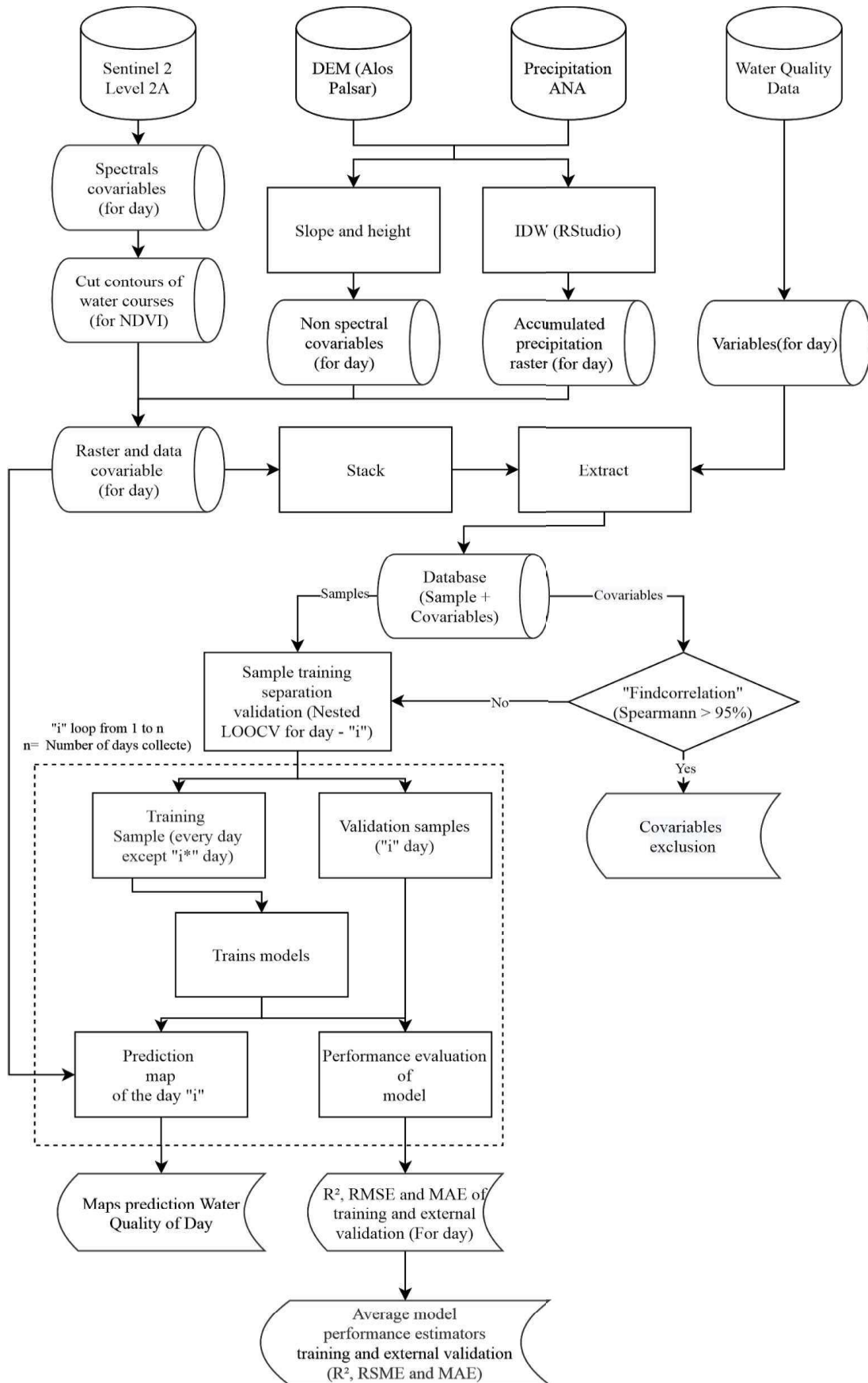


Figura 18: Fluxograma metodológico de modelagem e espacialização dos parâmetros de qualidade da água.

2.2.2 Aquisição dos parâmetros de qualidade da água

Os parâmetros pH e Temperatura foram determinados *in loco* nos pontos de coleta dos rios utilizando o equipamento PCD-PHE. A turbidez e Cor aparente foram determinadas em laboratório, sendo a turbidez utilizando o equipamento 2100AN Turbidimeter-Hach (unidade: NTU), e a Cor aparente utilizando o equipamento *Aqua Color Cor-Policontrol* (unidade: uC). Os teores de metais dissolvidos (Al, Zn, Cd, Fe, Mn, Cu, Cr), foram determinados segundo método EPA 3005A (EPA, 1992) sendo quantificados pelo aparelho AAS modelo 240FS AA da Agilent Technologies. As análises de SST foram realizadas pelo método gravimétrico (unidade: mg/L). Os pontos de coleta foram georreferenciados para posterior uso na modelagem utilizando receptor GPS.

2.2.3 Covariáveis espectrais

As covariáveis espectrais usadas neste trabalho foram adquiridas do sensor MSI (*Multi Spectral Instrument*) acoplado ao sistema orbital Sentinel-2, com processamento de Nivel 2A, disponíveis no repositório: <https://scihub.copernicus.eu/dhus/#/home>. Essas imagens apresentam correção geométrica, ortorretificação e registro espacial, além de reflectância corrigida abaixo da atmosfera. As datas utilizadas tentavam se manter próximas das datas de coletas mantendo no máximo 5 dias de diferença. Devido a limitação da predição, imagens que apresentavam taxa de coberturas de nuvens superiores a 80% foram eliminadas e por consequência as leituras de campo. A determinação da taxa de nuvens foi realizada utilizando a máscara de nuvens disponibilizada pelas imagens providas do satélite. Essas máscaras são baseadas nas bandas 1 ou 2 e as bandas SWIR - *Short-wave infrared* (banda 10 em conjunto com a banda 11 ou 12). Todas as bandas foram reamostradas na mesma resolução espacial (10 m) (Coluzzi et al., 2018).

Com os arquivos que passaram pelo processo de eliminação de nuvens foram extraídas as bandas 2, 3, 4, 8, 8a, 11 e 12 (Tabela 8). As bandas 8a, 11 e 12, com resolução de 20 m também foram interpoladas para a resolução de 10 m usando o método do vizinho mais próximo no software ArcGIS 10.8.1, para coincidir com a resolução das bandas mais detalhadas. O conjunto final de bandas foi agrupado a 28 relações de bandas criadas a partir do conjunto de bandas base (Tabela 9).

Tabela 8: Comprimentos de onda e resoluções espaciais dos instrumentos MSI (ESA, 2016)

Resolução espacial (m)	Bandas	Comprimento de onda central (nm)
10	2	490
	3	560
	4	665
	8	842
20	5	705
	6	740
	7	783
	8a	865
	11	1 610
	12	2 190
60	1	443
	9	945
	10	1 380

Tabela 9: Covariáveis espectrais utilizadas na modelagem dos parâmetros de qualidade da água

Covariável	Equação	Nome	Autor
bd8bd2	$(bd8-bd2)/(bd8+bd2)$		
bd8bd3	$(bd8-bd3)/(bd8+bd3)$	GNDVI	(Gitelson et al., 1996)
bd8bd4	$(bd8-bd4)/(bd8+bd4)$		
bd4bd2	$(bd4-bd2)/(bd4+bd2)$		
bd4bd3	$(bd4-bd3)/(bd4+bd3)$	NDTI	(Lacaux et al., 2007)
bd3bd2	$(bd3-bd2)/(bd3+bd2)$		
IIA	$(bd3-(4*bd8))/(bd3+(4*bd8))$	Índice Indicador de Água	(Polidorio et al., 2004)
NDWI	$(bd3-bd8)/(bd3+bd8)$	NDWI	(McFeeters, 1996)
NDVI	$(bd4-bd8)/(bd4+bd8)$	NDVI	(Tucker, 1979)
Toming	$(bd3/bd4)$	Toming	(Toming et al., 2016)
Clay	$(bd11/bd12)$	Clay	(Bousbih et al., 2019)
Iron	$(bd4/bd2)$	Simple Ratio Red/Blue Iron Oxide	(Hewson et al., 2001)
NDMI	$(bd8-bd11)/(bd8+bd11)$	NDMI	(Zhang et al., 2016)

Onde: bd corresponde à banda espectral utilizada da imagem Sentinel-2, nível 2A.

Tabela 9: Covariáveis espectrais utilizadas na modelagem dos parâmetros de qualidade da água (Continuação)

Covariável	Equação	Nome	Autor
moisture_10	$(bd8a - bd11)/(bd8a + bd11)$		(Zhang et al., 2016)
FII	$(bd12/bd8) + (bd3/bd4)$	FII	(Kalinowski and Oliver, 2004)
MNDWI_10	$(bd3 - bd11)/(bd3 + bd11)$	MNDWI	(Xu, 2006)
MSI	$(bd11/bd8)$	MSI	(Miller et al., 1990)
Ktwetness	$(0.1509 * bd2) + (0.1793 * bd3) + (0.3279 * bd4) + (0.3406 * bd8) + (0.7112 * bd11) + (0.4572 * bd12)$	K-T Wetness	(Kauth and Thomas, 1976)
Aweinsh	$(4 * (bd3 - bd12)) - ((0.25 + bd8) + (2.75 + bd12))$	AWEI	(Feyisa et al., 2014)
Aweish	$(bd2 + (2.5 * bd3)) - (1.5 * (bd8 + bd12)) - (0.25 * bd12)$	AWEI-SH	(Feyisa et al., 2014)
Wri	$(bd3 + bd8)/(bd8 + bd11)$	WRI	(Mukherjee and Samuel, 2016)
SCI	$(bd11 - bd8)/(bd11 + bd8)$	NDBI	(Li and Liu, 2008)
SAVI	$((bd8 - bd4)/(bd8 + bd4 + 0.5)) * (1.5)$	SAVI	(Huete, 1988)
FE_SI	$(bd12/bd11)$	Silicato de Ferro	(Henrich et al., 2011)
FE_OX	$(bd11/bd8)$	Óxidos de Ferro	(Henrich et al., 2011)
FE2	$(bd12/bd8) + (bd3/bd4)$	Ferro ²⁺	(Rowan and Mars, 2003)
FE3	$(bd4/bd3)$	Ferro ³⁺	(Rowan and Mars, 2003)
NDTI	$(bd11 - bd12)/(bd11 + bd12)$	NDTI	(Van Deventer et al., 1997)

Onde: bd corresponde à banda espectral utilizada da imagem Sentinel-2, nível 2A.

2.2.4 Covariáveis não espectrais e precipitação acumulada

Para avaliação dos efeitos de clima e relevo na modelagem dos parâmetros de qualidade de água, foram usados dados que representam estas duas características. Para o relevo foram usadas a elevação e a declividade. O Modelo Digital de Elevação (MDE) usado com base foi provido do sensor de micro-ondas PALSAR (<https://search.asf.alaska.edu/#/?dataset=ALOS>) contidas no Satélite ALOS com resolução de 12,5 m, ou seja, com resolução próxima à das imagens espectrais. A declividade foi gerada partir do MDE usando os pacotes raster e RSAGA no software R. Os arquivos do MDE e declividade foram recortados para a área de estudo assim como os demais dados. Dos metadados das imagens do sistema orbital Sentinel 2, foram utilizados os ângulos Zenital e Azimutal do período de aquisição da imagem.

Os dados de clima usados na avaliação do efeito da precipitação de chuva nos parâmetros de qualidade da água, foram os volumes de água precipitado, 10 e 20 dias antes da coleta de amostra, para tal foi utilizado como covariável pontual, a partir dos dados de quatro estações pluviométricas instaladas pela Fundação Renova, responsável pelas atividades de recuperação da área afetada pelo desastre, realizando a seleção dos dados tendo em vista a estação mais próxima ao ponto de coleta. Também foram coletados dados de precipitação disponíveis pela Agência Nacional das Águas (ANA), Instituto Nacional de Meteorologia (INMET) e Centro Nacional de Monitoramento e Alertas de Desastres Naturais (CMADEN), No total foram utilizados dados de 27 estações durante o período de agosto de 2017 a dezembro de 2019. Todos os dados foram posteriormente tabulados para o período de investigação. A base de dados foi utilizada para gerar mapas de precipitação acumulada, dois, cinco e oito dias anterior à data de coleta da amostra de água, para isso foi utilizado a interpolação pelo método IDW (*Inverse Distance Weighted*) em ambiente R. O IDW responde a influência local do dado de precipitação diminuindo na medida em que a distância ao ponto aumenta (De By, 2001).

Como resultados foram gerados arquivos matriciais com a densidade de precipitação para área da microbacia do rio Gualaxo do Norte e Carmo (Gemmer et al., 2004). O período seco e chuvoso, determinado segundo Lima et al. (2019) foi segmentando para abril a setembro como período seco, e o período chuvoso de outubro a março.

Para se obter o pixel que representa com fidelidade o retorno espectral do curso d'água, a largura de uma margem a outra do curso d'água deveria ser de pelo menos 30 m, para que não houvesse interferência de áreas ripárias no retorno espectral do pixel. Levando-se em conta que a resolução do Sentinel-2 possui 10 m para as bandas do visível e infravermelho próximo (Nir)

próximo, foi realizado o processamento dos membros final, que consiste na identificação de espectros de referência, no software ENVI por meio do algoritmo *Viper* para detectar o pixel que melhor representava o curso d'água, ou seja, pixel que possuía menor interferência de vegetação ou solo de área ripária, conforme Kilham et al., 2012.

A largura do rio para cada ponto de amostragem foi mensurada para fim de discussão dos resultados, sendo apresentado na Tabela 10.

Tabela 10: Largura do rio nos pontos de amostragem e suas respectivas coordenadas

Ponto	Largura (m)	Coordenadas		Ponto	Largura (m)	Coordenadas	
		Latitude	Longitude			Latitude	Longitude
P01	36,60	-20,3028	-43,0674	P08	36,50	-20,2815	-43,0252
P02	29,00	-20,3008	-43,0727	P09	16,70	-20,2602	-43,1184
P03	17,00	-20,2892	-43,0591	P10	29,40	-20,2694	-43,1704
P04	18,00	-20,2842	-43,0674	P11	14,30	-20,2983	-43,1995
P05	21,20	-20,2817	-43,0759	P12	18,30	-20,3031	-43,2321
P06	23,90	-20,2832	-43,0322	P13	15,90	-20,2943	-43,2586
P07	27,30	-20,2933	-43,0547	Média	23,39		

As coordenadas são expressas no DATUM SIRGAS 2000.

2.3 Modelagem de qualidade da água

O banco de dados foi separado em dois conjuntos onde os preditores foram separados e as covariáveis passaram por um processo de remoção por correlação. Este processo é necessário para a geração de modelos mais simples, redução do custo computacional e devido alguns modelos não serem otimizados para trabalhar com variáveis com alta colinearidade. O processo de remoção por correlação foi realizada com a remoção de covariáveis que obtiveram correlação de Spearman superior ou igual a 95%. A duplas acima desse valor foram avaliadas com relação entre as outras covariáveis do trabalho. Este passo é executado pela função *find correlation* do pacote *Caret* (Kuhn, 2020b). As covariáveis que passaram por esta fase sofreram o processo de treinamento, validação e espacialização.

O processo de treinamento, validação e espacialização dos parâmetros de qualidade de água foram feitos de forma especial, visto o número de dias de coleta pequeno é também o número de pontos de coletas (Hu et al., 2020; Vergara et al., 2018; Zhang et al., 2020). Dessa forma foi aplicado um método de validação *Nested LOOCV* com adaptações para este trabalho. Esse método considera processo de duplo *loop*, onde no primeiro *loop* é realizado o treinamento do modelo com conjunto de dados de tamanho n-1, e a validação é feita no segundo *loop* com

amostra removida, este processo é repetido pelo número de amostras do banco de dados. A aplicação do método *Nested LOOCV* com adaptações foi realizado considerando o segundo *loop* as datas amostradas, dessa forma o modelo rodou dez vezes onde em cada uma delas um dos dias foi colocado no segundo loop de validação. O resultado final da performance do algoritmo de *Machine Learning* foi a média dos dez resultados (Treinamento/Validação). Este método avalia a performance do algoritmo e detecta possíveis dados com problemas nas coletas ou de covariáveis. Em cada rodada também criado um mapa com a predição para toda a área de estudo.

O treinamento foi feito usando o pacote Caret (Khun, 2020). com cinco valores para cada uma dos *hype* parâmetros internos de cada um dos algoritmos (*tuneLength*) foram testados, sendo que esses podem ser consultados no manual do pacote Caret <https://topepo.github.io/caret/train-models-by-tag.html>. A métrica para escolha do melhor subconjunto de cada modelo foi a R^2 . Os algoritmos testados neste trabalho foram Cubist, Random Forest, Ranger, Generalized Boosted Regression Models (GBM) e Suport Vector Machine (SVM), sendo que neste trabalho foi utilizado o SVM baseado no Kernel que ajusta parâmetros de custo e o parâmetro Sigma do Kernel. Sendo avaliado, no máximo, seis valores do parâmetro Kernel. Isso permite uma pesquisa ampla sobre o parâmetro de custo e uma pesquisa relativamente estreita sobre Sigma (Amari and Wu, 1999).

2.4 Avaliação dos algoritmos de aprendizado de máquina

Os algoritmos de aprendizado de máquinas foram avaliados quanto sua performance para cada um dos parâmetros de qualidade de água. Foram escolhidos o algoritmo que apresentasse a melhores performances de acordo com coeficiente de correlação (R^2) no conjunto de validação (Equação 1). A raiz do erro quadrático médio – RMSE (Equação 2), erro absoluto médio – MAE (Equação 3), também foram avaliados.

$$\text{Equação 1:} \quad R^2 = 1 - \frac{\sum_{i=1}^n (P_i - P_m)^2}{\sum_{i=1}^n (P_i - O_m)^2}$$

Onde: P_i o valor predito, P_m o valor predito médio, P_i o valor observado, o O_m valor observado médio e o n o número de observação.

$$\text{Equação 2:} \quad RMSE = \left[\frac{1}{n} \sum_{i=1}^n (P_i - O_i)^2 \right]^{\frac{1}{2}}$$

Onde: P_i o valor predito, O_i o valor observado e o n o número de observação.

Equação 3:
$$MAE = \frac{\sum_{i=1}^n |P_i - O_i|}{n}$$

Onde: P_i o valor predito, O_i o valor observado e o n o número de observação.

3 RESULTADOS E DISCUSSÃO

Para as análises estabelecidas, verificou-se que ocorre correlação positiva do volume precipitado 10 e 20 dias antes da coleta de amostra com os valores de Zn dissolvido, o que pode ser relacionado com provável adubação nas áreas. O índice de umidade (*moisture_10*), covariável de índice espectral, também está correlacionada com os valores de Zn dissolvido.

Os parâmetros físicos de Sólidos Suspensos Totais (SST), Turbidez e Cor estão altamente correlacionadas com os parâmetros químicos Mn e Fe dissolvidos em água. Diversos índices espectrais estão inversamente correlacionados com os parâmetros Físicos (SST, Turbidez e Cor), por utilizarem a razão da banda 3 e banda 8. Também estão correlacionados diretamente com a reflectância das bandas puras 3, 4 e 8, e aos índices Iron e NDMI (Figura 19).

No trabalho de Long and Pavelsky (2013) foi encontrado correlação de Spearman de 0,44 para faixa de Turbidez de 1 a 5 NTU quando utilizaram dados de sistemas orbitais IKONOS, Landsat e MODIS, esses autores também encontraram valores de correlação de 0,87 para Turbidez com 12 NTU utilizando dados do sensor TM, bandas 1, 3 e 4 (Azul, Vermelho e Nir) do sistema orbital Landsat 5.

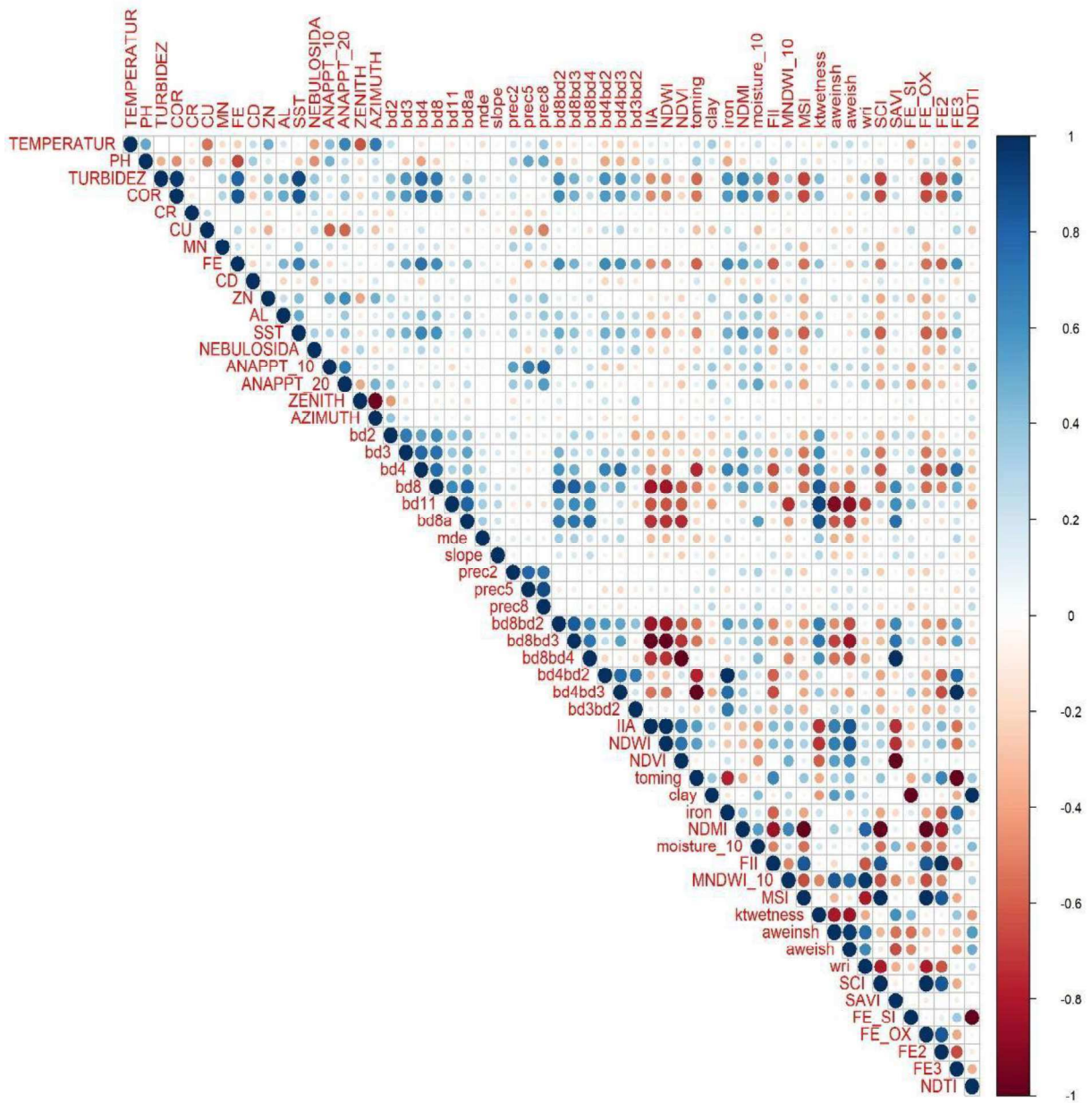


Figura 19: Matriz de correlação entre as variáveis e covariáveis.

Das 44 covariáveis, 13 foram eliminadas na predição devido à alta correlação entre elas, com valores acima de 0,95: bd8bd3; NDWI; aweish; MNDWI_10; NDMI; MSI; bd4bd3; FII; SAVI; bd8bd4; prec5; NDTI; Clay. Essa alta correlação pode estar associada a semelhança na fórmula de cálculo entre os índices, como por exemplo entre a relação de banda bd8bd3 com o índice NDWI, e com a relação de banda bd8bd3 com o índice SAVI.

Em geral as variáveis com melhor desempenho na modelagem foram aquelas que apresentaram correlação com maior número de covariáveis, dentre elas a variável

Temperatura, correlacionada positivamente com a precipitação pontual acumulada 10 e 20 dias antes da coleta de amostras, e como o ângulo Azimutal, e negativamente com a Nebulosidade e ângulo Zenital. Dessa forma, quanto maior a Nebulosidade do dia da coleta de amostra menor a Temperatura da amostra. O ângulo Azimutal faz referência à incidência do feixe de luz com o plano solar e o Zenital do feixe de luz com o ângulo Nadiral (INPE, 2019), logo a alta correlação entre o ângulo Azimutal e a Temperatura da amostra pode ser relacionada à elevação da Temperatura da amostra em decorrência da radiação solar que incide com menor ângulo sobre a superfície da água.

Turbidez e Cor correlacionaram positivamente com as bandas, 4 e 8, com as relações de bandas: $bd8/bd2$, $bd4/bd2$ e com os índices: iron, NDMI; negativamente com os índices espectrais toming, FII, MSI, SCI, FE_OX , FE2. A correlação positiva e negativa são decorrentes da correlação das bandas 4 e 8, e índices e relações de bandas em que as bandas 4 e 8 estão no numerador, correlações positivas, no denominador da equação, correlação negativa.

A variável química Fe, foi a que teve correlação com o maior número de covariáveis, sendo positiva com a banda 4, as relações de banda: $bd8/bd2$, $bd4/bd2$ e com os índices: iron, NDMI e FE3; negativamente com toming, FII, SCI, FE_OX e FE2.

Percebe-se que a correlação com índice FE3 e a variável Fe, foi superior a correlação com as variáveis Turbidez e Cor, indicando que talvez nem toda a modificação na Cor e Turbidez das amostras de água podem não ser apenas relacionados aos óxidos de Ferro dissolvidos ou em suspensão na fração argila e também por estarem mais correlacionados a banda 3 (vermelho) enquanto Turbidez e Cor, com a banda 8 (infravermelho).

Em relação aos resultados das modelagens realizadas, poucas variáveis foram eficientes, devido aos valores de R^2 ficarem abaixo de 0,30 em pelo menos um dos métodos testados. Dessa forma as variáveis com melhores resultados foram Turbidez, Cor, Temperatura e Fe.

Entretanto, de maneira geral os modelos preditivos foram classificados como satisfatórios, apesar dos valores de R^2 abaixo de 50 %. Uma vez que, valores entre 20 e 50% apresentam resultados mais informativos (Dobos, 2003; Malone et al., 2009).

A variável SST não foi bem modelada, devido aos valores de correlação dessa variável com as covariáveis serem diferentes das encontradas para Turbidez, Cor, Temperatura e Fe, valores próximos a 0,6 (Figura 20).

Na modelagem da Turbidez o algoritmo Ranger apresentou, melhores resultados. O Random Forest (RF) também obteve valores expressivos ao modelar a Turbidez, no entanto,

apresentou dispersão superior ao Ranger. O modelo que apresentou melhor desempenho na predição das variáveis químicas (Fe e Al) foi o RF (Figura 19).

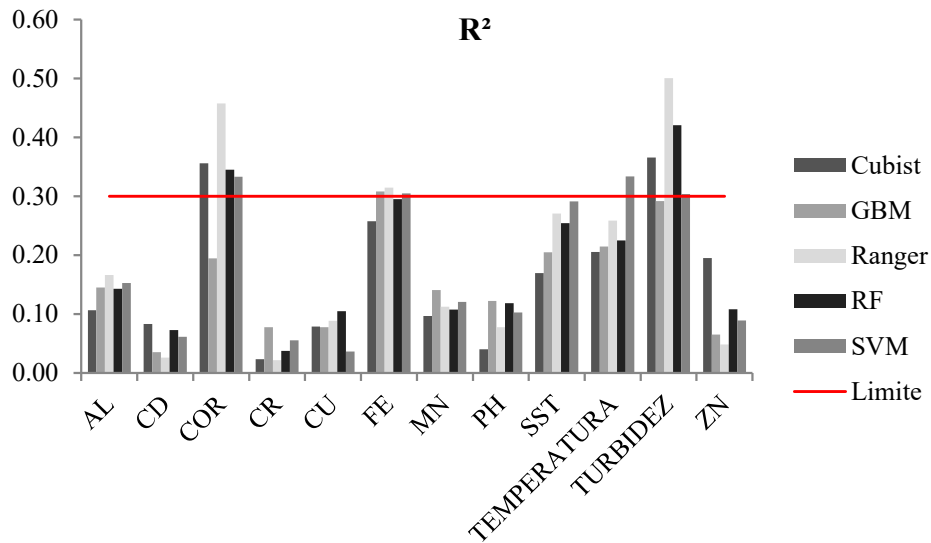


Figura 20: Valores médios para o coeficiente de determinação linear (R^2) ao utilizar diferentes modelos para determinação da regressão. Onde AL significa Alumínio dissolvido; CD Cádmio dissolvido; COR, a Cor aparente; CR, Cromo dissolvido; CU, Cobre dissolvido; Fe, Ferro dissolvido; Mn, Manganês dissolvido; pH, Potencial Hidrogeniônico; SST, Sólidos Suspensos Totais; Turbidez, a Turbidez da água; Zn, Zinco dissolvido.

Os parâmetros físicos foram melhor modelados quando comparados com os parâmetros químicos. Parâmetros químicos Fe e Al foram os que apresentaram melhor desempenho quando comparados com outros parâmetros químicos. Isso porque tiveram elevada correlação como por exemplo com o índice FE_OX, acima de 0,6.

Os resultados do RMSE e MAE para Cor, SST e Turbidez (Figuras 21 e 22) indicam maior variabilidade para esses parâmetros, erro na validação de 60 mg Pl l⁻¹ para Cor, em 20 mg l⁻¹ para o SST e em aproximadamente 50 NTU para Turbidez, logo, mesmo com o melhor ajuste desses parâmetros, o erro associado a validação também é maior, devido ao gradiente de variabilidade dessas variáveis. Dentre os metais dissolvidos o Fe, foi o que teve valores elevados de RMSE e MAE.

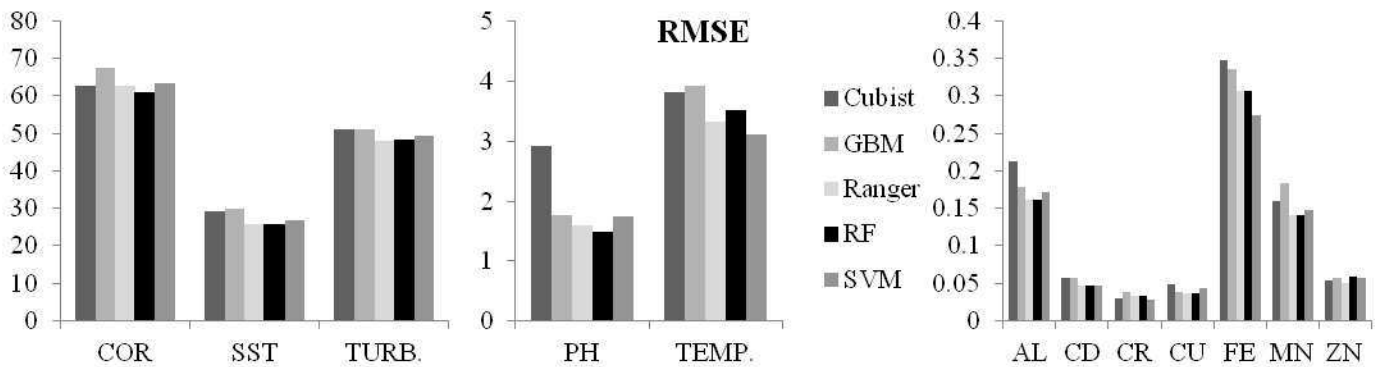


Figura 21: Raiz do erro médio quadrático (RMSE) da validação para os dados. Onde AL significa Alumínio dissolvido; CD Cádmio dissolvido; COR, a Cor aparente; CR, Cromo dissolvido; CU, Cobre dissolvido; Fe, Ferro dissolvido; Mn, Manganês dissolvido; pH, Potencial Hidrogeniônico; SST, Sólidos Suspensos Totais; Turb., a Turbidez da água; Temp., a Temperatura da amostra; e Zn, Zinco dissolvido.

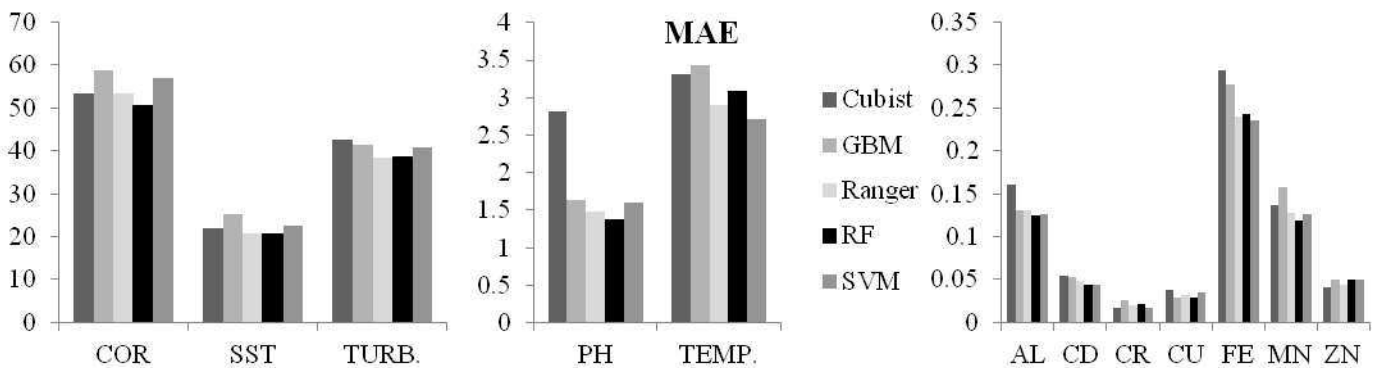


Figura 22: Erro absoluto médio (MAE) da validação para os dados.

A Turbidez obteve melhores resultados na validação dos cinco modelos utilizados. Dentre eles o Ranger apresentou coeficiente de determinação linear (R^2) com melhores resultados entre eles, com mediana acima de 0,5 de R^2 . O resultado do Ranger foi o que em média a predição ficou mais próximo da validação (Figura 23). Modelando o estoque de carbono em diferentes profundidades para o Brasil inteiro Gomes et al. (2019) encontraram valores de R^2 de 0,33, colaborando no avanço de pesquisas na distribuição geográfica de carbono por meio do aprendizado de máquina.

Os cinco modelos apresentaram resultados semelhantes de dispersão, sendo que a mediana para o SVM foi mais próxima do 3º Quartil, ou seja, mais distante da média, com alta frequência de valores abaixo da média.

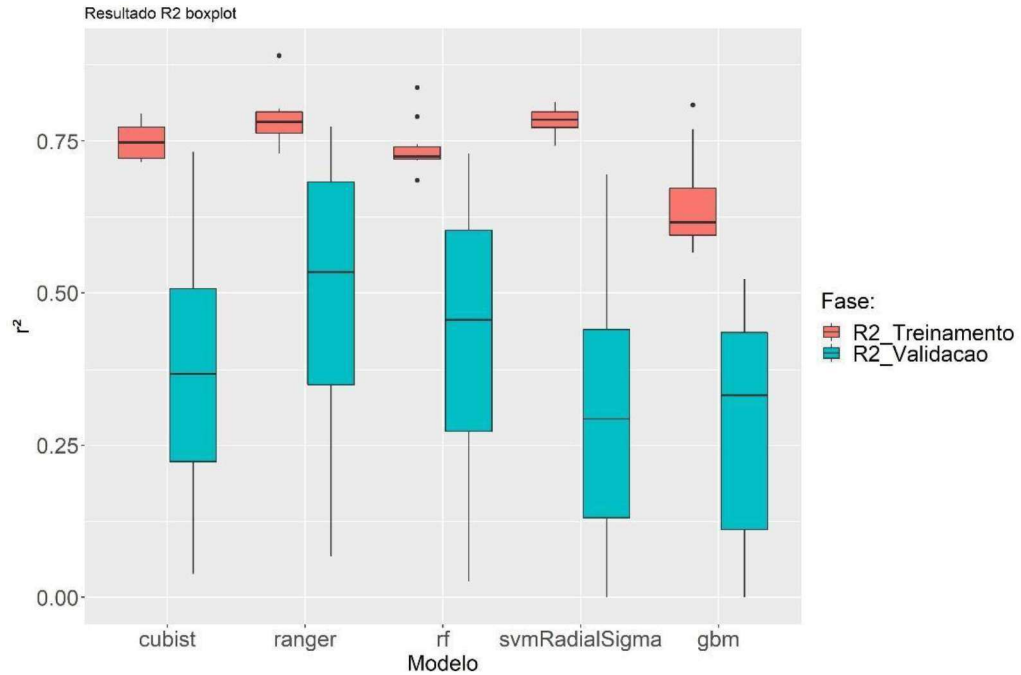


Figura 23: R^2 para os diferentes modelos aplicados no treinamento e na validação para a variável Turbidez.

Já para a Raiz Quadrática do Erro Médio (RMSE) os modelos GBM e SVM apresentaram menor dispersão dos dados, sendo o Cubist com maior variabilidade. Ranger e RF tiveram resultados semelhantes na validação, com médias semelhantes (Figura 24).

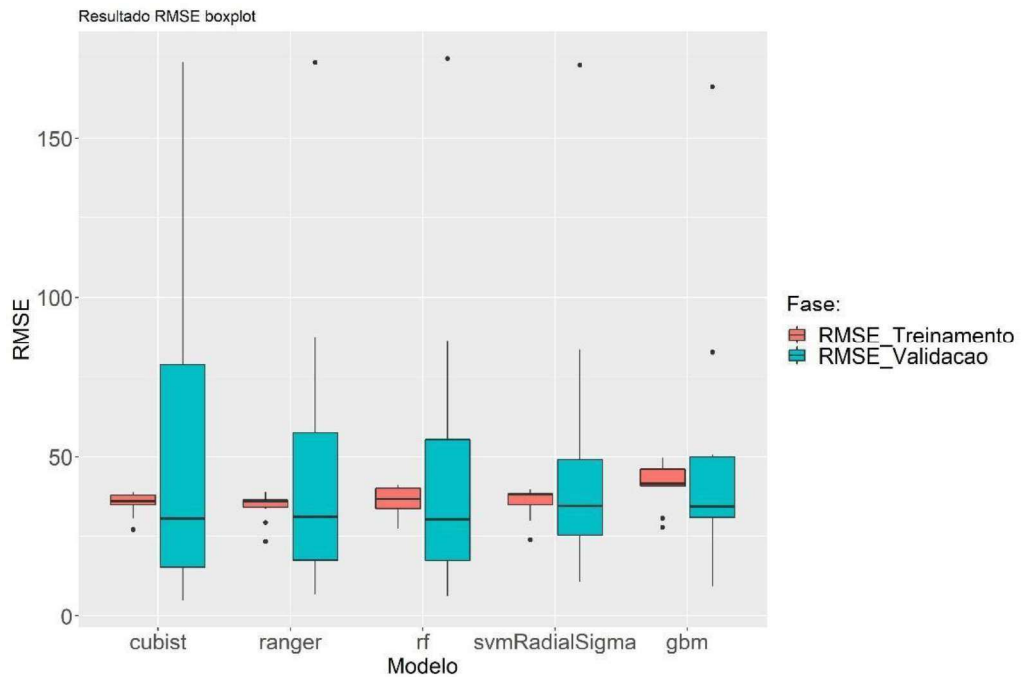


Figura 24: RMSE para os diferentes modelos aplicados no treinamento e na validação para a variável Turbidez.

O Erro Absoluto Médio (MAE) repetiu o comportamento do RMSE, como é uma métrica menos sensível aos outliers, o gradiente entre os valores máximos e mínimos foram reduzidos (Figura 25).

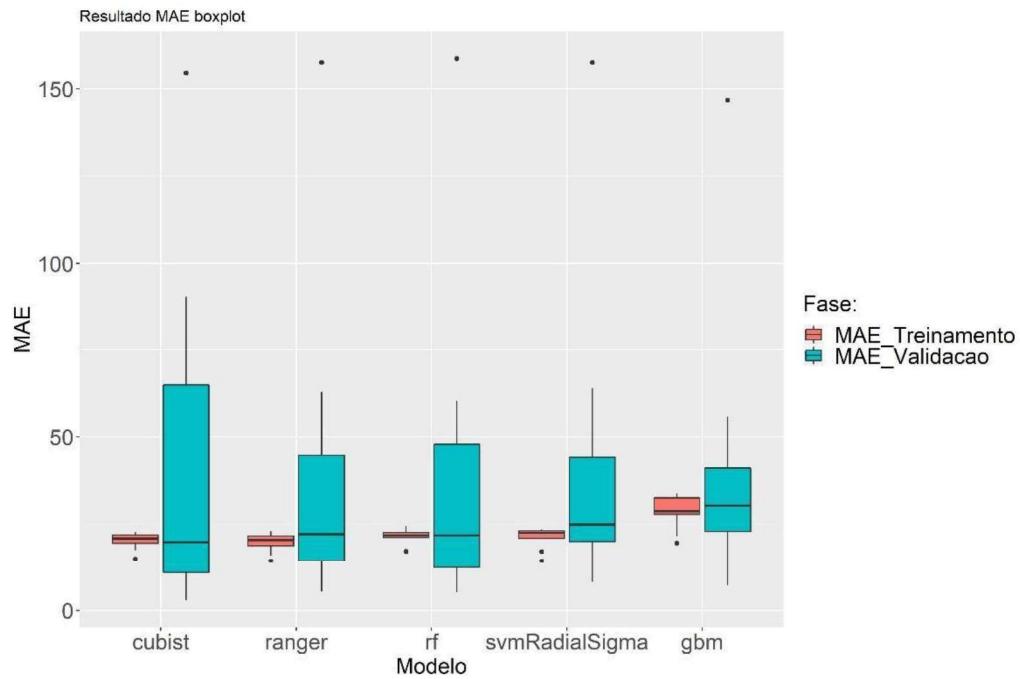


Figura 25: MAE para os diferentes modelos aplicados no treinamento e na validação para a variável Turbidez.

A modelagem do parâmetro Cor foi semelhante ao da Turbidez, com o modelo Ranger se destacando dos demais na validação. O RF apresentou os piores valores de validação, diferente de quando foi modelado para Turbidez (Figura 26).

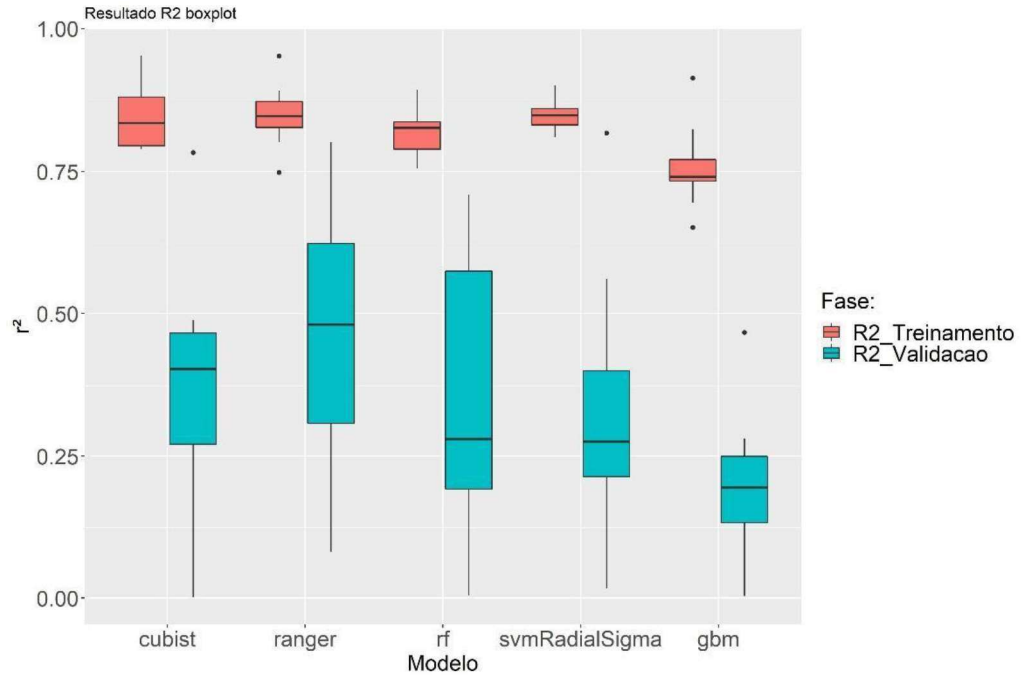


Figura 26: R^2 para os diferentes modelos aplicados no treinamento e na validação para a variável Cor aparente.

O RMSE da validação ao modelar a Cor aparente foi semelhante para os modelos Ranger e RF na validação, sendo semelhantes na dispersão dos resultados, entretanto, na predição para essa variável o RMSE do RF teve maior variabilidade (Figura 27).

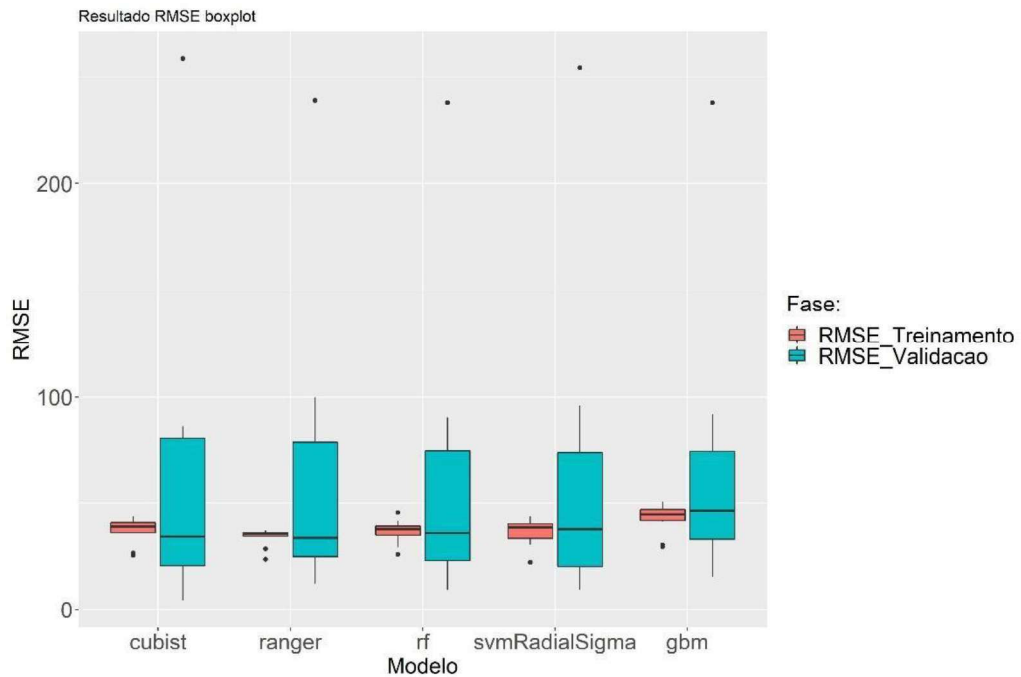


Figura 27: RMSE para os diferentes modelos aplicados no treinamento e na validação para a variável Cor aparente.

O MAE para o modelo Ranger volta a se destacar por ter média menor, menor dispersão dos dados e pela mediana próxima ao 1º Quartil, o que representa alta frequência dos valores abaixo da média (Figura 28).

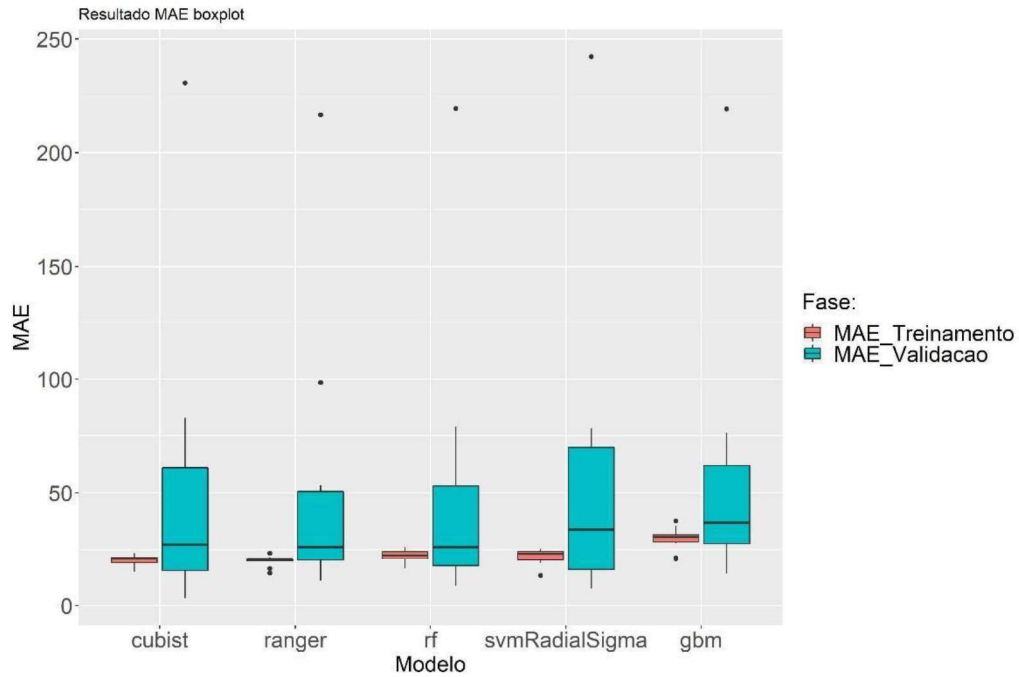


Figura 28: MAE para os diferentes modelos aplicados no treinamento e na validação para a variável Cor.

Para o parâmetro Fe três modelos tiveram resultados acima de 0,30 de R^2 em média (GBM, Ranger e SVM). O resultado do Ranger foi superior, por conta da assimetria positiva, em que os valores da média são superiores ao da mediana (Figura 29).

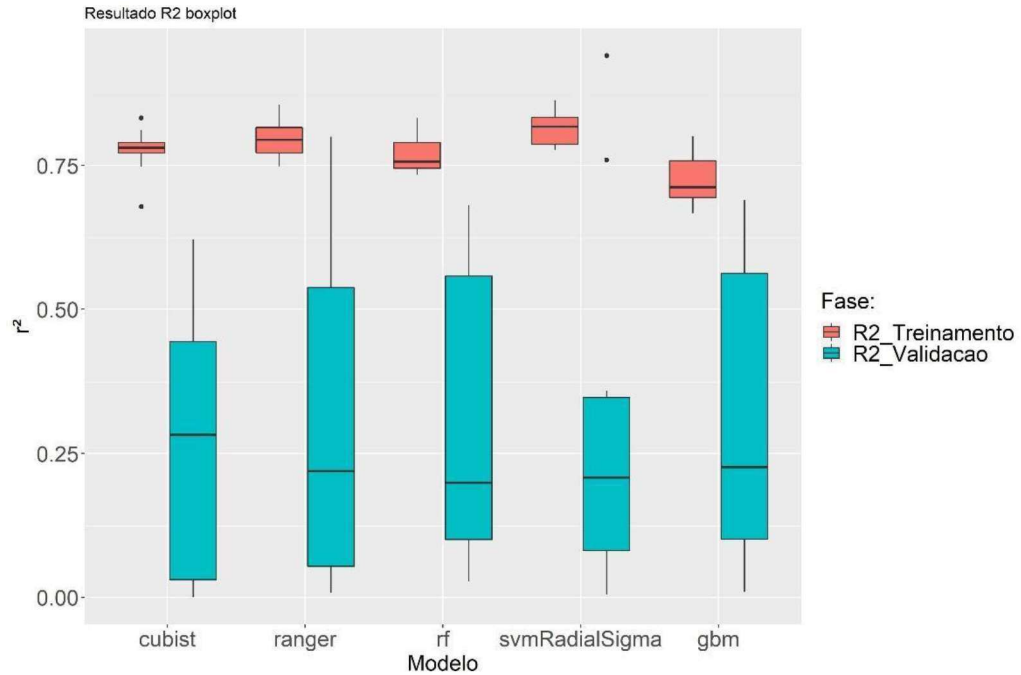


Figura 29: R^2 para os diferentes modelos aplicados no treinamento e na validação para a variável Fe.

Outro ponto que deve ser levado em consideração, e que demonstra melhor resultado do Ranger está no valor baixo para RMSE comparado com os outros modelos, sendo superior apenas ao SVM (Figura 30).

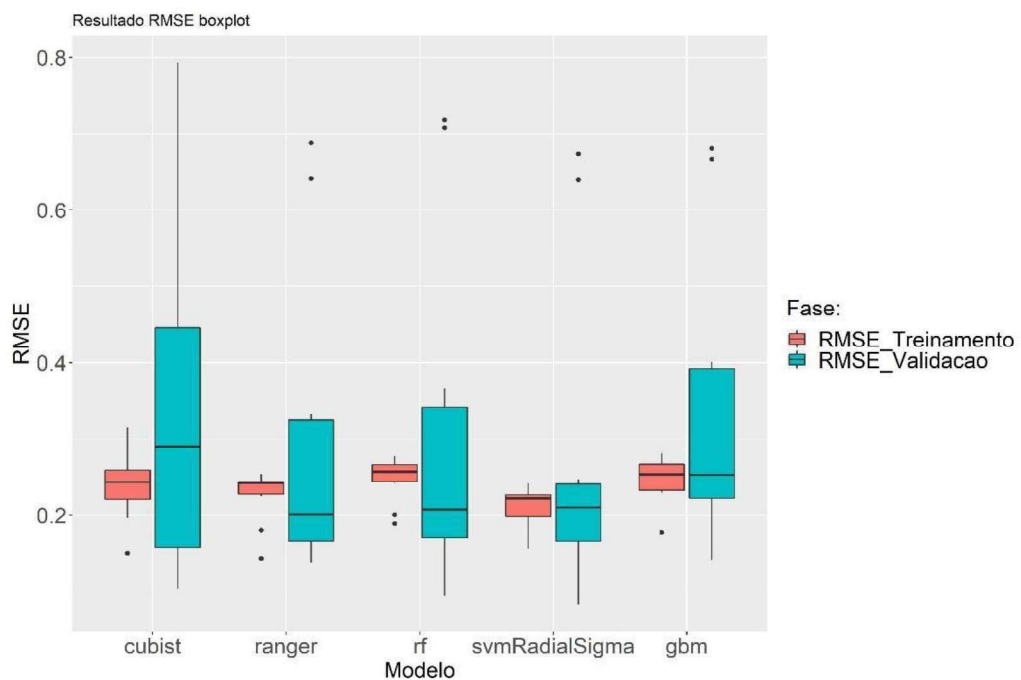


Figura 30: RMSE para os diferentes modelos aplicados no treinamento e na validação para a variável Fe.

O resultado do RMSE, repete no MAE, sendo novamente o Ranger inferior aos outros modelos que tiveram melhor R^2 (Figura 31).

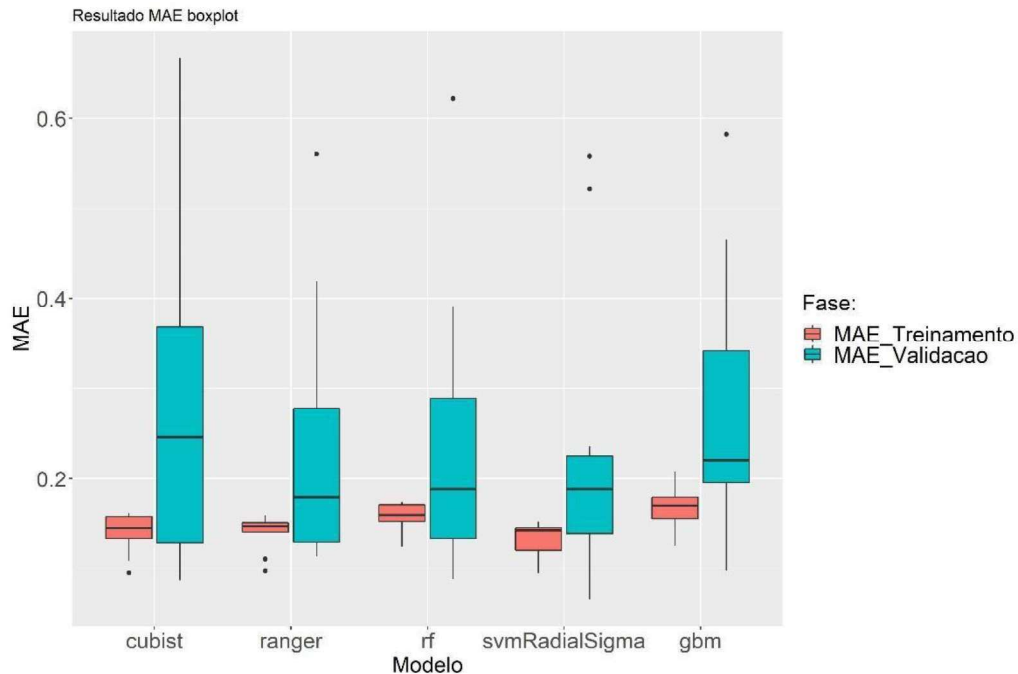


Figura 31: MAE para os diferentes modelos aplicados no treinamento e na validação para a variável Fe.

Valores de R^2 de 0,88 para Turbidez foram encontrados em rio na China ao comparar os valores de Turbidez com o comprimento de onda medido em campo (Wu et al., 2014). As metodologias usadas são diferentes nesse trabalho utilizamos a reflectância em uma célula de 10 m, (Wu et al., 2014) trabalhou com dados pontuais por espectroradiômetro, o que reduz a interferência da reflectância de outros elementos que não são o objeto de estudo.

Utilizando os valores de Turbidez em base logarítmica (Bustamante et al., 2009) obtiveram R^2 0,78 a modelar esses valores com a reflectância da banda 3 (630-690 nm) do sensor ETM+ do sistema orbital Landsat 7.

Os baixos valores de R^2 encontrados na modelagem podem estar associados à resolução da imagem utilizada na predição e validação dos modelos, que apresentam pixel de 10 m, sendo que as larguras do rio nos pontos de coleta das amostras de água não passam de 23,8 m em média. De acordo Isidro et al. (2018) a largura do rio interfere na resposta espectral do pixel sendo que oito dos treze pontos possuem largura inferior 20 m e os pontos 1 e 8 são os únicos que possuem largura maior que 30 m.

Dessa forma, além da reflectância das águas os pixels podem pegar informações que não representam os cursos d'água e sim os leitos dos rios que incluem solo exposto e vegetação

(Maillard and Pinheiro Santos, 2008; Tarasov and Tutubalina, 2018; Wang et al., 2020; Zhao et al., 2014)

Outra observação importante quanto à modelagem de recursos hídrico está em relação à profundidade, uma vez que a reflectância da imagem pode não estar associada a refletância de superfície do curso d'água, mas sim a reflectância do assoalho do rio (INPE, 2019; Isidro et al., 2018).

As formas contínuas de visualização dos resultados das variáveis melhores modeladas podem ser visualizadas nas Figuras 32 e 33, com aproximação em pontos em duas seções, rio Carmo e rio Gualaxo do Norte. A variável Turbidez, teve valores parecidos no período seco e úmido no ponto 8 (cidade de Barra longa, após a foz do rio Gualaxo) entre 20 e 30 NTU. Já no ponto 13 (primeiro ponto a montante) diferiu com valores de 30 a 40 NTU no período seco e de 50 a 60 NTU no período úmido. Essa diferença pode ser associada a maior perda de solo e sedimento acumulado no ponto 13, quando comparado ao ponto 8 (Schaefer et al., 2016).

Já para o SST, a diferença entre o período seco do período úmido é mais notória com valores menores que 10 mg/L no período seco no ponto 8 e 20 a 25 mg/L no período chuvoso nesse mesmo ponto. No ponto 13 de 15 a 20 mg/L no período seco e de 25 a 30 mg/L no período chuvoso. Com isso, percebe-se que em geral o ponto 13, é o mais influenciado pela carga de rejeito que é erodido dos terraços e planícies fluviais (Schaefer et al., 2016; Silva et al., 2016).

A cor aparente e o Fe dissolvido, são as variáveis que mais se distingue entre o período seco e chuvoso em ambos os pontos (Figura 33). Como ambas as bacias naturalmente possuem elevada concentração de ferro, em função da litologia formada por rochas itabiríticas (Costa et al., 2003), somado ao rompimento da barragem que disponibilizou ainda mais ferro tanto nos terraços quanto nas calhas dos rios (Guerra et al., 2017; Marta-Almeida et al., 2016) aumentam em 0,4 mg/L tanto no período seco para o chuvoso em ambos os pontos avaliados. A cor aparente aumenta em 20 uC no ponto 8 e em até 30 uC no ponto 13.

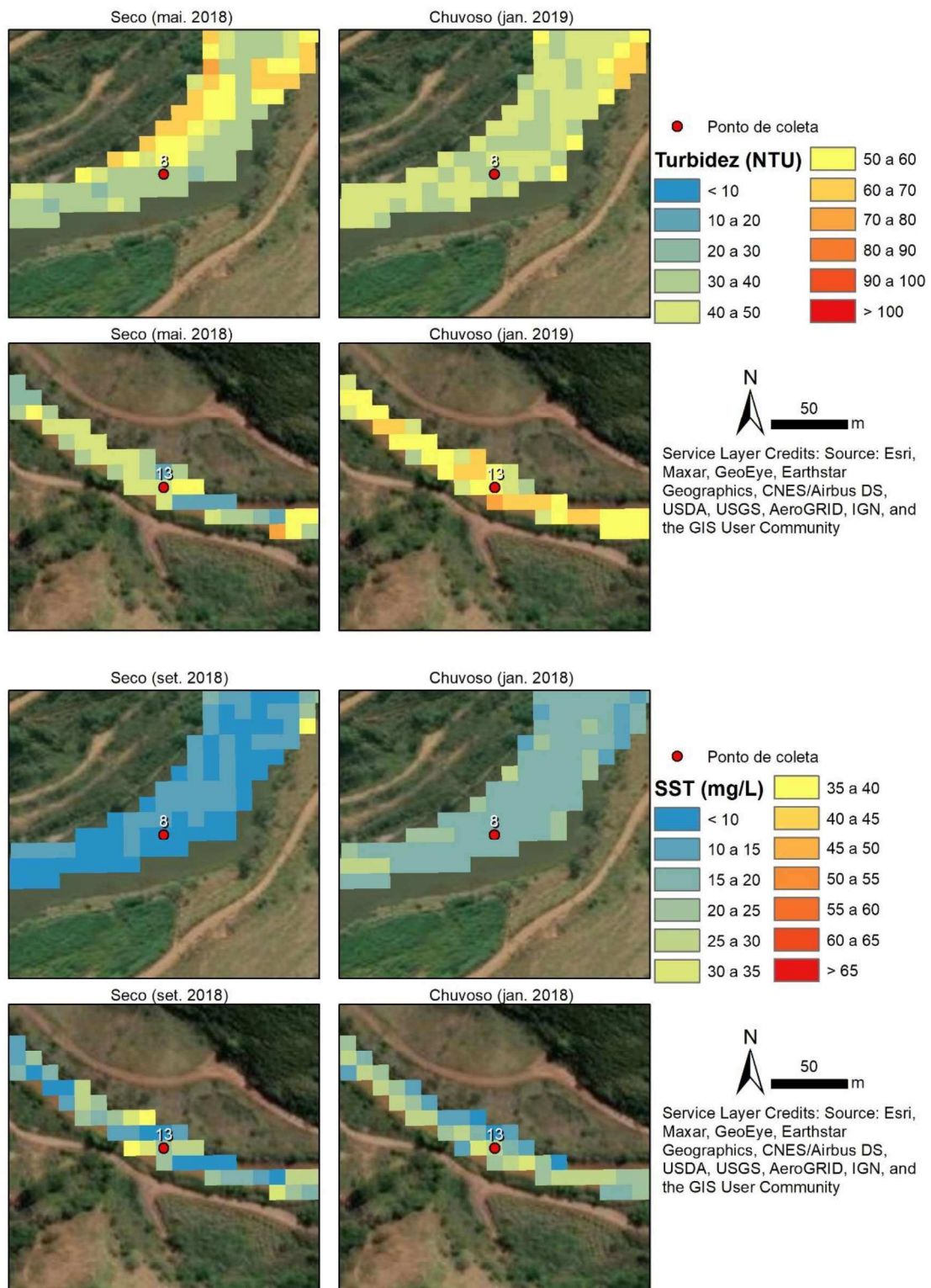


Figura 32: Espacialização dos resultados de Turbidez (NTU) e SST (mg/L) após a modelagem, sendo representado para as variáveis o resultado do modelo Ranger.

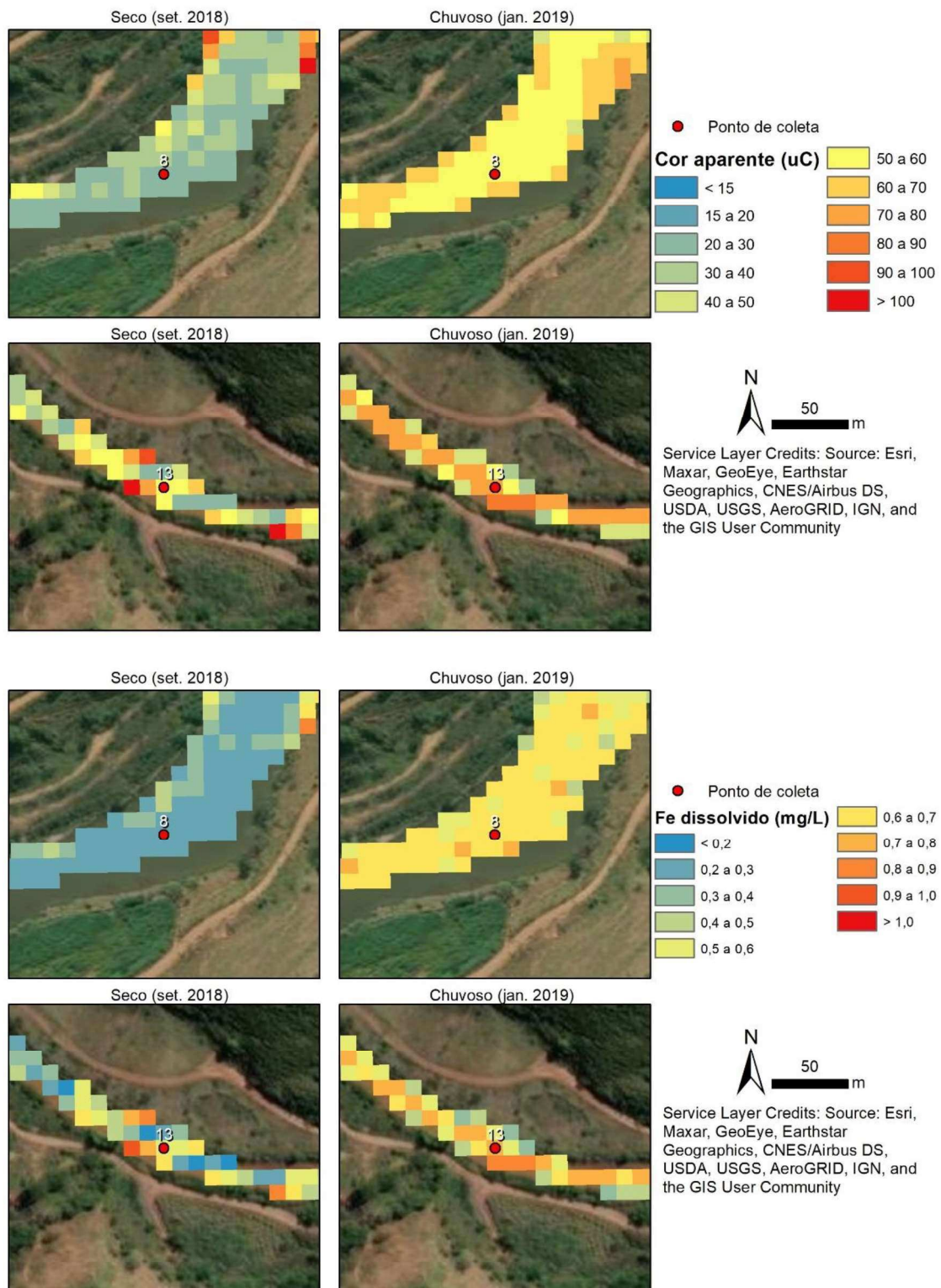


Figura 33: Espacialização dos resultados de Cor aparente (uC) e Fe dissolvido (mg/L) após a modelagem, sendo representado para as variáveis o resultado do modelo Ranger.

4 CONCLUSÕES

As relações de bandas desenvolvidas neste trabalho (bd8bd2 e bd4bd2) tiveram boa correlação acima de 0,75 com as variáveis Turbidez e Cor.

As variáveis Cor e Turbidez obtiveram melhores resultados em quatro dos cinco modelos testados, dentre eles Cubist, Ranger, RF e SVM.

A variável Temperatura obteve melhor resultado quando utilizado o modelo SVM.

Das variáveis químicas avaliadas, a que obteve melhor resultado na modelagem foi o Fe, onde dos cinco modelos testados obteve valores de R^2 acima de 0,30 em três, GBM, Ranger e SVM.

Valores abaixo de 0,30 de R^2 encontrados na validação da modelagem são decorrentes da resolução espacial das imagens Sentinel-2, originalmente com 10 m e a grande quantidade de pontos localizados em seções onde a largura dos cursos d'água são pouco maiores que essa resolução.

REFERENCIAS

- AllaouaSaadi, Campos, J.C.F., 2015. Geomorfologia do caminho da lama: contexto e consequências da ruptura da Barragem do Fundão (novembro 2015, Mariana–MG). *Arq. do Mus. História Nat. e Jard. Botânico - UFMG* 24, 63–103.
- Amari, S., Wu, S., 1999. Improving support vector machine classifiers by modifying kernel functions. *Neural Networks* 12, 783–789. [https://doi.org/10.1016/S0893-6080\(99\)00032-5](https://doi.org/10.1016/S0893-6080(99)00032-5)
- Bousbih, S., Zribi, M., Pelletier, C., Gorraab, A., Lili-Chabaane, Z., Baghdadi, N., Aissa, N. Ben, Mougenot, B., 2019. Soil texture estimation using radar and optical data from Sentinel-1 and Sentinel-2. *Remote Sens.* 11. <https://doi.org/10.3390/rs11131520>
- Breiman, L., 2001. Random forests. *Mach. Learn.* 45, 5–32. <https://doi.org/10.1023/A:1010933404324>
- Bustamante, J., Pacios, F., Díaz-Delgado, R., Aragonés, D., 2009. Predictive models of turbidity and water depth in the Doñana marshes using Landsat TM and ETM+ images. *J. Environ. Manage.* 90, 2219–2225. <https://doi.org/10.1016/j.jenvman.2007.08.021>
- Cardozo, F.A.C., Pimenta, M.M., Zingano, A.C., 2017. Métodos Construtivos De Barragens De Rejeitos De Mineração – Uma Revisão. *Holos* 8, 77. <https://doi.org/10.15628/holos.2016.5367>
- Carmo, Flávio Fonseca do, Kamino, L.H.Y., Junior, R.T., Campos, I.C. de, Carmo, Felipe Fonseca do, Silvino, G., Castro, K.J. da S.X. de, Mauro, M.L., Rodrigues, N.U.A., Miranda, M.P. de S., Pinto, C.E.F., 2017. Fundão tailings dam failures: the environment tragedy of the largest technological disaster of Brazilian mining in global context. *Perspect. Ecol. Conserv.* 15, 145–151. <https://doi.org/10.1016/j.pecon.2017.06.002>
- Clevers, J.G.P.W., 2007. Estimating grassland biomass using SVM band shaving of hyperspectral data. *Photogramm. Eng. Remote Sens.* 73, 1141–1148.
- Coluzzi, R., Imbrenda, V., Lanfredi, M., Simoniello, T., 2018. A first assessment of the Sentinel-2 Level 1-C cloud mask product to support informed surface analyses. *Remote Sens. Environ.* 217, 426–443. <https://doi.org/10.1016/j.rse.2018.08.009>
- Cortes, C., Vapnik, V.N., 1995. Support-Vector Networks. *Mach. Learn.* 20, 273–297.
- Costa, A.T., Nalini Júnior, H.A., Lena, J.C. de, Friese, K., Mages, M., 2003. Surface water quality and sediment geochemistry in the Gualaxo do Norte basin, eastern Quadrilátero Ferrífero, Minas Gerais, Brazil. *Environ. Geol.* 45, 226–235. <https://doi.org/10.1007/s00254-003-0870-6>

- De By, R., 2001. Principles of geographic information systems, ITC Educational Textbook Series 1. Enschede.
- Dobos, E., 2003. The application of remote sensing and terrain modeling to soil characterization. *Innov. Soil-Plant Syst. Sustain. Agric. Pract.* 328348.
- Donha, A.G., Souza, L.C. de P., Sugamoto, M.L., 2006. Determinação da fragilidade ambiental utilizando técnicas de suporte à decisão e SIG. *Rev. Bras. Eng. Agríc. Ambient. Bras. Eng. Agríc. Ambient.* 10, 175–181. <https://doi.org/10.1590/S1415-43662006000100026>
- EPA, 1992. Method 3005A - Acid digestion of waters for total recoverable or dissolved metals for analysis by FLAA or ICP spectroscopy.
- Evgeniou, T., Pontil, M., Elisseff, A., 2004. Leave One Out Error, Stability, and Generalization of Voting Combinations of Classifiers. *Mach. Learn.* 55, 71–97. <https://doi.org/10.1023/B:MACH.0000019805.88351.60>
- Feyisa, G.L., Meilby, H., Fensholt, R., Proud, S.R., 2014. Automated Water Extraction Index: A new technique for surface water mapping using Landsat imagery. *Remote Sens. Environ.* 140, 23–35. <https://doi.org/10.1016/j.rse.2013.08.029>
- Freitas, M.B., Freitas, C.M. de, 2006. A vigilância da qualidade da água para consumo humano: desafios e perspectivas para o Sistema Único de Saúde. *Cien. Saude Colet.* 10, 993–1004. <https://doi.org/10.1590/s1413-81232005000400022>
- Friedman, J.H., 2001. Greedy function approximation: A gradient boosting machine. *Ann. Stat.* 29, 1189–1232. <https://doi.org/10.2307/2699986>
- Gemmer, M., Becker, S., Jiang, T., 2004. Observed monthly precipitation trends in China 1951–2002. *Theor. Appl. Climatol.* 77, 39–45. <https://doi.org/10.1007/s00704-003-0018-3>
- Gitelson, A.A., Kaufman, Y.J., Merzlyak, M.N., 1996. Use of a green channel in remote sensing of global vegetation from EOS- MODIS. *Remote Sens. Environ.* 58, 289–298. [https://doi.org/10.1016/S0034-4257\(96\)00072-7](https://doi.org/10.1016/S0034-4257(96)00072-7)
- Golder Associates, 2017. Programa de Monitoramento Quali-Quantitativo Sistema de Água e Sedimentos.
- Gomes, L.C., Faria, R.M., de Souza, E., Veloso, G.V., Schaefer, C.E.G.R., Filho, E.I.F., 2019. Modelling and mapping soil organic carbon stocks in Brazil. *Geoderma* 340, 337–350. <https://doi.org/10.1016/j.geoderma.2019.01.007>
- Gong, W., Lin, Z., Chen, Y., Chen, Z., Shen, J., Zhang, H., 2018. Effect of waves on the dispersal of the Pearl River plume in winter. *J. Mar. Syst.* 186, 47–67. <https://doi.org/10.1016/J.JMARSYS.2018.05.003>

- Greenwell, B., Boehmke, B., Cunningham, J., 2019. Generalized Boosted Regression Models, CRAN Repository.
- Guerra, M.B.B., Teaney, B.T., Mount, B.J., Asunskis, D.J., Jordan, B.T., Barker, R.J., Santos, E.E., Schaefer, C.E.G.R., 2017. Post-catastrophe Analysis of the Fundão Tailings Dam Failure in the Doce River System, Southeast Brazil: Potentially Toxic Elements in Affected Soils. *Water, Air, Soil Pollut.* 228, 252. <https://doi.org/10.1007/s11270-017-3430-5>
- Henrich, V., Krauss, G., Gätze, C., Sandow, C., 2011. The Index Database.
- Hewson, R.D., Cudahy, T.J., Huntington, J.F., 2001. Geologic and alteration mapping at Mt fitton, South Australia, using ASTER satellite-borne data. *Int. Geosci. Remote Sens. Symp.* 2, 724–726. <https://doi.org/10.1109/igarss.2001.976615>
- Honeyborne, I., 2016. Profiling persistent tubercule bacilli from patient sputa during therapy predicts early drug efficacy, *BMC medicine*.
- Huete, A.R., 1988. A Soil-Adjusted Vegetation Index (SAVI). *Remote Sens. Environ.* 25, 295–309.
- Hu, J., Zhao, Y., Li, M., Liu, J., Wang, F., Weng, Q., Wang, X., Cao, D., 2020. Machine learning-based radiomics analysis in predicting the meningioma grade using multiparametric MRI. *Eur. J. Radiol.* 131, 109251. <https://doi.org/10.1016/j.ejrad.2020.109251>
- INPE, 2019. Introdução ao sensoriamento remoto de sistemas aquáticos: Princípios e aplicações, Sensoriamento remoto - Tecnologia. INPE, São José dos Campos.
- Isidro, C.M., McIntyre, N., Lechner, A.M., Callow, I., 2018. Quantifying suspended solids in small rivers using satellite data. *Sci. Total Environ.* 634, 1554–1562. <https://doi.org/10.1016/j.scitotenv.2018.04.006>
- Kalinowski, A., Oliver, S., 2004. ASTER Mineral Index Processing Manual Compiled by Aleks Kalinowski.
- Kauth, R.J., Thomas, G.S., 1976. Tasseled Cap - a Graphic Description of the Spectral-Temporal Development of Agricultural Crops As Seen By Landsat., in: *Machine Processing of Remotely Sensed Data* June. pp. 4B – 41.
- Kilham, N.E., Roberts, D., Singer, M.B., 2012. Remote sensing of suspended sediment concentration during turbid flood conditions on the Feather River, CaliforniaA modeling approach. *Water Resour. Res.* 48, 1–18. <https://doi.org/10.1029/2011WR010391>
- Kuhn, M., Wing, J., Weston, S., Williams, A., Keefer, C., Engelhardt, A., 2012. Caret: Classification and Regression Training [WWW Document]. <https://Cran.R->

Project.Org/Package=Caret.

- Lacaux, J.P., Tourre, Y.M., Vignolles, C., Ndione, J.A., Lafaye, M., 2007. Classification of ponds from high-spatial resolution remote sensing: Application to Rift Valley Fever epidemics in Senegal. *Remote Sens. Environ.* 106, 66–74. <https://doi.org/10.1016/j.rse.2006.07.012>
- Li, H., Liu, Q., 2008. Comparison of NDBI and NDVI as indicators of surface urban heat island effect in MODIS imagery. *Int. Conf. Earth Obs. Data Process. Anal.* 7285, 728503. <https://doi.org/10.1117/12.815679>
- Lima, R.P.C., Da Silva, D.D., Pereira, S.B., Moreira, M.C., Passos, J.B.M.C., Coelho, C.D., Elesbon, A.A.A., 2019. Development of an annual drought classification system based on drought severity indexes. *An. Acad. Bras. Cienc.* 91. <https://doi.org/10.1590/0001-3765201920180188>
- Long, C.M., Pavelsky, T.M., 2013. Remote sensing of suspended sediment concentration and hydrologic connectivity in a complex wetland environment. *Remote Sens. Environ.* <https://doi.org/10.1016/j.rse.2012.10.019>
- Maia, L.C., Pereira, A.R., 2017. Impactos no abastecimento de água decorrentes do rompimento da barragem de fundão: estudo de caso de Governador Valadares, in: Congresso ABES. pp. 1–6.
- Maillard, P., Pinheiro Santos, N.A., 2008. A spatial-statistical approach for modeling the effect of non-point source pollution on different water quality parameters in the Velhas river watershed - Brazil. *J. Environ. Manage.* 86, 158–170. <https://doi.org/10.1016/j.jenvman.2006.12.009>
- Malone, B.P., McBratney, A.B., Minasny, B., Laslett, G., 2009. Mapping continuous depth functions of soil carbon storage and available water capacity. *Geoderma* 154, 138–152. <https://doi.org/https://doi.org/10.1016/j.geoderma>
- Marta-Almeida, M., Mendes, R., Amorim, F.N., Cirano, M., Dias, J.M., 2016. Fundão Dam collapse: Oceanic dispersion of River Doce after the greatest Brazilian environmental accident. *Mar. Pollut. Bull.* 112, 359–364. <https://doi.org/10.1016/J.MARPOLBUL.2016.07.039>
- Martinez, J., Espinoza-villar, R., Armijos, E., Moreira, L.S., 2015. The optical properties of river and floodplain waters in the Amazon River Basin: Implications for satellite-based measurements of suspended particulate matter - Supplementary material. *J. Geophys. Res. Earth Surf.* 1, 1–11. <https://doi.org/10.1002/2014JF003404>.Received
- McFeeters, S.K., 1996. The use of the Normalized Difference Water Index (NDWI) in the

- delineation of open water features. *Int. J. Remote Sens.* 17, 1425–1432. <https://doi.org/10.1080/01431169608948714>
- Mechi, A., Sanches, D. Luiz, 2010. Impactos ambientais da mineração no estado de São Paulo. *Estud. Avançados* 24, 209–220.
- Meyer, D., 2019. The Interface to libsvm in package e1071, Support Vector Machines. <https://doi.org/10.4018/978-1-60960-557-5.ch007>
- Milanez, B., Losekann, C., 2016. Desastre no Vale do Rio Doce: Antecedentes, impactos e ações sobre a destruição., *Desastre no Vale do Rio Doce: antecedentes, impactos e ações sobre a destruição.* Folio Digital, Rio de Janeiro.
- Miller, J., Elvidge, C.D., Rock, B.N., Fremantle, J., 1990. An Airborne Perspective on Vegetation Phenology From the Analysis of AVIRIS Data Sets Over the Jasper Ridge Biological Preserve, in: *Geosciences and Remote Sensing Society/URSI*.
- Morgenstern, N.R., Vick, S.G., Viotti, C.B., Watts, B.D., 2016. *Fundão Tailings Dam Review Panel - Report on the Immediate Causes of the Failure of the Fundão Dam.*
- Mukherjee, N.R., Samuel, C., 2016. Assessment of the temporal variations of surface water bodies in and around Chennai using landsat imagery. *Indian J. Sci. Technol.* 9. <https://doi.org/10.17485/ijst/2016/v9i18/92089>
- Polidório, A.M., Imai, N.N., Tommaselli, A.M.G., 2004. Índice Indicador De Corpos D'Água Para Imagens Multiespectrais, in: *I Simpósio de Ciências Geodésicas e Tecnologias Da Geoinformação.*
- Queiroz, A.C.L., de Magalhães Cardoso, L.S., da Silva, S.C.F., Heller, L., Cairncross, S., 2012. Programa nacional de vigilância em saúde ambiental relacionada à qualidade da água para consumo humano (vigiagua): Lacunas entre a formulação do programa e sua implantação na instância municipal. *Saude e Soc.* 21, 465–478. <https://doi.org/10.1590/S0104-12902012000200019>
- Quinlan, J.R., 1992. Learning with continuous classes. *Mach. Learn.* 92, 343–348. <https://doi.org/10.1.1.34.885>
- Rowan, L.C., Mars, J.C., 2003. Lithologic mapping in the Mountain Pass, California area using Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) data. *Remote Sens. Environ.* 84, 350–366. [https://doi.org/10.1016/S0034-4257\(02\)00127-X](https://doi.org/10.1016/S0034-4257(02)00127-X)
- Rytky, S., 2020. Automating three-dimensional osteoarthritis histopathological grading of human osteochondral tissue using machine learning on contrast-enhanced micro-computed tomography. *Osteoarthr. Cartil.*
- Schaefer, C.E.G.R., Santos, E.E. dos, Souza, C.M. de, Neto, J.D., Filho, E.I.F., Delpupo, C.,

2015. Cenário histórico, quadro fisiográfico e estratégias para recuperação ambiental de Tecossolos nas áreas afetadas pelo rompimento da barragem do Fundão, Mariana, MG. *Arq. do Mus. História Nat. e Jard. Botânico - UFMG* 24, 104–135.
- Schaefer, C. E. G. R., Santos, E. E. dos, Fernandes Filho, E. I., & Assis, I. R. de. (2016). Paisagens de Lama: Os Tecossolos para recuperação ambiental de áreas afetadas pelo desastre da barragem do Fundão, em Mariana. *Boletim informativo da SBCS*, 1(1), 18–23.
- Silva, A.C., Cavalcante, L.C.D., Fabris, J.D., Júnior, R.F., Barral, U.M., Farnezi, M.M. de M., Viana, A.J.S., Ardisson, J.D., Fernandez-Outon, L.E., Lara, L.R.S., Stumpf, H.O., Barbosa, J.B.S., Silva, L.C. da, 2016. Chemical, mineralogical and physical characteristics of a material accumulated on the river margin from mud flowing from the collapse of the iron ore tailings dam in Bento Rodrigues, Minas Gerais, Brazil. *Rev. Espinhaço* 5, 44–53.
- Tarasov, M.K., Tutubalina, O. V., 2018. Estimating the Water Turbidity in the Selenga River and Adjacent Waters of Lake Baikal Using Remote Sensing Data. *Izv. - Atmos. Ocean Phys.* 54, 1353–1362. <https://doi.org/10.1134/S0001433818090372>
- Toming, K., Kutser, T., Laas, A., Sepp, M., Paavel, B., Nõges, T., Ni₂ges, T., 2016. First experiences in mapping lakewater quality parameters with sentinel-2 MSI imagery. *Remote Sens.* 8, 1–14. <https://doi.org/10.3390/rs8080640>
- Tucker, C.J., 1979. Red and photographic infrared linear combinations for monitoring vegetation. *Remote Sens. Environ.* 8, 127–150. [https://doi.org/10.1016/0034-4257\(79\)90013-0](https://doi.org/10.1016/0034-4257(79)90013-0)
- UFV, CETEC, UFLA, FEAM, 2010. Mapa de solos do Estado de Minas Gerais.
- Van Deventer, A.P., Ward, A.D., Gowda, P.M., Lyon, J.G., 1997. Using thematic mapper data to identify contrasting soil plains and tillage practices. *Photogramm. Eng. Remote Sensing* 63, 87–93.
- Vergara, V.M., Mayer, A.R., Kiehl, K.A., Calhoun, V.D., 2018. Dynamic functional network connectivity discriminates mild traumatic brain injury through machine learning. *NeuroImage Clin.* 19, 30–37. <https://doi.org/10.1016/j.nicl.2018.03.017>
- Wang, C., Li, W., Chen, S., Li, D., Wang, D., Liu, J., 2018. The spatial and temporal variation of total suspended solid concentration in Pearl River Estuary during 1987–2015 based on remote sensing. *Sci. Total Environ.* <https://doi.org/10.1016/j.scitotenv.2017.09.196>
- Wang, Z., Li, J., Lin, Y., Meng, Y., Liu, J., 2020. GrabRiver : Graph-Theory-Based River Width Extraction From Remote Sensing Imagery. *IEEE Geosci. Remote Sens. Lett.* 1–5. <https://doi.org/10.1109/LGRS.2020.3023043>
- Wright, M.N., Wager, S., Probst, P., 2020. A Fast Implementation of Random Forests, CRAN

Repository.

- Wu, J.L., Ho, C.R., Huang, C.C., Srivastav, A.L., Tzeng, J.H., Lin, Y.T., 2014. Hyperspectral sensing for turbid water quality monitoring in freshwater rivers: Empirical relationship between reflectance and turbidity and total solids. *Sensors (Switzerland)* 14, 22670–22688. <https://doi.org/10.3390/s141222670>
- Xu, H., 2006. Modification of normalised difference water index (NDWI) to enhance open water features in remotely sensed imagery. *Int. J. Remote Sens.* 27, 3025–3033. <https://doi.org/10.1080/01431160600589179>
- Zhang, K., Ross, M., Gann, D., 2016. Remote sensing of seasonal changes and disturbances in mangrove forest. *Ecosphere* 7, 1–23. <https://doi.org/10.1002/ecs2.1366>
- Zhang, Y., Liu, S., Yu, X., 2020. Longitudinal structural MRI analysis and classification in Alzheimer's disease and mild cognitive impairment. *Int. J. Imaging Syst. Technol.* 30, 421–433. <https://doi.org/10.1002/ima.22390>
- Zhao, D., Lv, M., Zou, X., Wang, P., Yang, T., An, S., 2014. What is the minimum river width for the estimation of water clarity using medium-resolution remote sensing images? *Water Resour. Res.* 50, 3764–3775. <https://doi.org/10.1002/2013WR015068>

CONSIDERAÇÕES FINAIS

Atividades de recuperação nas margens de rios afetados por rejeito de mineração colaboram para melhoria da qualidade da água, reduzindo principalmente o aumento elevado de turbidez em períodos chuvosos, em virtude da perda de solos de terraços e planícies aluviais.

Técnicas de sensoriamento remoto são promissoras para identificar qualitativamente valores de parâmetros de qualidade de água em locais onde não possuem monitoramento pontual.

Algoritmos de aprendizado de máquina do tipo de árvore de decisão são mais promissores que algoritmos de rede neural na modelagem de variáveis de qualidade de água. Contudo, estudos com dados de sensores remotos com melhor resolução espacial devem ser utilizados para melhorar a acurácia da modelagem.