



MODELAGEM PREDITIVA DA QUALIDADE DA ÁGUA EM BACIAS HIDROGRÁFICAS IMPACTADAS POR DESASTRES AMBIENTAIS

Luiz Claudio Alcântara da Silva - ⁽¹⁾ - Fábio Monteiro Cruz ⁽²⁾

RESUMO

Este estudo visa desenvolver uma metodologia para a predição de classes regionais de qualidade da água. A aplicabilidade do método é demonstrada através de um estudo de caso na bacia do Rio Doce, uma região profundamente impactada pelo desastre de Fundão. A metodologia utiliza séries temporais multivariadas (turbidez/nível) de seis estações. O modelo híbrido de Inteligência Artificial (IA) proposto visa prever a turbidez em três classes (Satisfatória, Moderada, Crítica), com uma abordagem de modelagem diferenciada para classes raras. Resultados iniciais confirmam o bom desempenho do modelo na previsão de todas as classes, validando a abordagem para apoio à tomada de decisão em sistemas de abastecimento público.

Palavras-chave: Sistemas de abastecimento. Segurança hídrica. Bacia hidrográfica. Machine learning.

1 INTRODUÇÃO

Emergências ambientais podem impactar diretamente os recursos hídricos e ecossistemas aquáticos. Alterações na qualidade da água, modificações na morfologia do sedimento do leito do rio e biodisponibilização de poluentes imobilizados no sedimento podem ocorrer (MILLER; GHISOLFI; BARROSO, 2023) Em 2015, o desastre de Fundão impôs sérios impactos a bacia do rio Doce, ao liberar 44 milhões de m³ de rejeitos de minérios de Ferro em uma pluma de contaminação que impactou cerca de 663 Km de mananciais, desde as cabeceiras do Rio Doce até sua foz no Estado do Espírito Santo (CUNHA RICHARD, DA *et al.*, 2020).

Além de impactos ambientais diretos de grande proporção como mortalidade de peixes, profundas alterações na qualidade da água, perda de biodiversidade, contaminação por metais e metaloides, fragmentação de habitats, destruição de áreas de preservação permanente e desestruturação físicas de rios de pequeno porte (IBAMA, 2015); sistemas de abastecimento

1 – Curso de Bacharelado em Engenharia Ambiental e Sanitária. IFMG *campus* Governador Valadares

2 - Engenheiro Ambiental, Doutor em Recursos Hídricos. IFMG *campus* Governador Valadares



de grandes centros, situado na bacia, como Governador Valadares e Colatina tiveram a operação de sistemas de tratamento e distribuição de água suspensos por 12 dias, em decorrência da elevadíssima concentração de sólidos, que inviabilizou as atividades, deixando uma população estima de 424.000 pessoas desassistidas de acesso a água, gerando um quadro crítico de calamidade pública (SÁNCHEZ *et al.*, 2018).

Esse cenário decorreu da dependência básica dessas cidades do Rio Doce, como fonte de abastecimento e demonstram a grande fragilidade da segurança hídrica, dessas regiões (ANA, 2016). Apesar de, posteriormente, com a recuperação da infraestrutura e a retomada das operações o abastecimento ter sido retomado ainda há grande incerteza sobre a capacidade desses sistemas em responder a aumentos da turbidez no Rio Doce, pois os efeitos do desastre ainda são imprevisíveis a médio e longo prazo e há fortes evidências de aumentos da turbidez na bacia em decorrência não apenas do elevado transporte de sedimentos na bacia, mas também pela ressuspensão da lama depositada ao longo de toda a extensão por onde a pluma se deslocou (CARVALHO, DE *et al.*, 2018).

Dessa forma, essa pesquisa pretende construir um modelo para previsão de classes regionais de qualidade da água em mananciais impactados por desastres ambientais baseado em modelos de inteligência artificial, como forma de apoiar a tomada de decisão em sistemas de abastecimento público e a garantir a segurança hídrica de comunidades vulneráveis aos efeitos da poluição hídrica.

2 DESENVOLVIMENTO

A área de estudo compreendeu o município de Governador Valadares (GV), situado na região do médio Rio Doce, constituindo o maior município da bacia, o centro mais densamente habitado e que sofreu os impactos mais severos do desabastecimento durante a fase aguda do desastre de Fundão (ENGEOPRPS, 2023). Foram obtidas as séries temporais dos parâmetros turbidez e nível do Rio Doce da estação RDO-08, situada em GV e de cinco outras estações distribuídas à montante. São elas: RDO-05, RDO-03, RGN-01M, RGN-06 e RCA-02M. Essas estações operam de forma automática e fazem parte da rede do *Programa de Monitoramento Quali-quantitativo Sistemático de Água e Sedimento*, implementado na bacia do Rio Doce



como parte das ações de reparação incluídas no termo de transação e ajustamento de conduta firmado após o desastre (RENOVA, 2017).

Os dados previamente consistidos pela fundação Renova foram filtrados, mantendo apenas aqueles considerados válidos. Das séries subsidiárias de cada parâmetro foram obtidas as séries parciais de turbidez máxima e nível máximo. Não foi feita limpeza das séries originais para remoção de *outliers* pois isso poderia alterar o comportamento distributivo dos dados e eliminar o valores extremos, que tem maior relevância nos objetivos dessa pesquisa. A série de variável turbidez máxima diária da estação RDO-08 foi discretizada em classes, para utilização de modelos de *machine learning* com fins de classificação.

Em decorrência da natureza do sistema modelado foi necessário defasar as séries das preditoras para encontrar o *lag* ideal, simulando o efeito do tempo de transporte da água e sedimentos em um contexto real de propagação da vazão em rios. Essa etapa usou como ferramenta a informação mútua, derivada da teoria da informação (SINGH, 1997), por permitir determinar estruturas de dependência entre par de variáveis sem basear-se em pressupostos de normalidade e relacionamento linear, o que ocorre com a abordagem clássica da correlação cruzada.

Definidos os *lags* ideais optou-se por utilizar apenas a preditora com maior valor de informação mútua de cada estação. Por fim, os dados das preditoras selecionadas foram defasados. A modelagem proposta baseou-se na construção de um modelo híbrido de predição de classes de turbidez máxima. Em decorrência da elevada assimetria da turbidez coerente com distribuições de cauda longa à direita a abordagem de modelagem baseou-se no pressuposto que a classe majoritária poderia ser modelada com melhor desempenho mesmo por modelos clássicos (*baseline*), como redes *Multilayer Perceptron* (MLP) ou árvores de decisão (DT) no *dataset* integral. Por outro lado, para as classes raras optou-se por gerar uma série derivada, eliminando as instâncias da classe majoritária, levando ao *desigm* de um problema de classificação binário. Essa abordagem buscou melhorar o desempenho na previsão final utilizando na medida em que explora as características dos modelos em situações diversas.

Toda a modelagem baseou-se no *framework* modular Tidymodels, desenvolvido para a linguagem R, usando *tune* para definição do melhor conjunto de parâmetros para cada modelo, *recipes* para preparação dos dados pré-modelagem, *parnisp* para configuração dos modelos, *rsample* para validação cruzada e *yardstick* para avaliação de desempenho.



As categorias adotadas para a discretização da variável resposta foram definidas a partir do teste de várias combinações diferentes. A princípio utilizou-se métodos baseados na distribuição dos dados como histograma e quebras naturais, priorizando quatro classes. Essa abordagem não trouxe resultados satisfatórios no treinamento dos modelos, pois inevitavelmente foram geradas classes raras com pequeno número de instâncias, que dificultaram a coleta de dados dessas classes no processo de validação cruzada.

De forma geral, os modelos treinados apesar de conseguir bons resultados na previsão das classes majoritárias tiveram desempenho pífio nas classes raras. A solução definitiva a esse problema foi reduzir o número de classes a três e definir os intervalos de classe com base tanto nos objetivos de apoio aos gestores de sistemas de abastecimento, quanto pela distribuição dos dados, de forma a reduzir o desbalanceamento das classes.

A classes de turbidez adotadas foram: satisfatória (0-50 UNT), moderada (50-200 UNT) e crítica (). Abaixo estão descritas cada classes, à luz da relação com a operação de sistemas de tratamento e distribuição de água (Quadro 1).

CLASSE DE TURBIDEZ	DESCRIÇÃO
SATISFATÓRIA	Condição ideal em que as ETAs operam em condições normais sem necessidade de ajustes nos protocolos de tratamento da água.
MODERADA	Condição que põe em alerta as equipes das ETAs. Há demanda pelo aumento da frequência de <i>jar test</i> .e eventualmente ajuste na dosagem de coagulantes.
CRÍTICA	Recursos de ajuste mais profundo na operação das ETAs devem ser preparados, como: coagulantes poliméricos. Aumento da frequência de <i>jar test</i> . Preparação de manobras para acionamento de fontes alternativas de água bruta. Em casos extremos deve-se considerar a suspensão da operação da ETA com objetivo de preservar os equipamentos.

4 CONCLUSÃO

A pesquisa realizada está em curso e tem avançado de forma satisfatória. Há bons resultados no ajuste dos modelos para previsão tanto da classe satisfatória (majoritária), quanto das demais classes, que possuem menor número de instâncias, e por consequência, do modelo híbrido. Apesar dos bons resultados ainda pretende-se investigar a aplicação da mesma abordagem utilizando um *dataset* derivado do original, com menor número de estações, com objetivo de tornar os recursos de previsão mais parcimoniosos e ampliar o horizonte de



previsão, uma vez que a abordagem descrita nesse trabalho é capaz de prever a classe de turbidez com apenas um dia de antecedência. Além disso, está em desenvolvimento uma etapa de aplicação de ferramentas que visam explicar os modelos, permitindo identificar as variáveis mais importantes, dentre outras coisas.

REFERÊNCIAS

ANA. **Encarte Especial sobre a Bacia do Rio Doce Rompimento da Barragem em Mariana/MG Conjuntura dos Recursos Hídricos no Brasil**. Brasília: [s.n.].

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

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

ENGECOPRPS. **Atualização do plano diretor de recursos hídricos da circunscrição hidrográfica do Rio Suaçuí**. Barueri: [s.n.].

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

MILLER, M. E.; GHISOLFI, R. D.; BARROSO, G. F. Remote sensing monitoring of mining tailings in the fluvial-estuarine-coastal ocean continuum of the Lower Doce River Valley (Brazil). **Environmental monitoring and assessment**, v. 195, n. 5, p. 542, 2023.

RENOVA. **Programa de monitoramento quali-quantitativo sistemático de água e sedimentos – PMQQS Relatório técnico dezembro 2017**. Belo Horizonte: [s.n.].

SÁNCHEZ, L. E.; ALGER, K.; ALONSO, L.; BARBOSA, F. A. R.; BRITO, M. C. W.; LAUREANO, F. V.; MAY, P.; ROESER, H.; KAKABADSE, Y. **Os impactos do rompimento da Barragem de Fundão: O caminho para uma mitigação sustentável e resiliente**. [s.l.: s.n.].

SINGH, V. P. The use of entropy in hydrology and water resources. **Hydrological Processes**, v. 11, n. 6, p. 587–626, 1997.