

TARCILA NEVES GENEROSO

**MODELAGEM HIDROLÓGICA VISANDO A MELHORIA DA QUALIDADE DE
DADOS DIÁRIOS DE VAZÃO EM ESTAÇÕES FLUVIOMÉTRICAS
CONVENCIONAIS LOCALIZADAS A JUSANTE DE RESERVATÓRIOS DE
REGULARIZAÇÃO**

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

Orientador: Demetrius David da Silva

Coorientadores: Lineu Neiva Rodrigues
Ricardo Santos Silva Amorim

**VIÇOSA – MINAS GERAIS
2021**

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

T

G326m
2021
Generoso, Tarcila Neves, 1988-
Modelagem hidrológica visando a melhoria da qualidade de dados diários de vazão em estações fluviométricas convencionais localizadas a jusante de reservatórios de regularização / Tarcila Neves Generoso. – Viçosa, MG, 2021.
1 tese eletrônica (114 f.): il. (algumas color.).

Orientador: Demetrius David da Silva.
Tese (doutorado) - Universidade Federal de Viçosa,
Departamento de Engenharia Agrícola, 2021.

Inclui bibliografia.

DOI: <https://doi.org/10.47328/ufvbbt.2022.080>

Modo de acesso: World Wide Web.

1. Vazões Defluentes. 2. Cuva-chave. 3. Aprendizado do máquinas. 4. Análise de regressão logística. I. Silva, Demetrius David da, 1966-. II. Universidade Federal de Viçosa. Departamento de Engenharia Agrícola. Programa de Pós-graduação em Engenharia Agrícola. III. Título.

CDD 22. ed. 551.483

Bibliotecário(a) responsável: Alice Regina Pinto CRB6 2523

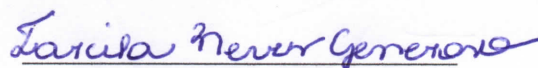
TARCILA NEVES GENEROSO

**MODELAGEM HIDROLÓGICA VISANDO A MELHORIA DA QUALIDADE DE
DADOS DIÁRIOS DE VAZÃO EM ESTAÇÕES FLUVIOMÉTRICAS
CONVENCIONAIS LOCALIZADAS A JUSANTE DE RESERVATÓRIOS DE
REGULARIZAÇÃO**

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

APROVADA: 27 de outubro de 2021

Assentimento:



Tarcila Neves Generoso
Autora



Demetrius David da Silva
Orientador

*A Deus.
A minha família.
Dedico*

AGRADECIMENTOS

A evolução do Ser é um processo que une aprendizado e reconhecimento. Reconhecimento a tudo e todos que contribuíram/contribuem para que o sucesso possa se materializar das mais diversas formas. A conclusão deste trabalho leva a um sentimento, dos mais nobres, que é o de gratidão. Gratidão sobre todas as coisas, a Deus, minha fonte de fé, esperança e confiança, ELE, que sabendo o quão difícil estava, agiu de diversas maneiras para que eu não desistisse do caminho.

Agradeço imensamente a minha família, meu sustentáculo, minha força, minha paz, minha fonte de amor, responsáveis pelo apoio incondicional. São o meu destino certo, minha maior riqueza.

Obrigada ao meu orientador Demetrius, que é uma inspiração de profissionalismo e ainda maior de humanidade. Um ser de luz, que além de todo o conhecimento técnico que detém ainda é capaz de reconhecer fraquezas demonstrando generosidade e confiança.

Aos meus coorientadores Lineu e Ricardo, que aceitaram o desafio e não pouparam esforços para contribuir com o desenvolvimento do trabalho, muito obrigada.

À banca examinadora, os professores, Elpídio Inácio, Celso Bandeira, Ricardo Amorim e Roberto Cecílio pela participação e contribuições ao trabalho.

A Universidade Federal de Viçosa e ao Departamento de Engenharia Agrícola, pela oportunidade, apoio e compreensão frente as dificuldades impostas, para a realização do doutorado.

Ao Conselho de Desenvolvimento Científico e Tecnológico (CNPq) e à Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (Capes) – Código de Financiamento 001, pela concessão da bolsa de estudos.

Um agradecimento especial aos amigos Erli e Daniel, por se disponibilizaram a ajudar num momento delicado e de grande dificuldade, mais uma vez obrigada.

Agradeço ainda aos demais amigos do CRRH pelo apoio, pelos momentos de descontração, pelo compartilhamento das experiências, pelas contribuições, por tornarem todo o processo mais leve e mais feliz. Obrigada pela convivência maravilhosa, em especial a Micael, Fernanda, Camila, Rayssa, Laura, Felipe, Roberto, Robson, Josiane, Lígia, Luane, Priscila, Lucas, Jéssica, Rafael Petruceli, Arnaldo, Clívia, Athur Kolling, Athur Calegário, Cássio, Cristiane e Laisi.

Por fim, meu sincero agradecimento a todos que de alguma forma contribuíram para a realização desse trabalho.

BIOGRAFIA

TARCILA NEVES GENEROSO, filha de Maria do Socorro Neves Generoso e Tarcísio Generoso Neto, nasceu em 15 de outubro de 1988, na cidade de São Francisco, estado de Minas Gerais, Brasil.

Iniciou, em 2007 o curso de graduação em Engenharia Ambiental na Universidade Estadual do Sudoeste da Bahia (UESB), graduando-se em outubro de 2012.

Em 2014 ingressou no Programa de Pós-Graduação em Engenharia Agrícola da Universidade Federal de Viçosa (UFV), área de concentração em Recursos Hídricos e Ambientais, a nível de mestrado, submetendo-se a defesa de dissertação em 2016.

Em 2016 ingressou no Programa de Pós-Graduação em Engenharia Agrícola da Universidade Federal de Viçosa (UFV), área de concentração em Recursos Hídricos e Ambientais, a nível de doutorado, submetendo-se à defesa de tese em 2021.

RESUMO

GENEROSO, Tarcila Neves, D.Sc., Universidade Federal de Viçosa, outubro de 2021. **Modelagem hidrológica visando a melhoria da qualidade de dados diários de vazão em estações fluviométricas convencionais localizadas a jusante de reservatórios de regularização.** Orientador: Demetrius David da Silva. Coorientadores: Lineu Neiva Rodrigues e Ricardo Santos Silva Amorim.

A estimativa de vazões em cursos d'água é essencial para que seja possível o estabelecimento de uma rede adequada de informações hidrológicas cuja finalidade é dar suporte para o planejamento e gestão dos recursos hídricos. O método do balanço hídrico é utilizado para monitoramento das vazões em situações em que se tem reservatórios de regularização em usinas hidrelétricas, de maneira que o cálculo da vazão defluente é realizado com base na vazão turbinada e na vazão vertida, em intervalos horários. Entretanto, na quase totalidade das estações fluviométricas da Rede Hidrometeorológica Nacional (RHN), a estimativa das vazões é feita com base no método da curva-chave, que embora seja amplamente utilizado apresenta como desvantagens a dependência da área da seção transversal de escoamento, que comumente é alterada nos ciclos de cheias e de secas, e as dificuldades operacionais para medições de vazões em eventos extremos máximos, fazendo-se necessário o uso de técnicas de extrapolação. Além disso, nas estações da RHN pertencentes à Agência Nacional de Águas e Saneamento Básico (ANA), as leituras são realizadas em apenas dois momentos do dia, às 7 e 17 horas, sendo a média entre os horários utilizada para a estimativa da vazão diária. Dessa forma, o objetivo geral do presente estudo foi, por intermédio da modelagem hidrológica, buscar técnicas e realizar análises que pudessem servir de base para a melhoria da qualidade de dados diários de vazão obtidos em estações fluviométricas convencionais localizadas a jusante de reservatórios de regularização de vazões, visando identificar e corrigir possíveis inconsistências. Todo o estudo foi realizado na bacia do rio Preto, adotando-se o reservatório da Usina Hidrelétrica de Queimado (UHE Queimado) como referência para a realização das análises. No primeiro capítulo é demonstrada uma aplicação dos modelos de aprendizado de máquinas e de regressão linear com o objetivo de realizar previsões de dados diários de vazão em estações fluviométricas convencionais localizadas a jusante de um reservatório de regularização. Os resultados indicaram que, além de serem encontradas inconsistências nas series de dados das estações convencionais analisadas, ainda foi observado que o modelo baseado em Redes Neurais (BRNN) foi o que apresentou melhor desempenho para as estações mais próximas do reservatório. Já o modelo de Regressão Linear Múltipla (RLM) foi o que

melhor se ajustou para a estação mais distante do reservatório da UHE Queimado. No segundo capítulo foram utilizadas técnicas estatísticas para analisar a relação entre as vazões afluentes e defluentes ao reservatório da UHE Queimado, assim como entre vazões defluentes ao reservatório da UHE Queimado e as vazões das estações fluviométricas convencionais da RHN localizadas a jusante da UHE Queimado. Também foram analisadas as tendências temporais nas séries de vazões mínimas, médias e máximas. Os resultados mostraram que a operacionalização do reservatório está fortemente relacionada ao comportamento das vazões afluentes, especialmente ao se considerar análises em períodos trimestrais, afetando também o escoamento nas estações fluviométricas a jusante. Além disso, foi constatada tendência de redução em todas as variáveis analisadas relacionadas às vazões afluentes e defluentes ao reservatório da UHE Queimado, bem como das estações a jusante do barramento, com exceção da vazão máxima na estação mais distante ao reservatório, que apresentou comportamento estacionário. No terceiro capítulo foi avaliado o comportamento das vazões horárias defluentes ao reservatório da UHE Queimado e, também, a utilização de modelos de regressão logística para comparar as séries diárias de vazões defluentes ao reservatório e da estação fluviométrica convencional localizada imediatamente a jusante do barramento, com o objetivo de identificar discrepâncias entre dados obtidos pelos métodos do balanço hídrico e da curva-chave. Os resultados mostraram que os valores médios de vazão defluente obtidos a partir a partir da média de duas leituras feitas às 7 e 17 horas (padrão da RHN) ou de 24 leituras (medições horárias feitas nas UHEs) foram muito próximos, considerando assim que os dados estimados a partir da média entre as duas leituras diárias apresentam boa representatividade da variabilidade das vazões na região. Também foi observado que, embora os métodos do balanço hídrico e da curva-chave tenham apresentado desempenho estatístico semelhante para a maioria dos dados considerados, com a regressão logística foi possível a identificação de dados atípicos nas séries históricas de vazões analisadas, associados principalmente aos métodos de extrapolação da curva-chave nos períodos chuvosos.

Palavras-chave: Vazões defluentes. Curva-chave. Aprendizado de máquinas. Regressão logística.

ABSTRACT

GENEROSO, Tarcila Neves, D.Sc., Universidade Federal de Viçosa, October, 2021. **Hydrological modeling aiming to improve the quality of daily flow data in conventional fluviometric stations located downstream of regularization reservoirs.** Advisor: Demetrius David da Silva. Co-advisers: Lineu Neiva Rodrigues and Ricardo Santos Silva Amorim

The estimation of flow in water courses is essential so that it is possible to establish an adequate network of hydrological information whose purpose is to support the planning and management of water resources. The water balance method is used to monitor flows in situations where there are regularization reservoirs in hydroelectric plants, so that the calculation of the effluent flow is performed based on the turbine flow and the spilled flow, at hourly intervals. However, in almost all fluviometric stations of the National Hydrometeorological Network (RHN), the estimation of flows is made based on the rating curve method, which, although widely used, has as disadvantages the dependence of the flow cross-sectional area, which it is commonly changed in flood and drought cycles and operational difficulties for flow measurements in maximum extreme events, making it necessary to use extrapolation techniques. In addition, at RHN stations belonging to the National Water and Basic Sanitation Agency (ANA), readings are taken at only two times of the day, at 7 am and 5 pm, with the average between the times used to estimate the daily flow. Thus, the general objective of this study was, through hydrological modeling, to seek techniques and perform analyzes that could serve as a basis for improving the quality of daily flow data obtained from conventional fluviometric stations located downstream of flows regularization reservoirs, aiming to identify and correct possible inconsistencies. The entire study was carried out in the Preto River basin, using the reservoir of the Queimado Hydroelectric Power Plant (UHE Queimado) as a reference for carrying out the analyses. The first chapter demonstrates an application of machine learning and linear regression models in order to perform daily flow data predictions in conventional fluviometric stations located downstream of a regularization reservoir. The results indicated that, in addition to inconsistencies being found in the data series of the analyzed conventional stations, it was also observed that the model based on Neural Networks (BRNN) presented the best performance for the stations closer to the reservoir. The Multiple Linear Regression (MRL) model was the best fit for the station farthest from the UHE Queimado reservoir. In the second chapter, statistical techniques were used to analyze the relationship between the inflows and outflows to the reservoir of the UHE Queimado, as well as between the outflows into the reservoir of the UHE Queimado and the flows of conventional fluviometric stations of RHN

located downstream of the UHE Queimado. Temporal trends in the series of minimum, average and maximum flows were also analyzed. The results showed that the operationalization of the reservoir is strongly related to the behavior of the inflows, especially when considering analyzes in quarterly periods, also affecting the flow in the fluviometric stations downstream. In addition, a downward trend was observed in all analyzed variables related to the inflows and outflows to the UHE Queimado reservoir, as well as the stations downstream of the dam, with the exception of the maximum flow in the station furthest from the reservoir, which presented a stationary behavior. In the third chapter, the behavior of the hourly flows effluent to the reservoir of the UHE Queimado was evaluated, as well as the use of logistic regression models to compare the daily series of effluent flows to the reservoir and the conventional fluviometric station located immediately downstream of the dam, with the objective of identifying discrepancies between data obtained by the water balance and the rating curve methods. The results showed that the outflows mean values obtained from the average of two readings taken at 7 and 17 hours (RHN standard) or 24 readings (hourly measurements taken at the UHEs) were very close, considering that the data estimated from the average between the two daily readings show good representation of the variability of flows in the region. It was also observed that, although the water balance and rating curve methods have presented similar statistical performance for most of the data considered, with logistic regression it was possible to identify atypical data in the historical series of analyzed flows, mainly associated with the methods of extrapolation of the rating curve in rainy periods.

Keywords: Effluent flows. Rating curve. Machine learning. Logistic regression.

SUMÁRIO

INTRODUÇÃO GERAL	12
REFERÊNCIAS	15
CAPÍTULO 1:	16
Predição de vazões diárias a jusante de reservatórios com regularização de vazões utilizando modelos de aprendizado de máquina e de regressão linear.....	16
1.1. INTRODUÇÃO	17
1.2. MATERIAL E MÉTODOS.....	19
1.2.1. Área de Estudo	19
1.2.2. Dados hidrológicos	21
1.2.3. Cenários e critérios para a predição de vazões diárias.....	24
1.2.4. Modelos de regressão linear e aprendizado de máquina utilizados.....	28
1.2.5. Validação dos modelos	30
1.2.6. Desempenho e análise estatística dos modelos	30
1.2.7. Importância das covariáveis	31
1.3. RESULTADOS E DISCUSSÃO.....	32
1.3.1. Dados obtidos para aplicação dos modelos.....	32
1.3.2. Desempenho dos modelos de aprendizado de máquinas e de regressão linear para Q1.....	36
1.3.3. Desempenho dos modelos de aprendizado de máquinas e de regressão linear para Q2.....	44
1.3.4. Desempenho dos modelos de aprendizado de máquinas e de regressão linear para Q3.....	48
1.4. CONCLUSÕES	52
1.5. REFERÊNCIAS.....	53
CAPÍTULO 2:	59
Análise do comportamento de vazões sob a influência de reservatório com regularização de vazões.....	59
2.1. INTRODUÇÃO	60
2.2. MATERIAL E MÉTODOS.....	62
2.2.1. Área de estudo	62
2.2.2. Dados hidrológicos	64
2.2.3. Análise exploratória da relação entre as vazões	67
2.2.4. Análises de tendência.....	68
2.3. RESULTADOS E DISCUSSÃO.....	72
2.3.1. Análise exploratória das vazões afluentes e defluentes	72
2.3.2. Análises de tendência.....	78

2.4. CONCLUSÕES	86
2.5. REFERÊNCIAS.....	87
CAPITULO 3:	91
Proposta metodológica para avaliar a estimativa de vazões pelos métodos do balanço hídrico e da curva-chave baseado no método de regressão logística	91
3.1. INTRODUÇÃO	92
3.2. MATERIAL E MÉTODOS.....	94
3.2.1. Área de estudo	94
3.2.2. Dados hidrológicos	96
3.2.3. Análise do comportamento horário das vazões defluentes	96
3.2.4. Comparação dos métodos para estimativa de vazão	97
3.3. RESULTADOS E DISCUSSÃO.....	102
3.3.1. Análise do comportamento horário das vazões defluentes ao reservatório UHE Queimado.....	102
3.3.2. Comparação dos métodos para estimativa de vazão	105
3.4. CONCLUSÕES	112
3.5. REFERÊNCIAS.....	112

INTRODUÇÃO GERAL

A determinação das vazões em cursos d'água pode ser feita de diferentes formas, considerando as especificidades apresentadas para cada tipo de situação. Quanto maior a precisão envolvida no método de estimativa dessas vazões maior é a eficácia na implementação das ações que envolvem a gestão e planejamento dos recursos hídricos. Nesse sentido, contribuições metodológicas vêm sendo desenvolvidas buscando aumentar a qualidade dos dados ou a forma como a vazão é obtida para os rios, o que pode promover grandes avanços quanto à distribuição, preservação e previsão de eventos hidrológicos (Junqueira et al., 2021; Farnoush e Maghrebi, 2021; Durga Rao et al., 2020).

Considerando as características do escoamento, a quantificação das vazões em cursos d'água pode ser feita de diferentes maneiras. Em situações em que se dispõe da presença de reservatórios de regularização em usinas hidrelétricas, o monitoramento normalmente é feito utilizando o método do balanço hídrico, no qual a vazão defluente é calculada a partir da vazão turbinada (relação entre potência, queda e vazão) e da vazão vertida medida (Hidalgo et al, 2012), devendo seus valores serem registrados em intervalos máximos de uma hora, conforme determinação estabelecida pela Resolução Conjunta ANA/ANEEL nº 03/2010.

Já em cursos d'água em condições naturais do escoamento, a estimativa da vazão normalmente é feita em estações fluviométricas convencionais com o uso do método da curva-chave ou curva de calibração, baseada na relação entre cota do nível de água e vazão. A curva-chave é calibrada através de medições de vazões e cota do nível de água em campo, o que torna essa estimativa fortemente dependente das características geométricas associadas a área da seção transversal do escoamento. Conforme explica Li et al. (2019), é um método trabalhoso, muitas vezes perigoso ou mesmo impossível de se realizar quando se tem eventos extremos máximos.

Embora o método da curva-chave seja amplamente utilizado em nível mundial, as estimativas de vazão podem apresentar problemas expressivos. Isso acontece em função do número limitado de pontos usados para calibração, sendo comum, portanto, a derivação de valores altos de vazão por extrapolação, o que pode produzir uma quantidade significativa de incerteza nas estatísticas e previsões de eventos críticos de enchentes (Manfreda, et al, 2020). Além disso, para a medida de vazão existe uma dependência da área da seção transversal de escoamento, que comumente é alterada nos ciclos de cheias e de secas e ainda os dados de vazão das estações fluviométricas convencionais, sob responsabilidade da ANA (Agência Nacional de Águas e Saneamento Básico), são obtidos a partir dos valores médios observados

às 7 e 17 horas, sendo assim, estimados a partir de uma média de apenas duas leituras ao longo do dia.

Em função das informações apresentadas em relação aos métodos do balanço hídrico e da curva-chave, é possível considerar que as estimativas das vazões defluentes em reservatórios de regularização poderiam apresentar menores imprecisões que aquelas realizadas em estações fluviométricas convencionais, uma vez que sua determinação envolve menores variações das características hidráulicas da seção de escoamento e sua determinação é feita considerando, necessariamente, a média de 24 leituras, isto é, em menores intervalos de tempo quando comparado ao método da curva-chave.

Neste contexto, na existência de reservatórios de regularização de vazões, séries de vazões defluentes poderiam ser utilizadas como apoio para análise de séries de vazão de estações fluviométricas convencionais localizadas a jusante de reservatórios, podendo proporcionar melhorias quanto às técnicas de previsão, preenchimento de falhas, compreensão das tendências do escoamento ou mesmo auxiliar na identificação de vazões discrepantes e/ou incorretas nessas seções, haja vista situações onde essa comparação seja aplicável.

Sendo assim, o objetivo geral do presente estudo foi, por intermédio da modelagem hidrológica, buscar técnicas e realizar análises que pudessem servir de base para a melhoria da qualidade de dados diários de vazão obtidos em estações fluviométricas convencionais localizadas a jusante de reservatórios de regularização de vazões, visando identificar e corrigir possíveis inconsistências.

No primeiro capítulo (Artigo 1) do presente trabalho, intitulado “Predição de vazões diárias a jusante de reservatórios com regularização de vazões utilizando modelos de aprendizado de máquina e de regressão linear” é avaliada a utilização de modelagem hidrológica objetivando a melhoria da previsão de dados diários de vazão em estações fluviométricas convencionais localizadas a jusante de reservatórios de regularização.

No segundo capítulo (Artigo 2), intitulado “Análise do comportamento de vazões sob a influência de reservatório com regularização de vazões”, focou-se em analisar as relações entre o comportamento das vazões afluentes e defluentes de um reservatório de regularização assim como entre as vazões defluentes e as vazões das estações fluviométricas convencionais posicionadas a jusante do barramento. Avaliou-se, ainda, as tendências de comportamento das vazões mínimas, médias e máximas em cada uma das seções estudadas.

Por fim, no terceiro capítulo (Artigo 3), intitulado “Avaliação da estimativa de vazões pelos métodos do balanço hídrico e da curva-chave com base na utilização do modelo de regressão logística” foi feita uma avaliação visando compreender o comportamento das vazões

horárias defluentes a um reservatório de regularização e apresentada uma metodologia, baseada na modelagem via regressão logística, para comparar e avaliar a estimativa das vazões diárias defluentes a um reservatório de regularização, estimadas pelo método do balanço hídrico, com as vazões diárias de uma estação fluviométrica localizada imediatamente a jusante, estimadas pelo método da curva-chave.

REFERÊNCIAS

- Durga Rao, K. H. V., Shravya, A. e Dadhwal, V. K. 2020. A novel method of satellite based river discharge estimation using river hydraulic geometry through genetic algorithm technique. *Journal of Hydrology*, v. 589, doi: <https://doi.org/10.1016/j.jhydrol.2020.125361>.
- Farnoush, H e Maghrebi, M. F. 2021. Application of surface velocimetry in estimating rating in compound channels. *Flow Measurement and Instrumentation*, doi: <https://doi.org/10.1016/j.flowmeasinst.2021.102049>.
- Hidalgo, I. G., Soares F. S.; Fontane, D. G., Gicogna, M. A. e Lopes, J. E. G. 2012. Procedimentos para consolidação de dados de usinas hidrelétricas baseados em informações registradas. *Revista Brasileira de Recursos Hídricos*, v. 17, n.1, p.9-21.
- Junqueira, A. M., Mao, F., Mendes, T. S. G., Simões, J. C., Balestieri, J. A. P., Hannah, D. M. 2021. Estimation of river flow using CubeSats remote sensing, *Science of The Total Environment*, v. 788, doi: <https://doi.org/10.1016/j.scitotenv.2021.147762>.
- Li, W., Liao, Q. e Ran, Q. 2019. Stereo-imaging LSPIV (SI-LSPIV) for 3D water surface reconstruction and discharge measurement in mountain riverflows. *Journal of Hydrology*, v.578, doi: <https://doi.org/10.1016/j.jhydrol.2019.124099>.
- Manfreda, S., Pizarro, A., Moramarco, T., Cimorelli, L., Pianese, D. e Barbeta, S. 2020. Potential advantages of flow-area rating curves compared to classic stagedischarge-relations. *Journal of Hydrology*, v. 585, doi: <https://doi.org/10.1016/j.jhydrol.2020.124752>.

CAPÍTULO 1:

Predição de vazões diárias a jusante de reservatórios com regularização de vazões utilizando modelos de aprendizado de máquina e de regressão linear

RESUMO: Estações fluviométricas convencionais podem apresentar estimativas diárias de vazão inconsistentes, em decorrência da quantidade de leituras realizadas pelo método da curva-chave, ao longo da série histórica. Para estações convencionais localizadas a jusante de reservatórios de regularização, a utilização das séries históricas de vazão defluente ao reservatório pode servir de referência para o aprimoramento desses registros, já que os dados diários são estimados pelo método do balanço hídrico, cujo registro é feito a partir da média de 24 leituras. Diante disso, o objetivo desse trabalho foi ajustar modelos de aprendizado de máquinas e de regressão linear para fins de previsão de dados diários de vazão em estações fluviométricas convencionais localizadas a jusante de reservatórios de regularização. Os resultados encontrados, além de indicarem a presença de inconsistências nas series de dados das estações convencionais analisadas, ainda mostraram que para as estações mais próximas ao reservatório o modelo baseado em Redes Neurais (BRNN) foi o que apresentou melhor desempenho para as estações mais próximas do reservatório e para aquela mais distante, o modelo de Regressão Linear Múltipla (RLM) foi o que melhor se ajustou.

Palavras-chave: preenchimento de falhas, vazões defluentes, modelagem

ABSTRACT: Conventional fluviometric stations may present inconsistent daily flow estimates, due to the number of readings performed by the rating curve method, throughout the historical series. For conventional stations located downstream of regularization reservoirs, the use of historical series of effluent flow to the reservoir can serve as a reference for the improvement of these records, as daily data are estimated by the water balance method, whose record is made from the average of 24 readings. Therefore, the objective of this study was to adjust machines learning and linear regression models for the purposes of predicting daily flow data in conventional fluviometric stations positioned downstream of regularization reservoirs. The results found, in addition to indicating the presence of inconsistencies in the data series of the analyzed stations, still available that for stations closer to the reservoir of the model based on Neural Networks (BRNN) was the one that presented the best performance for stations

nearest to the reservoir and for the farthest station, the Multiple Linear Regression (MRL) model was the best fit.

Keywords: fault filling, effluent flows, modeling

1.1. INTRODUÇÃO

A vazão de cursos d'água constitui-se em uma das variáveis hidrológicas mais complexas de serem obtidas em nível de campo, pois demanda procedimentos muitas vezes demorados, perigosos e de alto custo (Corato et al., 2014). Sendo assim, na grande maioria das vezes faz-se uso de métodos indiretos para a estimativa das vazões, sendo a curva-chave ou curva de calibração a metodologia mais usualmente empregada (Oliveira et al., 2016).

No Brasil, a Agência Nacional de Águas e Saneamento Básico (ANA) é responsável pela manutenção da Rede Hidrometeorológica Nacional (RHN) que, por sua vez, utiliza o método da curva-chave para a estimativa das vazões nas estações fluviométricas. Nessas estações, é necessário que uma pessoa faça diariamente a leitura de cota do nível de água na régua linimétrica. Nas estações sob responsabilidade da RHN, a leitura é feita diariamente em dois períodos, às 7 e 17 horas, e os valores anotados de cota são posteriormente convertidos em vazão.

Nas situações em que há presença de usinas hidrelétricas com reservatórios de regularização, a Resolução Conjunta ANEEL/ANA nº 03, de 10 de agosto de 2010, determina que todas as estações hidrométricas com monitoramento pluviométrico, linimétrico e fluviométrico sejam automatizadas e telemetrizadas. Os dados coletados por essas estações devem ser registrados em intervalo máximo de uma hora.

Nesses casos, para o cálculo da quantidade de água que entra e sai de um reservatório, em certo período de tempo, normalmente é usado o método do balanço hídrico, sendo que para a determinação das vazões defluentes é necessária a obtenção de duas outras variáveis, a vazão turbinada, retirada do reservatório pelo conduto forçado e que passa pelas turbinas produzindo energia, e a vazão vertida, que não passa pelas turbinas e, portanto, não gera energia (Molina, 2016).

Em decorrência da baixa frequência dos horários de medições das vazões nas estações da RHN, comportamentos importantes como ondas de cheia, captações diversas, como por exemplo, para fins de irrigação e regime de operação de reservatórios podem não ser identificados, quando ocorridos em horários diferentes daqueles das medições nas estações

fluviométricas convencionais. Já a coleta de dados de vazão em intervalos horários permite que 24 leituras distribuídas ao longo do dia sejam realizadas, o que melhora a estimativa do comportamento médio das vazões ao longo tempo, proporcionando a minimização de erros causados pela variabilidade do escoamento.

A estimativa de vazões nas estações fluviométricas convencionais pode apresentar inconsistências em função da limitação dos horários das leituras, resultando em dados não representativos do regime de variação de vazões no curso d'água. Além disso, quando se trata de dados hidrológicos é comum em países em desenvolvimento, como no Brasil, a presença de falhas, que podem ser atribuídas a vários fatores, como a interrupção de medições devido a falhas do equipamento, efeito de fenômenos naturais extremos ou fatores induzidos pelo homem como manuseio incorreto de registros de campo ou perda acidental de arquivos de dados no sistema de computador (Elshorbagy et al, 2000), influenciando diretamente assim na confiabilidade dos resultados.

Para contornar essas situações existem algumas técnicas estatísticas que podem ser aplicadas e, segundo Naghettini e Pinto (2007), as teorias da regressão e correlação são as que ocupam lugar de destaque por serem de uso mais difundido. A regressão linear vem sendo utilizada em diversos estudos como um método adequado para o preenchimento de falhas de dados hidrológicos. Granemann et al. (2018) selecionaram 6 diferentes estações fluviométricas na bacia do rio Iguaçu e aplicaram o método da regressão linear às múltiplas falhas existentes.

Mosavi et al. (2018) afirmam que os métodos estatísticos, apesar de serem amplamente utilizados em situações como previsão de inundações, são inadequados devido à falta de precisão e robustez dos métodos para previsão de curto prazo e, por isso, necessitam de melhorias. Segundo os mesmos autores, a utilização de técnicas de aprendizado de máquina surge como uma alternativa viável aos métodos estatísticos convencionais, uma vez que podem modelar numericamente a não linearidade dos eventos, utilizando para isso dados históricos, sem a obrigatoriedade de conhecimento sobre os processos físicos envolvidos.

O aprendizado de máquina é uma técnica de inteligência artificial que vem sendo utilizada como ferramenta para interpretação e previsão de variáveis hidrológicas a partir da existência de dados anteriormente coletados (Bui et al. 2020; Althoff et al., 2020a). Para Shortridge et al. (2016), o desenvolvimento das técnicas de aprendizado de máquina, combinado com o rápido aumento da capacidade computacional, tem levado a uma extensa aplicação na previsão de dados de vazões nas últimas décadas. Essas técnicas estão baseadas no conceito de aprender a partir de dados específicos, para então identificar padrões relevantes (Saravanan e Nagoor Meeran, 2020). Esses algoritmos são capazes de aprender sistemas não

lineares complexos a partir de um conjunto de observações, geralmente com alto grau de precisão, uma vez que não são afetados por processos subjacentes no sistema (Kingston et al., 2008).

Os métodos de aplicação dos algoritmos de aprendizado de máquina pertencem a três categorias distintas: o aprendizado supervisionado, não supervisionado e de reforço (Mishra, 2019). O aprendizado supervisionado, segundo Guimarães et al. (2019), ocorre, quando o processamento desejado é especificado, utilizando uma ou mais entradas e as saídas desejadas correspondentes. Durante esse processamento é feita uma comparação entre o valor desejado e o valor de saída, gerando um valor de erro, que é utilizado para ajustar os pesos. Já no aprendizado não supervisionado os algoritmos, conforme explicam Pini et al. (2020), são treinados sem que para isso se inclua uma saída desejada, utilizando somente dados de entrada. Na aprendizagem de reforço os algoritmos são treinados de forma que eles aprendam uma ação adequada, pela qual recebem uma recompensa, no ambiente determinado.

Diante da importância do desenvolvimento de estudos que contribuam para o melhor entendimento do comportamento hidrológico de uma região, a obtenção de séries mais contínuas, com minimização de possíveis inconsistências e redução de falhas, promove melhoria na qualidade aos dados, o que implica em maior confiabilidade aos resultados obtidos em quaisquer análises hidrológicas. Assim, os modelos de aprendizado de máquina podem contribuir com o desenvolvimento de estratégias que permitam reduzir as limitações associadas ao processo de medição de vazão em estações fluviométricas convencionais, especialmente para aquelas localizadas a jusante de reservatórios com regularização de vazões.

Nesse contexto, objetivou-se, no presente trabalho, ajustar modelos de aprendizado de máquinas e de regressão linear para fins de previsão de dados diários de vazão em estações fluviométricas convencionais localizadas a jusante de reservatórios de regularização.

1.2. MATERIAL E MÉTODOS

1.2.1. Área de Estudo

O estudo foi desenvolvido na bacia hidrográfica do rio Preto (Figura 1), que possui uma área de drenagem de 10.325 km² e uma variação de elevação de 500 a 1.188 m. O rio Preto nasce no município de Formosa (GO) e faz divisa entre Goiás, Distrito Federal e Minas Gerais. A bacia do rio Preto é uma sub-bacia do rio Paracatu que, por sua vez, pertence à bacia do rio São Francisco.

A região está totalmente inserida no Cerrado e é caracterizada por forte sazonalidade, com temperaturas médias mensais do ar variando de 20,5 a 25,1 °C, e precipitação total anual de aproximadamente 1.300 mm. Devido a fatores favoráveis, como estação chuvosa bem definida e topografia adequada, a bacia tem forte vocação agropecuária, com vários núcleos rurais (ANA, 2019).

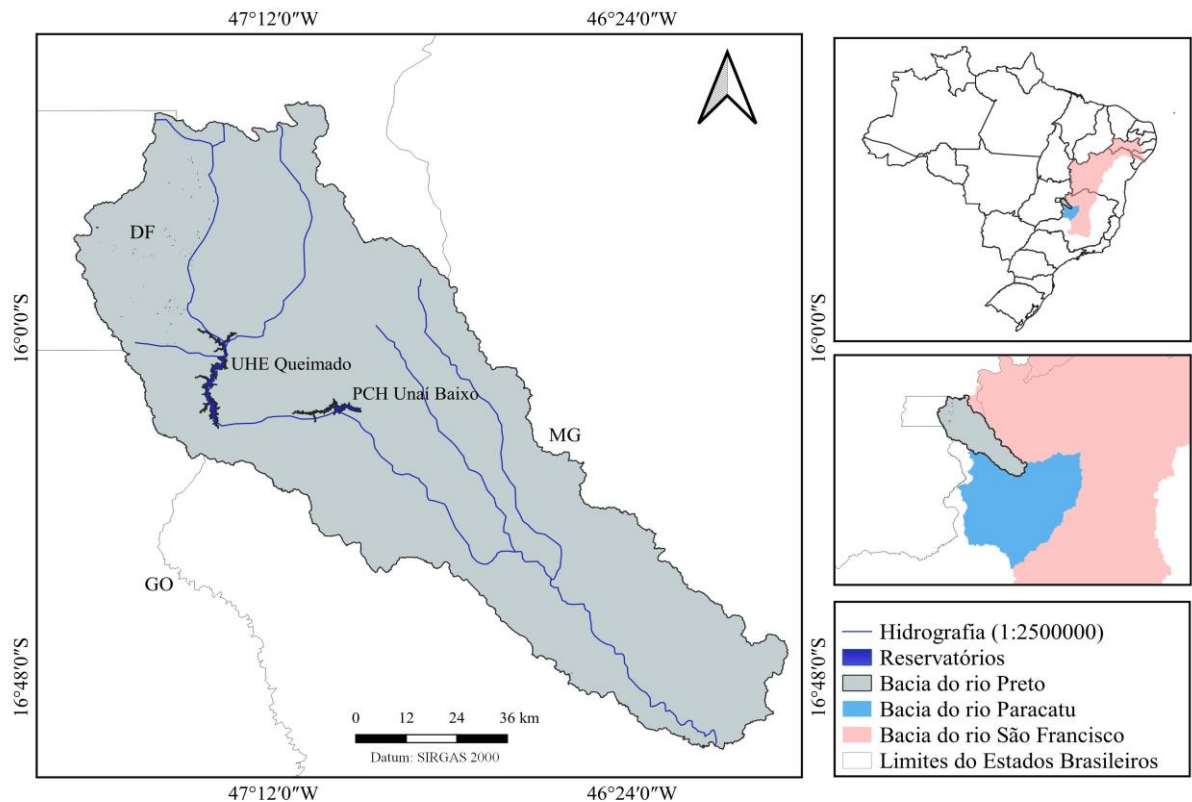


Figura 1 – Localização da bacia hidrográfica do rio Preto.

Segundo a Embrapa (2005), o relevo na região da bacia do rio Preto é formado, principalmente, de chapadas e cristas. A atividade agropecuária é desenvolvida, predominantemente, nas chapadas, sendo as áreas de bordas não utilizadas para agricultura. Nas cristas o uso é restrito por causa do relevo montanhoso.

A região é marcada pela presença da atividade de irrigação, com aumento, segundo levantamento realizado pela ANA, da área irrigada por pivô central, entre 1985 e 2017, de 47.365,3 ha (ANA, 2017). Em função disso, houve um aumento muito expressivo do número de pequenos reservatórios de acumulação de água na bacia (Althoff e Rodrigues, 2019).

Na bacia do rio Preto localiza-se a usina hidrelétrica de Queimado (UHE Queimado), operada pela Companhia Energética de Minas Gerais (CEMIG) e situada no Rio Preto, entre os municípios de Cabeceira Grande (MG), Unai (MG) e Cristalina (GO). A usina possui

reservatório com espelho d'água de 36,26 km², capacidade para acumular 477,98 milhões de m³ e conta com três unidades geradoras, totalizando 105 MW de potência instalada. A construção do reservatório, que proporciona regularização das vazões no rio Preto, iniciou no ano 2000 e sua operação começou em 2004 (ANA, 2019; CEMIG, 2020).

Trata-se de uma usina para aproveitamento hidrelétrico a partir de um reservatório de acumulação de água que exerce papel fundamental na regularização das vazões do rio Preto. As condições para sua operação devem respeitar os usos múltiplos conforme estabelece a Resolução nº 147, de 2 de março de 2015, com descarga mínima de 8,8 m³/s no período úmido (novembro a abril) e 17 m³/s no período de estiagem (maio a outubro). Na bacia do rio Preto também está localizada a Pequena Central Hidrelétrica (PCH) Unai Baixo, entre os municípios de Unai e Cabeceira Grande, com queda d'água de 31 m, cuja potência instalada é de 26 MW. A área inundada é de 10,15 km² e capacidade de armazenamento estimada em 112,19 milhões de m³ (SEMAD, 2017). A regra operacional da PCH Unai é a fio d'água, com previsão de geração ininterrupta, aproveitando ao máximo a vazão regularizada da UHE Queimado, sem deplecionamento de seu reservatório.

1.2.2. Dados hidrológicos

A partir dos inventários disponibilizados pela ANA, foram identificadas 76 estações fluviométricas na bacia do rio Preto. No entanto, apenas 22 dessas estações foram avaliadas, por serem estações consideradas, em função da localização, aquelas que provavelmente apresentam influência das atividades de operacionalização do reservatório da UHE Queimado. Entretanto, ao se verificar a disponibilidade de dados observados para estas estações, apenas 3 delas atenderam aos seguintes critérios: i) estar localizada na bacia do rio Preto; ii) a jusante do reservatório da UHE Queimado; e iii) possuir, no mínimo, 10 anos de dados, considerando o período base de 2005 a 2019, posterior à construção e estabilização das atividades operacionais da UHE Queimado. As estações fluviométricas selecionadas foram: Fazenda Limeira (código 42460000), Unai (Código 42490000) e Porto dos Poções (código 42600000), cujas vazões foram denominadas por Q1, Q2 e Q3, respectivamente (Figura 2 e Tabela 1).

Os dados históricos das estações fluviométricas da rede hidrometeorológica da ANA foram obtidos na plataforma do Hidroweb (<https://www.snirh.gov.br/hidroweb/serieshistoricas>), que inclui dados da RHN (Rede Hidrometeorológica Nacional) de responsabilidade da ANA.

Os registros históricos de vazão associados à operacionalização do reservatório da UHE Queimado foram também obtidos na ANA, através da plataforma SAR (Sistema de Acompanhamento de Reservatórios) que fornece, entre outras informações, as vazões médias diárias defluentes, do período de 2003 a 2019, obtidas das medições horárias realizadas de vazões defluentes (Qd) e de um processo de consistência em primeiro grau aplicado pelo ONS.

A obtenção das vazões defluentes diárias, segundo informações coletadas junto à Companhia Energética de Minas Gerais (CEMIG), são calculadas a partir de informações de rendimento do conjunto turbina-gerador, função de queda bruta ou líquida pela potência ou vazão defluente, e da curva descarga do vertedouro, obtida por modelos de escala reduzida ou modelagem matemática.

As principais informações para essa estação telemétrica, a partir do qual foi obtida as vazões defluentes (Qd) e também para as outras três estações fluviométricas convencionais utilizadas no estudo (Q1, Q2 e Q3), estão apresentadas na Tabela 1.

Foram também identificadas 35 estações pluviométricas (Figura 2) localizadas dentro da bacia do rio Preto e no seu entorno (buffer de 50 km), com disponibilidade mínima de 10 anos de dados dentro do período base de 2005 a 2019.

Os dados históricos de precipitação também foram obtidos na plataforma Hidroweb da ANA (<https://www.snirh.gov.br/hidroweb/serieshistoricas>), sendo selecionadas apenas aquelas estações que satisfizeram aos mesmos critérios estabelecidos para as estações fluviométricas, no que se refere ao período de dados considerado. Foram, portanto, utilizadas informações relativas a 35 estações pluviométricas no total, porém apenas 8 delas localizadas dentro da bacia do rio Preto (Figura 2).

Os dados de precipitação foram utilizados para apoiar às análises hidrológicas realizadas e, para tal, foram consideradas dois tipos de variáveis: precipitação média diária na área de drenagem das estações fluviométricas (PPT_m); e precipitação média diária das séries provenientes das oito estações pluviométricas localizadas dentro da bacia (PPT_i). A diferente seleção dos dados de precipitação foi feita com o objetivo de testar se haveria diferenças consideráveis nos resultados quando se considera um dado pontual e um dado representativo das regiões próximas a estação.

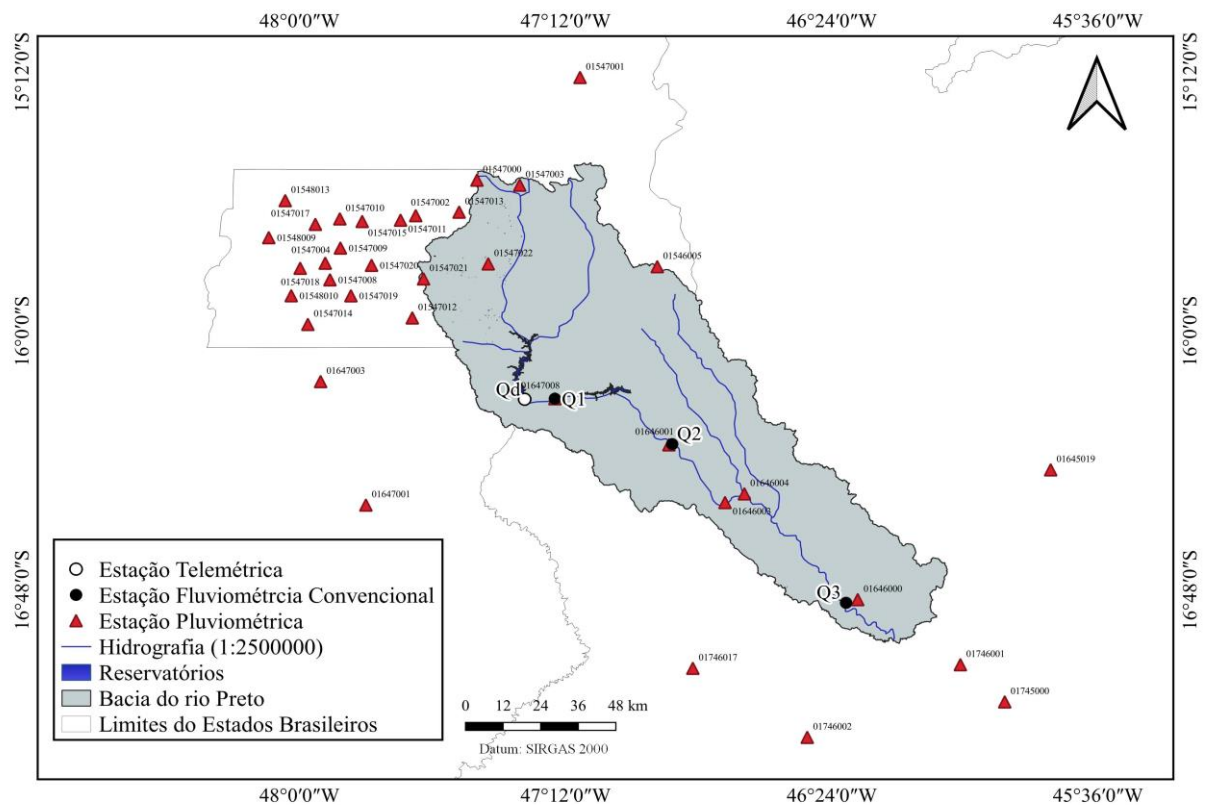


Figura 2 – Localização das estações fluviométricas e pluviométricas da rede hidrometeorológica da ANA e da estação telemétrica de vazões defluentes do reservatório UHE Queimado selecionadas para o estudo na bacia do rio Preto.

Tabela 1 – Características das estações fluviométricas convencionais e da estação telemétrica localizada na UHE Queimado

Código	Nome	Variáveis	Latitude	Longitude	Área (km ²)	Tipo
42460000	Fazenda Limeira	Q1	-16,2089	-47,2325	3.890	Convencional
42490000	Unaí	Q2	-16,3494	-46,88	5.360	Convencional
42600000	Porto dos Poções	Q3	-16,8397	-46,3572	9.400	Convencional
42459080	UHE Queimado Barramento	Qd	-16,2106	-47,3233	3.657	Telemétrica

A PPT_m foi estimada a partir de técnicas de interpolação de dados provenientes das 35 estações selecionadas e para isso foi utilizado o método de interpolação IDW (Inverse Distance Weighted).

O IDW consiste em uma média ponderada pelo inverso de uma potência da distância dos valores dos dados dentro de uma vizinhança usada para estimar o valor da variável de interesse em local não amostrado (Mello e Silva, 2013). Segundo Lu e Wong (2008) é um método considerado relativamente rápido e fácil de calcular, de simples interpretação. No estudo realizado por Xavier et al. (2015) foram avaliados seis diferentes métodos de

interpolação e o IDW foi identificado como o de melhor desempenho ao se utilizar diversas variáveis meteorológicas, incluindo a precipitação.

No processo de interpolação, realizado no software R, foram testados diferentes valores de potência da distância, sendo escolhido aquele que apresentou menor raiz do quadrado médio do erro (RMSE) para cada uma das interpolações diárias realizadas, sendo para isso encontrado ajustes que variaram entre 1, 2, 3, 4, 5 e 6, conforme as características dos diferentes conjuntos de dados.

1.2.3. Cenários e critérios para a predição de vazões diárias

Após a obtenção das séries históricas de vazão das estações Q1, Q2 e Q3, foi dado início aos procedimentos de modelagem para a predição de vazões diárias nas respectivas estações fluviométricas convencionais localizadas a jusante do reservatório da UHE Queimado, utilizando para tal dados das vazões defluentes (Q_d), observadas na saída do reservatório.

Além das vazões defluentes coletadas na estação telemétrica localizada na UHE Queimado, foram utilizadas na modelagem das vazões das estações Q1, Q2 e Q3 os valores das precipitações diárias (PPT_i) das estações pluviométricas consideradas, bem como os dados das precipitações médias diárias na área de drenagem (PPT_m).

No caso da estação Q3 foi considerado, ainda, para o modelo de predição dos valores diários de vazão, além das variáveis já mencionadas, a inserção das séries de dados observados nas estações Q1 e Q2 como apoio para os processos de modelagens aplicados. Essa consideração se deu em função do incremento da área de drenagem que a estação Q3 apresenta em relação a Q_d , o que pode significar uma maior influência associada às variáveis climáticas ou antrópicas, além da própria operação do reservatório no que se refere ao comportamento do escoamento nesse trecho.

Foram associadas às séries de vazões de Q_d , Q1 e Q2, quando utilizadas como variáveis explicativas na modelagem das vazões, períodos de defasagem temporal no escoamento que variaram de 1 a 4 dias, levando em consideração a distância percorrida pela água e, por consequência, o tempo de deslocamento da vazão entre as seções consideradas. Sendo assim, a título de exemplo, além da variável Q_d , foram consideradas ainda as variáveis Q_d_1 (1 dia de defasagem), Q_d_2 (2 dias de defasagem), Q_d_3 (3 dias de defasagem) e Q_d_4 (4 dias de defasagem), com análise semelhante para Q1 e Q2.

Com base nessas informações, as variáveis explicativas consideradas nos modelos para predição de vazões diárias nas séries históricas para cada uma das estações foram organizadas

em dois grupos (G_{ppti} e G_{pptm}), que se distinguem apenas em relação ao tipo de dado associado aos valores de precipitação que foram utilizados como variáveis explicativas. A Tabela 2 apresenta as variáveis adotadas em cada um dos Grupos, para cada uma das estações estudadas.

Tabelas 2 – Variáveis explicativas consideradas nos modelos para predição de vazões diárias nas diferentes estações estudadas e grupos associados aos valores de precipitação

Estações em análise	Variáveis Explicativas (G_{ppti})	Variáveis Explicativas (G_{pptm})	Dados Utilizados
Q1	Qd, Qd_1, Qd_2, Qd_3, Qd_4, PPT ₁ , PPT ₂ , PPT ₃ , PPT ₄ , PPT ₅ , PPT ₆ , PPT ₇ , PPP ₈	Qd, Qd_1, Qd_2, Qd_3, Qd_4, PPT _{m_Qd} , PPT _{m_Q1}	2005 a 2014
Q2	Qd, Qd_1, Qd_2, Qd_3, Qd_4, PPT ₁ , PPT ₂ , PPT ₃ , PPT ₄ , PPT ₅ , PPT ₆ , PPT ₇ , PPP ₈	Qd, Qd_1, Qd_2, Qd_3, Qd_4, PPT _{m_Qd} , PPT _{m_Q2}	2005 a 2019
Q3	Qd, Qd_1, Qd_2, Qd_3, Qd_4, PPT ₂ , PPT ₃ , PPT ₄ , PPT ₅ , PPT ₆ , PPT ₇ , PPP ₈	Qd, Qd_1, Qd_2, Qd_3, Qd_4, PPT _{m_Qd} , PPT _{m_Q4}	2005 a 2019
Q3	Qd, Qd_1, Qd_2, Qd_3, Qd_4, Q1, Q1_1, Q1_2, Q1_3, Q1_4, Q2_1, Q2_2, Q2_3, Q2_4, PPT ₂ , PPT ₃ , PPT ₄ , PPT ₅ , PPT ₆ , PPT ₇ , PPP ₈	Qd, Qd_1, Qd_2, Qd_3, Qd_4, Q1, Q1_1, Q1_2, Q1_3, Q1_4, Q2_1, Q2_2, Q2_3, Q2_4, PPT _{m_Qd} , PPT _{m_Q4}	2005 a 2014

Q₁, Q₂ e Q₃: vazão das estações 1, 2 e 3, respectivamente; Q_{1.1}, Q_{1.2}, Q_{1.3} e Q_{1.4}: vazão da estação 1 com 1, 2, 3 e 4 dias de defasagens, respectivamente, Q_{2.1}, Q_{2.2}, Q_{2.3} e Q_{2.4}: vazão da estação 2 com 1, 2, 3 e 4 dias de defasagem, respectivamente; Q_d: vazão defluente ao reservatório; Q_{d.1}, Q_{d.2}, Q_{d.3} e Q_{d.4}: vazão defluente ao reservatório com 1, 2, 3 e 4 dias de defasagem, respectivamente; PPT_m: precipitação média da área de drenagem; e PPT₁, PPT₂, PPT₃, PPT₄, PPT₅, PPT₆, PPT₇ e PPT₈: precipitações das estações 1 (01547003), 2 (01547021), 3 (01547022), 4 (01646004), 5 (01647008), 6 (01646001), 7 (01646003) e 8 (01646000), respectivamente.

Embora os dados disponíveis de Qd correspondam ao período de 2003 a 2019 e não apresentam falhas em seus registros, decidiu-se em utilizar no presente estudo o período de 2005 a 2019. Essa decisão deu-se em função de observações de vazões atípicas nos anos de 2003 e 2004, registrando em alguns momentos valores iguais a zero (0), ocasionados, muito provavelmente, pelas atividades pontuais decorrentes da construção do reservatório.

O período de dados utilizados para a estação Q2 também não apresentou falhas, sendo adotado o período de 2005 a 2019. Para a estação Q1, os registros de vazões utilizados foram de 2005 a 2014, embora seja possível evidenciar registros nos anos de 2015 e 2016. Os valores encontrados para os dois últimos anos apresentaram magnitudes extremamente elevadas comparativamente à maioria dos valores apresentados na série, não sendo evidenciado nenhum evento hidrológico ou de qualquer outra natureza que justificasse tal comportamento. Sendo assim, os anos de 2015 e 2016 foram removidos da modelagem das vazões na estação Q1, assim

como para a modelagem das vazões na estação Q3 em que foi utilizado dados de Q1 dentre as variáveis explicativas analisadas.

Já para a estação Q3, foram utilizados dois conjuntos de dados distintos, isto é, foram testadas dois conjuntos de dados diferentes para a realização das modelagens. Uma considerando apenas dados de vazão defluente ao reservatório e outra utilizando dados também das estações fluviométricas convencionais. Dessa forma, a disponibilidade de dados para cada conjunto de variáveis adotadas é o que justifica a diferença de período entre as modelagens. O uso apenas de Qd como variável explicativa associada a vazão possibilitou a adoção do período de análise de 2005 a 2019, entretanto ao inserir Q1 na nova modelagem, cujos dados apresentam-se disponíveis apenas entre os anos de 2005 a 2014, é necessário a consideração de um novo período de dados (2005 a 2014).

Os grupos, para cada uma das estações, foram ainda organizados em seis cenários, construídos a partir de diferentes critérios, adotados para a seleção dos dados utilizados nos modelos de predição, estando apresentados na Tabela 3.

Tabela 3 – Cenários e critérios adotados para a seleção do conjunto de dados a serem utilizados na modelagem do comportamento médio das vazões nas estações fluviométricas

Grupos	Cenários	Critérios
G _{PPTi}	C1	Utilização de todos os dados disponíveis, exceto dias com falhas identificados em alguma das covariáveis utilizadas
	C2	Remoção de dados que apresentaram relação $Q1/Qd \leq 0,90$, $Q2/Qd_{-1} \leq 0,90$ e $Q3/Qd_{-1} \leq 1,0$
	C3	Remoção de dados que apresentaram relação $Q1/Qd \leq 0,90$ e $Q2/Qd_{-1} \leq 0,90$, e daqueles identificados como outliers pelo boxplot
G _{PPTm}	C4	Utilização de todos os dados disponíveis, exceto dias com falhas identificados em alguma das covariáveis utilizadas
	C5	Remoção de dados que apresentaram relação $Q1/Qd \leq 0,90$, $Q2/Qd_{-1} \leq 0,90$ e $Q3/Qd_{-1} \leq 1,0$
	C6	Remoção de dados que apresentaram relação $Q1/Qd \leq 0,90$, $Q2/Qd_{-1} \leq 0,90$ e daqueles identificados como outliers pelo boxplot

Para C1 e C4 foram considerados todos os dados disponíveis nas séries históricas selecionadas, sendo excluídos do processo de modelagem apenas os dados dos dias que apresentaram falhas em qualquer uma das variáveis explicativas.

Para o estabelecimento dos cenários C2 e C5 foi realizada, previamente, uma análise de correlação entre a série de vazões de cada uma das estações selecionadas e as séries de vazões

defluentes defasadas ou não no tempo (Q_d , Q_{d_1} , Q_{d_2} , Q_{d_3} e Q_{d_4}), a fim de identificar o período de defasagem temporal que melhor se ajustaria ao comportamento das vazões nas seções escolhidas. Deste modo, foi verificado maior correlação de Q_1 com Q_d , Q_2 com Q_{d_1} e Q_3 com Q_{d_1} , significando que a vazão defluente do reservatório UHE Queimado é detectada nas estações Q_2 e Q_3 um dia depois, em função da distância e da declividade existentes entre os pontos. Em contrapartida, para Q_1 a resposta é imediata, ou seja, no mesmo dia em que a vazão defluente é liberada do reservatório, ela é registrada na estação Q_1 .

Posteriormente foram adotadas relações entre as estações analisadas e a série de vazões defluentes defasadas que mais se correlacionaram entre si. Estações mais a jusante tendem a apresentar valores de vazões diárias superiores aos das estações de montante, ou seja, a razão entre elas deve apresentar valores maiores que 1 (um) e, por esse motivo, o estabelecimento dos cenários C2 e C5 levou em consideração a remoção de dados que apresentaram uma relação $Q_1/Q_d \leq 0,90$ e $Q_2/Q_{d_1} \leq 0,90$. Sendo assim, 10% foi o limite de permissividade incorporado em função do nível de precisão dos equipamentos utilizados na medição de vazão e/ou possíveis captações que ocorrem ao longo do trecho do curso d'água entre as duas seções consideradas. Já a relação estabelecida para remoção dos dados em Q_3 foi $Q_3/Q_{d_1} \leq 1,0$, em função da área de drenagem incremental que neste caso ultrapassa em 60% o valor da área correspondente à medição da Q_d na UHE Queimado. Isso significa que a área de Q_3 supera em quase 2,5 vezes a de Q_d , tornando, dessa forma, improvável que Q_3 registre um valor de vazão inferior a Q_{d_1} .

Para C3 e C6, além dos critérios já estabelecidos para C2 e C5, foram removidos ainda os dados cujas relações Q_1/Q_d , Q_2/Q_{d_1} e Q_3/Q_{d_1} comportaram-se como outliers a partir da construção do boxplot mensal para cada uma das relações.

É importante destacar que para a modelagem de Q_3 os critérios estabelecidos para C3 e C6 não foram aplicados devido ao aumento expressivo da área de drenagem em relação a Q_d , sendo identificados maiores variabilidades quanto ao tipo de solo (Santos et al, 2020), declividade, geomorfologia (Silva, 2012) e, ainda, quanto à demanda de água para irrigação, conforme aponta os dados publicados no Atlas de Irrigação (ANA, 2017), o que dificulta e torna subjetiva e arriscada a determinação de valores de vazões classificados como outliers. Por se tratar de estações que apresentam grande distância entre si, a remoção de outliers em séries temporais, conforme destacam Maillard e Pereira (2017), pode ser complexa devido à natural ocorrência de valores extremos (secas e cheias) que são muito importantes de serem preservados. Em função de todas essas avaliações, nesse estudo não foi considerado prudente a remoção dos valores acima do limite máximo estabelecido pelo boxplot para a estação Q_3 .

1.2.4. Modelos de regressão linear e aprendizado de máquina utilizados

A modelagem do comportamento das vazões médias diárias, a partir das variáveis consideradas na Tabela 2, foi feita com base nos algoritmos de regressão linear simples (RLS), regressão linear múltipla (RLM) e três algoritmos de aprendizado de máquina, denominados Random Forest (RF), Cubist (CB) e Rede Neural com Regularização Bayesiana (RBNN).

A RLS é usada para estabelecer uma relação funcional entre uma variável dependente e uma explicativa (Naghetini e Pinto, 2007), sendo a equação para essa relação expressa pela Equação 1.

$$y = \beta_0 + \beta_1 x \quad \text{Equação 1}$$

Em que, y é a variável a ser explicada ou dependente, β_0 é o parâmetro a ser estimado, que representa o intercepto da reta, local onde a reta corta o eixo y , β_1 é o parâmetro a ser estimado que representa a inclinação da reta, medindo o ângulo que ela forma com o eixo x e x é a variável explicativa ou independente.

Na RLM é estabelecido a relação funcional linear da variável dependente com duas ou mais variáveis independentes, sendo o modelo geral expresso pela Equação 2.

$$y = \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_p x_p \quad \text{Equação 2}$$

Em que y é a variável a ser explicada ou dependente, $\beta_1, \beta_2, \dots, \beta_p$ são os coeficientes de regressão. x_1, x_2, \dots, x_p são as variáveis explicativas ou independentes.

Os modelos de regressão são reconhecidamente utilizados para modelagem hidrológica, incluindo a previsão de vazão com base em uma ou mais observações de vazão em estações de apoio. Aplicações como essa podem ser observadas em trabalhos como o de Kamwaga et al (2018), que buscaram avaliar diferentes métodos de preenchimento de falhas de vazão na bacia de Little Ruaha. Entre os resultados encontrados por esses autores, a regressão linear múltipla aparece como um dos modelos que alcançaram melhor desempenho estatístico.

O modelo Random Forest (RF) é um classificador que combina árvores de regressão e classificação e cada uma possui o valor de um vetor aleatório amostrado independentemente, com a mesma distribuição para todas as árvores na floresta, de modo que a classificação final é produzida a partir da classe com maior número de votos dentre todas as árvores existentes no

modelo de previsão (Breiman, 2001). Conforme explica Souza (2017), procurou-se criar e treinar cada árvore a partir da amostragem de dados originais, selecionando aleatoriamente um subconjunto das variáveis de entrada para que assim fosse realizada a divisão para cada nó.

Para determinar o número ideal de variáveis selecionadas aleatoriamente para construção de cada árvore do RF, foram construídos modelos utilizando diferentes números de variáveis e selecionando aquele que apresentou menor raiz do erro quadrático médio (RMSE) para a validação cruzada no conjunto de treinamento (Willmott e Matssura, 2015; Althoff, 2020b). O número mínimo de variáveis de entrada avaliadas para divisão de dados na geração de cada nó (`mtry`) foi considerado como quatro. O pacote Random Forest (Liaw e Wiener 2002) em R foi usado para ajustar os modelos de RF.

O modelo Cubist (CB) também é baseado na construção árvores de regressão sucessivas. Diferente das árvores de regressão tradicionais, as várias regras geradas estão cada uma associada a um modelo linear diferente. O algoritmo não recupera um modelo final como o RF, mas um conjunto de regras associado a um conjunto de modelos multivariados (Noi et al, 2017). Dessa forma, a predição final foi feita por meio da combinação de diversos modelos lineares, e não por valores discretos nos nós terminais de cada árvore.

O conjunto de dados de uma árvore foi inicialmente dividido em subconjuntos homogêneos, onde o critério de divisão foi baseado no desvio padrão como uma medida de erro para observações definidas em um determinado nó (Pal e Deswal, 2009). Posteriormente, para superar o problema de o processo produzir árvores demasiadamente grandes, a árvore foi podada e as subárvores podadas foram substituídas por modelos de regressão linear (Rahimikhoob, 2014). Assim, sempre que um preditor cumpriu um conjunto de regras, a previsão foi realizada com base no modelo linear associado a ele, conforme descrito em Althoff et al. (2020b).

A determinação do número de árvores subsequentes favorável foi feita construindo modelos utilizando números diferentes de árvores e selecionando o modelo que apresentasse menor RMSE. A análise de regressão CB foi realizada com o pacote Cubist (Kuhn et al, 2018) no software R.

Já o modelo baseado em redes neurais (BRNN) é um método que representa relações não lineares, apresentando habilidades em aprender e generalizar as características importantes em séries hidrológicas, como sazonalidade e periodicidade. O diferencial do BRNN é a penalização por excesso de crescimento e redes complexas, em que pesos de conexões de neurônios desnecessários são considerados como zero (Ticknor, 2013).

O pacote BRNN (Rodriguez e Gianola, 2018) em R foi usado para ajustar os modelos BRNN, em que o número de neurônios na camada foi otimizado com base na função objetivo SDR (Equação 3), sendo selecionado aquele que apresentasse menor RMSE.

$$F = \beta E_D + \alpha E_w \quad \text{Equação 3}$$

Em que F é a função objetiva minimizada; E_D é o erro da soma de quadrados; E_w é a soma de quadrados dos pesos e vieses e α e β são parâmetros da função objetivo.

1.2.5. Validação dos modelos

Os dados referentes a cada conjunto de variáveis foram divididos em dois subconjuntos. O primeiro subconjunto, composto por 80% dos dados (anos iniciais das séries), foi utilizado para treinamento dos modelos, e o segundo, com 20% (anos finais das séries), foi considerado no processo de teste dos modelos gerados (Mohanty et al, 2016).

Visando obter valores de hiper parâmetros livres do efeito de “over-training” para cada modelo, foi avaliado o desempenho do conjunto de treinamento através da validação cruzada k-folds (Vehtari et al., 2017), com repetições ($k=10$, $n=10$).

Portanto, a partir do desempenho das predições realizadas pelos modelos para o conjunto de teste, ou seja, com 20% dos dados, foi realizada o teste para as predições de vazões, usando um conjunto de dados não utilizado no processo de treinamento, fornecendo assim, novos critérios de desempenho.

Cabe ressaltar que a modelagem, treinamento, validação e teste dos modelos foi realizada usando recursos do pacote caret (Kuhn, 2020).

1.2.6. Desempenho e análise estatística dos modelos

A fim de avaliar o desempenho dos modelos, foram consideradas as seguintes métricas estatísticas: índice de eficiência de Nash-Sutcliffe (NSE) (Equação 4), coeficiente de determinação (R^2) (Equação 5), raiz do erro quadrático médio (RMSE) (Equação 6), erro médio absoluto (MAE) (Equação 7), índice de eficiência de Kling-Gupta (KGE) (Equação 8) e viés percentual (PBIAS) (Equação 9), conforme descrito em Yaseen et al. (2020).

$$NSE = 1 - \left(\frac{\sum((Q_{obsi} - Q_{esti})^2)}{\sum((Q_{obsi} - \bar{Q}_{obs})^2)} \right) \quad \text{Equação 4}$$

$$R^2 = \frac{(\sum_{i=1}^n (Q_{est} - \bar{Q}_{est})(Q_{obs} - \bar{Q}_{est}))^2}{\sum_{i=1}^n (Q_{est} - \bar{Q}_{est})^2 \sum_{i=1}^n (Q_{obs} - \bar{Q}_{est})^2} \quad \text{Equação 5}$$

$$RMSE = \sqrt{\frac{1}{n} \sum (Q_{esti} - Q_{obsi})^2} \quad \text{Equação 6}$$

$$MAE = \frac{1}{n} \sum |Q_{esti} - Q_{obsi}| \quad \text{Equação 7}$$

$$KGE = 1 - \sqrt{(r - 1)^2 + \left(\frac{\sigma_{est}}{\sigma_{obs}} - 1\right)^2 + \left(\frac{\mu_{est}}{\mu_{obs}} - 1\right)^2} \quad \text{Equação 8}$$

$$PBIAS (\%) = \frac{\sum(Q_{obsi} - Q_{esti})}{\sum Q_{obsi}} \times 100 \quad \text{Equação 9}$$

Em que n é o número de observações, Q_{obs} é a média das vazões observadas, Q_{obsi} e Q_{esti} são as vazões observada e estimada no dia i , respectivamente, r é a correlação linear entre a vazão observada e estimada, σ_{est} e σ_{obs} representam o desvio padrão dos dados observados e estimados, respectivamente, e μ_{est} e μ_{obs} correspondem à média dos dados observados e estimados, respectivamente.

1.2.7. Importância das covariáveis

Foi avaliada, ainda, a importância de cada umas das covariáveis utilizadas para os diferentes modelos, buscando-se identificar as de maior influência para o desempenho das diferentes modelagens das predições de vazões. Para isso, foi atribuído, à covariável mais importante 100% e à menos importante 0% (Reis et al., 2021)

O valor de importância das variáveis para os modelos lineares RLS e RLM foi extraído a partir dos valores absolutos encontrados na estatística t para cada parâmetro.

Para o modelo BRNN foi utilizado o método dos pesos, proposto por Garson (1991), em que cada variável de entrada conecta a um neurônio por um peso, de tal forma que, quanto maior o peso, maior será a importância.

O modelo CB fornece a porcentagem de vezes em que cada variável foi usada para uma condição. Em cada divisão da árvore, o algoritmo salva um modelo linear que pode conter termos para cada variável usada na divisão atual ou qualquer divisão acima dela (Kuhn, et al,

2012), resultando em uma função de todos os modelos lineares do nó inicial ao nó terminal. As porcentagens mostradas na saída cubista refletem todos os modelos envolvidos na previsão (em oposição aos modelos de terminal mostrados na saída). Portanto, a importância é baseada na porcentagem de vezes que essas variáveis aparecem em modelos lineares recursivamente ao longo das árvores.

Para o RF é avaliado o out-of-bag, que é um conjunto de dados criado no processo de treinamento no momento de construção das árvores. Uma parte constrói a árvore e a outra parte é o out-of-bag daquela árvore. Segundo explica Cutler et al (2007), para avaliar a importância de uma variável preditora específica, os valores da variável são permutados aleatoriamente para as observações out-of-bag e, em seguida, os dados out-of-bag modificados são passados para baixo nas árvores para obter novas previsões. A diferença entre a taxa de classificação incorreta para os dados out-of-bag originais e modificados, dividido pelo erro padrão, é uma medida de importância da variável.

1.3. RESULTADOS E DISCUSSÃO

1.3.1. Dados obtidos para aplicação dos modelos

Tomando como base os critérios adotados para cada um dos cenários considerados e das séries históricas disponíveis, o número de dados obtidos e utilizados na modelagem das vazões diárias foi diferente em cada uma das estações fluviométricas convencionais selecionadas, sendo apresentados na Tabela 4.

Tabela 4 – Quantidade de dados diários utilizados, para cada um dos cenários, considerando as três variáveis de interesse (Q1, Q2 e Q3)

Dados Utilizados	Variável Resposta	G _{PPTi}			Variável Resposta	G _{PPTm}		
		Cenários	Quantidade de dados	*Dados Descartados (%)		Cenários	Quantidade de dados	*Dados Descartados (%)
2005 a 2014	Q1	C1	3515	-	Q1	C4	3648	-
	Q1	C2	3328	5,3%	Q1	C5	3462	5,1%
	Q1	C3	3127	11,0%	Q1	C6	3234	11,3%
2005 a 2019	Q2	C1	5187	-	Q2	C4	5474	-
	Q2	C2	4710	9,2%	Q2	C5	4992	8,8%
	Q2	C3	4411	15%	Q2	C6	4651	15%
2005 a 2014**	Q3	C1	3516	-	Q3	C4	3648	-
	Q3	C2	3426	2,6%	Q3	C5	3558	2,5%
2005 a 2019	Q3	C1	5122	-	Q3	C4	5409	-
	Q3	C2	4902	4,3%	Q3	C5	5188	4,1%

*Contagem dos dados descartados considerando o total presente em C1 e C4, para cada variável e período selecionados

**Modelagem adicional realizada para Q3 considerando a inserção das variáveis associadas às estações fluviométricas de apoio Q1 e Q2, que em função do período de dados disponíveis para Q1, apresentou redução na quantidade de informações a serem submetidas aos testes.

Considerando as diferenças entre os períodos-base adotados para cada uma das variáveis de interesse (Q1, Q2 e Q3), foram identificadas variações no quantitativo de dados utilizados para cenários submetidos a um mesmo critério de seleção. As diferenças de informações entre C1 e C4, C2 e C5 e C3 e C6 estão associadas à substituição das séries históricas de precipitação das estações pluviométricas inseridas dentro da bacia pelas precipitações médias diárias nas áreas de drenagens das estações fluviométricas.

Ao se substituir os dados das séries de precipitação diária observados nas estações pluviométricas localizadas dentro da bacia do rio Preto por um único valor de precipitação diária obtido para cada área de drenagem a partir do processo de interpolação, é possível a estimativa de um valor médio de precipitação, mesmo existindo falhas em alguma das estações utilizadas. Isso ocorre porque, na presença de falha em alguma estação, a interpolação é feita considerando apenas os dados das estações sem falhas, permitindo, ainda assim, a obtenção de um valor de precipitação média para cada dia.

Na Figura 3 são apresentados, para as três estações fluviométricas em estudo, apenas os dados de vazões considerados incoerentes para o C2. Isso significa que essas vazões, cujas relações $Q1/Qd$ e $Q2/Qd_1$ foram menores ou iguais a 0,90 e $Q3/Qd_1$ menor ou igual a 1,0, puderam ser identificadas e eliminadas do conjunto de dados submetidos à modelagem das vazões para esse cenário.

Observa-se ainda, na Figura 3, que os dados de vazões incoerentes foram evidenciados tanto nos meses de estiagem da região (junho, julho, agosto e setembro) como nos meses chuvosos (fevereiro, março e abril), reforçando a baixa confiabilidade dessas estimativas, especialmente quando se analisa Q1, cujo incremento da área de drenagem em relação a Qd não ultrapassa 6%.

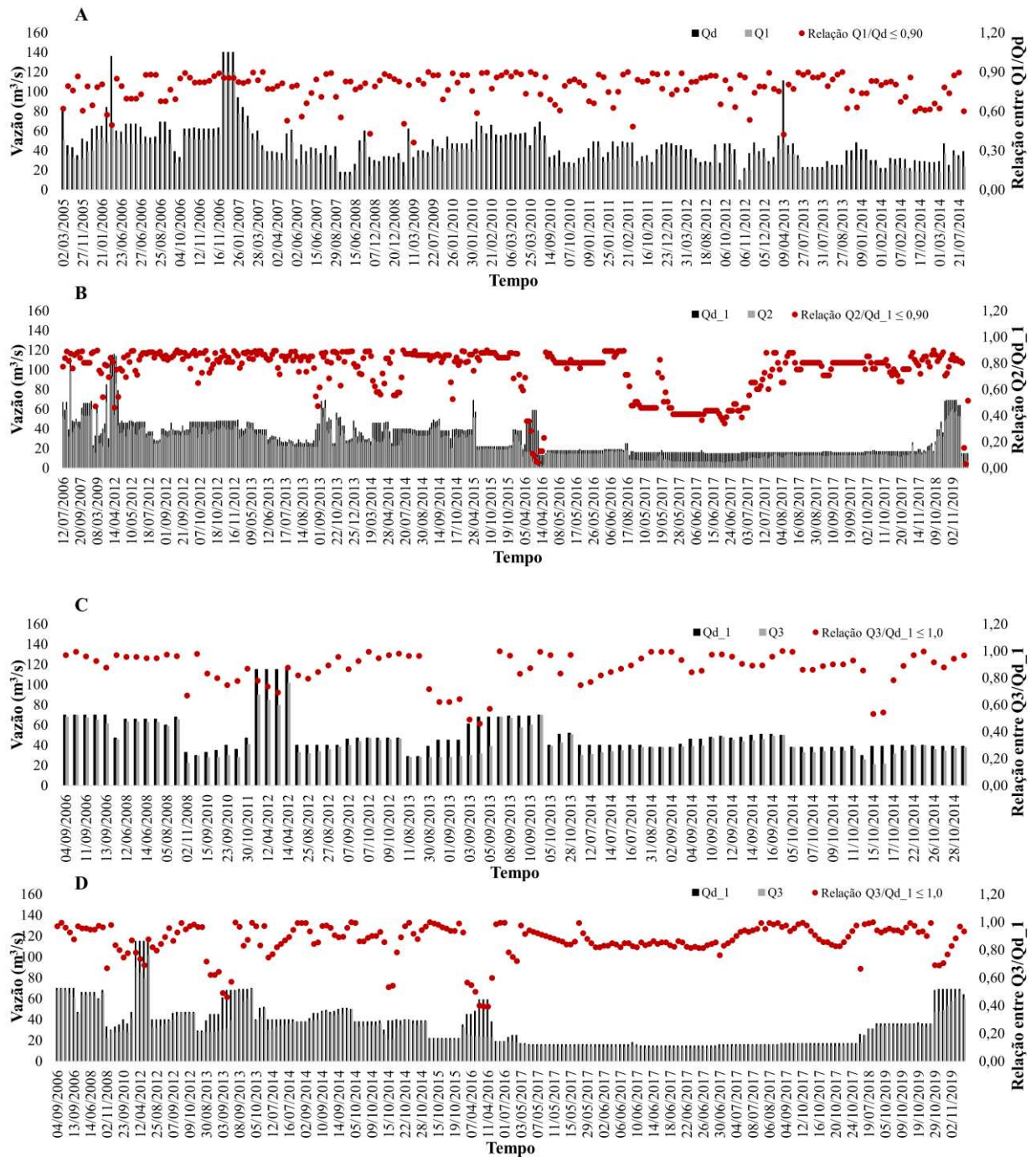


Figura 3 – Dados de vazão que apresentaram relações entre Q1 e Qd $\leq 0,90$ – período de 2005 a 2014 (A), Q2 e Qd_1 $\leq 0,90$ – período de 2005 a 2019 (B), Q3 e Qd_1 $\leq 1,0$ – período de 2005 a 2014 (C) e Q3 e Qd_1 $\leq 1,0$ – período de 2005 a 2019 (D), para o Cenário 2, dentro do GPPTi.

Análise semelhante foi realizada para o cenário C5, apresentando diferenças apenas em função da quantidade de dados removidos (Tabela 4).

Com base nos valores observados nas séries históricas das estações estudadas, foi verificada a existência de vazões a jusante do reservatório menores que o próprio valor de Qd, o que levou à consideração dos Cenários 2 e 5. Mesmo assim, foram ainda observados valores de vazão nas estações em proporções muito superiores quando comparadas com a vazão defluente do reservatório UHE Queimado, para determinados dias. Esses valores podem ter sido registrados de maneira inapropriada em função de erros de observação/equipamentos ou, ainda, associados à variação horária de operação do reservatório de regularização, localizado a montante das estações, uma vez que as estações fluviométricas convencionais realizam apenas duas leituras de cota do nível de água ao longo de todo o dia.

Visando minimizar esses efeitos é que foi aplicado o boxplot para se avaliar as relações entre as vazões diárias $Q1/Qd$ e $Q2/Qd_1$. Os valores considerados outliers neste estudo, isto é, acima do limite superior da caixa de boxplot foram identificados e removidos para então se estabelecer os cenários C3 e C6.

A distribuição das estações pluviométricas com disponibilidade de dados, dentro da bacia do rio Preto, conforme se observa na Figura 2, é consideravelmente escassa e heterogênea, estando abaixo das recomendações feitas pela OMM (1994) acerca da densidade de estações por km^2 para regiões caracterizadas como planas, relevo que mais prevalece na região, ou seja, 18 unidades ($575 km^2/estação$). Entretanto, na área da bacia foram obtidas apenas oito estações. Além disso, as estações estão situadas predominantemente em posições próximas às estações fluviométricas estudadas e nas regiões limites da área da bacia (montante), como se verifica na Figura 2, o que demonstra a baixa representatividade dos dados quanto à distribuição espacial das chuvas, especialmente para as estações representativas de Q1 e Q2.

Ainda assim, antes da remoção das vazões identificadas como outliers pelo boxplot a partir das relações entre $Q1/Qd$ e $Q2/Qd_1$, foram realizadas análises quanto ao comportamento das vazões em relação aos registros de chuva disponíveis nas séries pluviométricas. Entretanto, não se conseguiu estabelecer uma relação de causa e efeito entre ambas, uma vez que há grande variabilidade de comportamento entre as variáveis. Valores de vazões maiores estão associados tanto a precipitações de grande quanto de baixa ou nenhuma magnitude e vice-versa, seja considerando chuvas de 24 ou 48 horas. Diante dessa inconstância, optou-se por excluir todos os valores considerados outliers pelo boxplot, estabelecendo-se assim o conjunto de dados para a modelagem das vazões nos cenários C3 e C6.

Após o preparo e eliminação dos dados para cada um dos cenários estabelecidos, foram utilizados modelos de regressão linear e algoritmos de aprendizado de máquina visando modelar o comportamento médio diário das vazões em cada uma das estações fluviométricas analisadas.

1.3.2. Desempenho dos modelos de aprendizado de máquinas e de regressão linear para Q1

São apresentadas na Figura 4 os desempenhos dos modelos de aprendizado de máquinas e de regressão linear simples e múltipla obtidas na modelagem das vazões diárias para a estação fluviométrica Q1, incluindo os índices estatísticos tanto no treinamento quanto no teste dos modelos, considerando os cenários C1 e C4. Os valores apresentados para a RLS se referem ao modelo cuja variável explicativa de melhor desempenho foi atribuída a Qd.

Analisando o desempenho dos modelos nos cenários C1 e C4, verifica-se que os resultados com melhores desempenhos foram obtidos quando se utilizou a precipitação média da área de drenagem (PPT_m), entre as variáveis explicativas dos modelos, indicando assim, superioridade, ainda que pequena (maiores valores de NSE, KGE, R^2 e menores valores de RMSE e MAE), em relação à utilização da série histórica de precipitação das estações pluviométricas localizadas dentro da bacia (PPT_i).

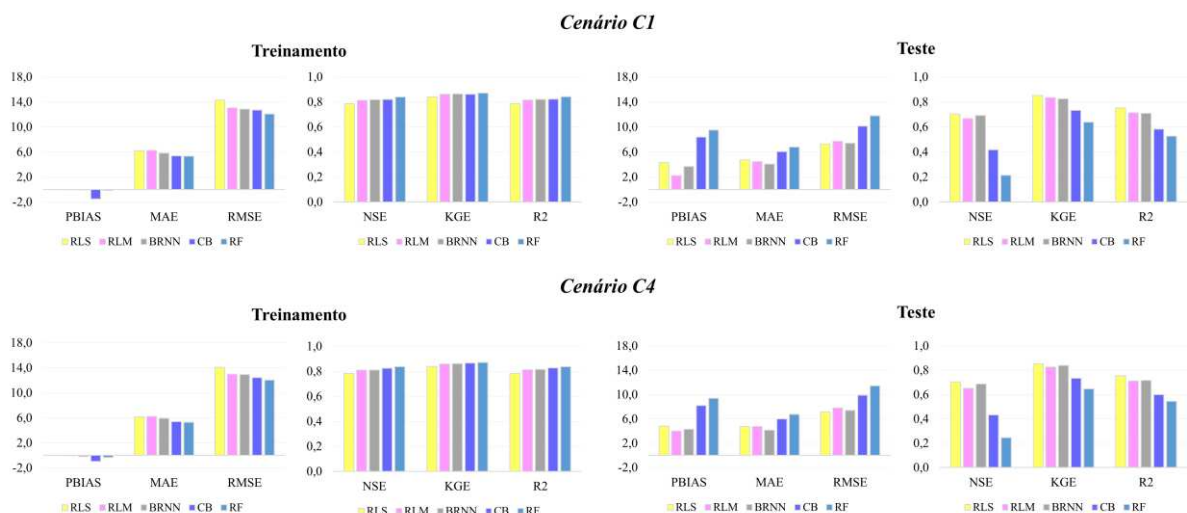


Figura 4 - Índices estatísticos obtidos no treinamento e no teste dos modelos de aprendizado de máquina (BRNN, CB e RF) e métodos de regressão linear (RLS e RLM) para estimativa da variável Q1, considerando os cenários C1 e C4

Para o treinamento, embora o valor do PBIAS para a maior parte dos modelos tenha apresentado tendência de superestimativa, os valores ficaram abaixo de 15% para todos os modelos, caracterizando assim, como desempenho aceitável, segundo a classificação proposta em Ma et al. (2012), sendo o melhor valor, obtido para a RLS (0,0%). Para o RMSE, os valores ficaram abaixo da metade do desvio padrão dos dados observados em Q1 (28,79 m³/s), para todos os modelos. Já o MAE apresentou uma variação de 5,34 a 6,23 m³/s no C1 e 5,28 a 6,22 m³/s no C4.

De acordo com Richter et al. (2011), um NSE > 0,50 indica uma performance adequada para um modelo. Sendo assim, para o treinamento pode-se observar que, para os dois cenários, todos os modelos apresentaram bom desempenho, com valores de NSE, maiores que 0,75. Os valores de KGE se mantiveram acima de 0,80, o que, segundo estudo de Orellana-Alvear et. al (2020), representa um resultado satisfatório para a modelagem.

Em geral, Bozzoni et al. (2020) utilizam faixas que servem de guias para avaliar o comportamento da correlação entre variáveis. Segundo eles a correlação é moderada para valores de R acima de 0,5, forte para valores acima de 0,7 e muito forte para valores acima de 0,9. Tomando por base que o coeficiente de correlação R representa a raiz quadrada do coeficiente de determinação (R²) para modelos lineares, os valores encontrados para todos os modelos foram maiores que 0,70 e, portanto, podem ser considerados bons os ajustes obtidos. Fazendo-se uma análise comparativa de todos os modelos, verifica-se que o RF foi o que apresentou melhor desempenho para a maioria das métricas estatísticas analisadas no treinamento, tanto no C1 quanto no C4.

No teste, como era esperado, os resultados encontrados indicam uma piora de desempenho dos modelos nos cenários C1 e C4 para a maioria das métricas, quando comparado ao treinamento. O valor de PBIAS, apesar de ser pior para todos os modelos, na etapa de teste, em comparação com a etapa de treinamento para ambos os cenários (C1 e C4), manteve-se abaixo de 15%, ou seja, desempenho ainda considerado como aceitável. Apesar disso, para o RMSE, os valores encontrados foram inferiores ao do treinamento, enquanto o MAE apresentou variação de comportamento, com melhora para os modelos RLS, RLM e BRNN e alguma piora para os modelos CB e RF. Os valores do NSE, KGE e R² foram maiores que 0,5, 0,8 e 0,7, respectivamente, apenas para os modelos RLS, RLM e BRNN, demonstrando que os modelos CB e RF não apresentaram bons desempenhos na fase de teste.

Nota-se assim, que modelos mais simplificados como a RLS e a RLM por apresentarem performances satisfatórias, poderiam ser utilizados para fins de estimativa das vazões diárias na estação Q1.

Yaseen et al. (2016), aplicando modelos de aprendizado de máquina para previsão de séries temporais hidrológicas, também encontraram melhores desempenhos na fase de treinamento dos modelos em comparação com a fase de teste. Niu e Feng (2021) evidenciaram a mesma diferença de comportamento entre treinamento e teste em alguns dos modelos testados para previsão de séries temporais de vazões diárias.

De maneira geral, embora o modelo RF tenha apresentado melhores resultados para a etapa de treinamento, foi evidenciado que na etapa de teste se obteve melhor desempenho das previsões de vazões diárias utilizando o modelo RLS. Enquanto entre os modelos de aprendizado de máquinas trabalhados, o BRNN tendeu a superar os demais testados.

Esse resultado corrobora com os obtidos por Carbonneau et al. (2008), que também encontraram melhores desempenhos para regressão linear simples em comparação às técnicas de aprendizado de máquina para modelagem da demanda nas cadeias de abastecimento de produtos manufaturados. Porém no trabalho de Rassouli et al. (2011), no qual foram comparados três métodos de aprendizado de máquina (BRNN, regressão de vetor suporte – SRV e processo gaussiano – GP) com o modelo de regressão linear múltipla (RLM), eles observaram que os modelos não lineares em geral superaram a RLM.

No caso deste estudo, a elevada relação linear existente entre as vazões observadas nas estações fluviométricas Qd e Q1 se dá em decorrência da proximidade entre ambas, não sendo evidenciados fatores que pudessem ocasionar interferências nas características do escoamento de uma seção para outra, o que ajuda a explicar o melhor desempenho do modelo RLS em relação aos demais.

Na Figura 5 estão apresentados os valores encontrados para os índices estatísticos tanto no treinamento quanto no teste dos modelos para a estação fluviométrica Q1, considerando a adoção de critérios associados aos cenários C2 e C5. Os valores apresentados para a RLS se referem ao modelo cuja variável explicativa de melhor desempenho foi atribuída a Qd.

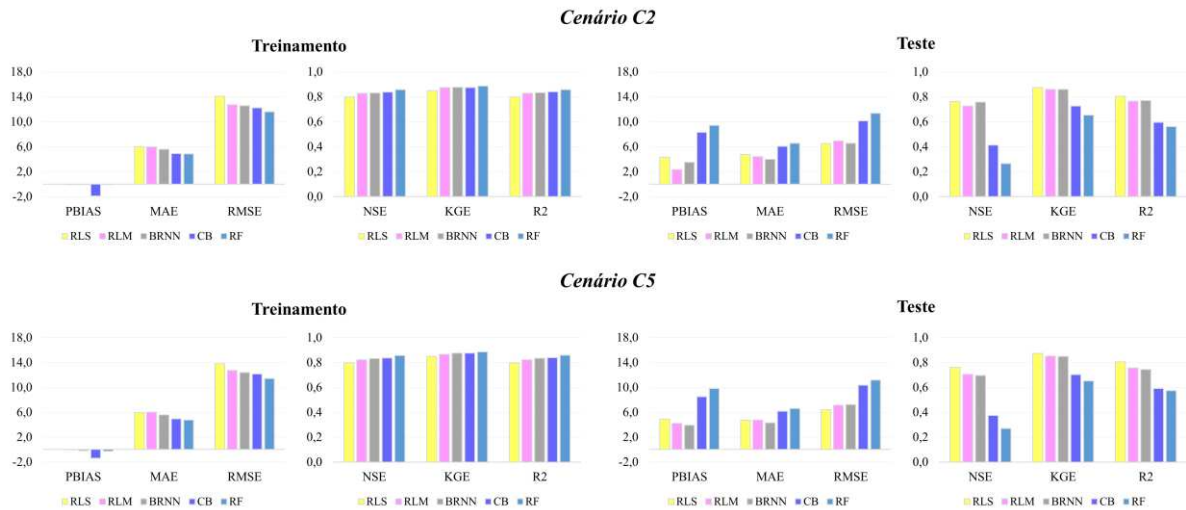


Figura 5 - Índices estatísticos obtidos no treinamento e no teste dos modelos de aprendizado de máquina (BRNN, CB e RF) e métodos de regressão linear (RLS e RLM) para estimativa da variável Q1, considerando os cenários C2 e C5.

A adoção do critério para seleção de dados nos cenários C2 e C5, ou seja, a remoção de registros de vazões a jusante (Q1) menores ou iguais a 90% das vazões a montante, provenientes da operacionalização do reservatório de regularização (Qd), refletiu positivamente no comportamento das métricas estatísticas avaliadas para a maioria dos modelos testados. O modelo RF, semelhante aos cenários C1 e C4, apresentou melhor performance na previsão das vazões no treinamento, apresentando menor RMSE (11,57 para C2 e 11,40 para C5) e maiores NSE (0,86 para C2 e C5), KGE (0,89 para C2 e C5) e R^2 (0,86 para C2 e C5).

Porém, para o conjunto de dados utilizados, os modelos RF e CB na fase de teste não apresentaram desempenhos satisfatórios para a previsão de vazões diárias na estação Q1, evidenciando valores de $NSE < 0,50$, $KGE < 0,80$ e $R^2 < 0,70$.

Na fase de teste, os modelos RLS, RLM e BRNN apresentaram desempenhos satisfatórios para a previsão de vazões diárias na estação Q1, considerando os cenários C2 e C5.

Dentre os modelos que apresentaram desempenho satisfatório, o BRNN apresentou o segundo melhor desempenho, para o qual foram observados os menores valores de MAE entre os modelos (3,95 m^3/s para C2 e 4,31 m^3/s para C5), e valores de RMSE iguais a 6,54 m^3/s para C2 e 7,22 m^3/s para C5, NSE de 0,76 para C2 e 0,70 para C5, KGE de 0,86 para C2 e 0,85 para C5 e R^2 de 0,77 para C2 e 0,74 para C5.

O modelo RLS foi aquele, dentre os favoráveis, que apresentou maior capacidade preditiva para a maioria das métricas avaliadas com RMSE de 6,47 m^3/s para o C2 e 6,40 m^3/s

para C5, NSE de 0,76 para C2 e C5, KGE de 0,87 para C2 e C5 e R^2 de 0,80 para C2 e 0,81 para C5.

Na Figura 6 são apresentados os índices estatísticos encontrados no treinamento e no teste dos modelos para a estação fluviométrica Q1, considerando a adoção dos critérios estabelecidos para os cenários C3 e C6.

Para os cenários C3 e C6, que contemplaram a eliminação de pouco mais de 10% dos dados utilizados em C1 e C4, observou-se uma melhora ainda mais significativa quanto ao desempenho de todos os modelos. Os valores apresentados para a RLS se referem ao modelo cuja variável explicativa de melhor desempenho foi atribuída a Qd.

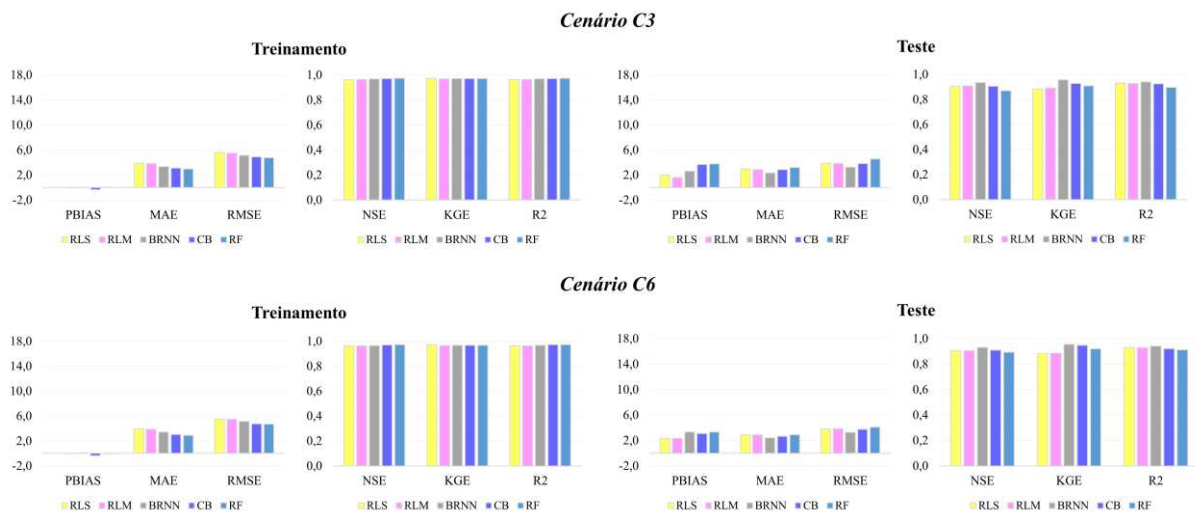


Figura 6 - Índices estatísticos obtidos no treinamento e no teste dos modelos de aprendizado de máquina (BRNN, CB e RF) e métodos de regressão linear (RLS e RLM) para estimativa da variável Q1, considerando os cenários C3 e C6

O modelo RF continuou a apresentar os melhores ajustes na fase do treinamento, para todas as métricas, com valor de PBIAS igual a 0%, MAE de 2,93 m³/s, para o C3 e 2,91 m³/s para C6, RMSE de 4,76 m³/s para C3 e de 4,71 m³/s para C6. No que se refere ao NSE, KGE e R^2 , foram encontradas alta performance para esse modelo, com valores de 0,97, 0,97 e 0,97, respectivamente e semelhantes tanto para o C3 quanto para o C6.

Na fase de teste os melhores desempenhos para a maioria da métricas avaliadas foram encontrados para o BRNN, apresentando os menores valores de MAE (2,31 m³/s para C3 e 2,41 m³/s para C6) e RMSE (3,22 m³/s para C3 e 3,27 m³/s para C6) e os maiores valores de NSE (0,93 para C3 e C6), R^2 (0,94 para C3 e C6) e KGE (0,96 para C3 e C6). Esse resultado demonstra, neste caso, certa superioridade do modelo de aprendizado de máquina em relação

aos de regressão linear, embora todos tenham apresentado métricas satisfatórias quanto a aplicabilidade das modelagens das vazões diárias na estação Q1, incluindo os modelos CB e RF.

Em função da expressiva melhora na performance dos modelos para os cenários C3 e C6 em relação aos C1, C2, C4 e C5, que se deu em função da remoção de dados de vazão diárias considerados inconsistentes para o comportamento do escoamento, é que foi considerado nesse estudo, como os cenários mais representativos da condição real, de maneira que a predição de vazões diárias na estação Q1 possa apresentar maior nível de confiabilidade.

Na Figura 7 é possível verificar que para a modelagem da vazão na estação Q1, considerando C3 e C6 como os melhores cenários encontrados, que a covariável Qd apresentou o mais alto valor de importância para todos os modelos testados, reforçando que as vazões a jusante do reservatório da UHE Queimado influenciam diretamente e fortemente o comportamento das vazões que escoam na estação Q1.

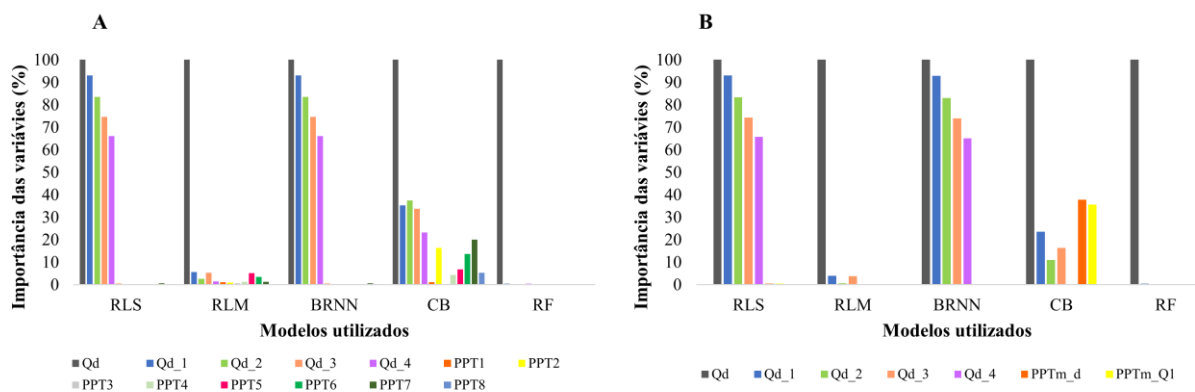


Figura 7 – Importância relativa das covariáveis selecionadas para cada modelo de aprendizado de máquina e de regressão linear utilizado para a predição de vazões na estação Q1, considerando os cenários C3 (a) e C6 (b).

Analisando especificamente o modelo BRNN, que apresentou melhor performance para os cenários C3 e C6, verifica-se que para o C3 e C6 as variáveis explicativas relacionadas a vazão defluente (Qd, Qd_1, Qd_2, Qd_3 e Qd_4) foram as que apresentaram maior importância para a modelagem da vazão na estação Q1. Apesar disso, a variável Qd foi a que apresentou importância superior em relação as demais, sendo a mais relevante para o desempenho do modelo.

Destaca-se, ainda, que para os cenários C3 e C6 a utilização de dados relacionados a precipitação (PPT) apresentaram importância reduzidas para os modelos, com maiores participações para o CB nos cenários C3 (Figura 7A) e C6 (Figura 7B), embora mantendo-se

pouco significativa. Os resultados das métricas estatísticas pouco diferiram entre si, indicando que, neste caso, a utilização de covariáveis associadas às precipitações não apresentou grandes vantagens para a melhoria das predições. Esse resultado não significa que o comportamento da precipitação não seja importante dentro do processo de formação das vazões, o que acontece é que nesse caso o comportamento do escoamento na seção onde se encontra localizada Q1 é muito mais influenciada pela proximidade da operação do reservatório a montante do que pela própria precipitação.

Na Figura 8 é apresentado, a título de exemplo, o desempenho dos modelos que mais se destacaram nos cenários C1, C2 e C3, por meio da comparação entre as séries históricas de vazões estimadas e observadas na estação Q1, tomando como base o ano de 2013, que foi utilizado em função da baixa quantidade de dados observados excluídos na estação fluviométrica Q1 para os três cenários.

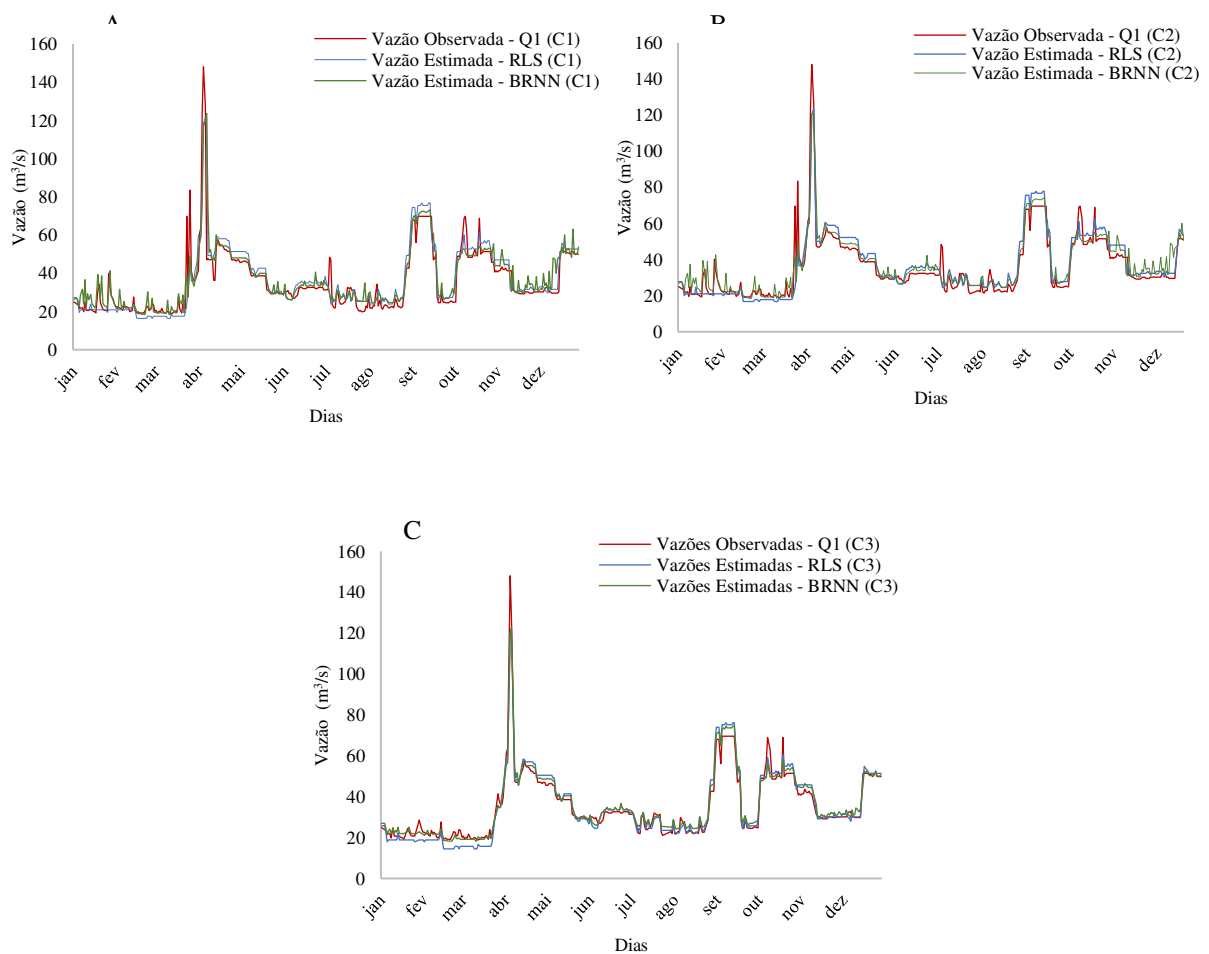


Figura 8 – Vazões observadas e estimadas para os modelos BRNN e RLS, na estação Q1, para os cenários C1 (A), C2 (B) e C3 (C) no ano de 2013.

Analisando a variação das vazões observadas e estimadas pelos modelos no ano de 2013, nota-se que, para o C1, o modelo RLS conseguiu se ajustar melhor aos valores de vazão da série histórica observada, quando comparado ao BRNN. Para alguns valores, o BRNN apresentou superestimativa do comportamento da vazão em relação ao RLS, especialmente de novembro a dezembro, além de subestimativa de valores máximos como no mês de outubro.

A situação encontrada no C1 se mostra bastante similar ao C2, de tal maneira que apesar da pequena diferença de performance (Figuras 4 e 5) entre os modelos BRNN e RLS, o RLS, de maneira geral, ainda se apresenta como o de melhor ajuste à série observada. Entretanto, ao observar o C3, se percebe uma visível melhora de desempenho do modelo BRNN em relação à RLS, com redução de valores superestimados e melhor previsão para os valores mínimos, resultando em um modelo com maior capacidade preditiva dentro desse cenário.

O C3, como mencionado apresentou melhor performance em função da remoção de dados considerados, nesse estudo, como incompatíveis ao escoamento da água, quando submetidos a atividade de regularização de reservatório. A implantação de reservatórios de regularização altera a distribuição das vazões nos rios e modifica o fornecimento de água a jusante da estrutura (Asfora e Cirilo, 2005), sendo que a frequência e magnitude das vazões em uma bacia, ou em parte dela, pode não estar associada apenas às condições hidrológicas, uma vez que o regime de operação do reservatório altera o volume de água a ser liberado, em um determinado intervalo de tempo, visando atender às demandas necessárias pelo sistema, especialmente quando se trata da geração contínua de energia.

Sendo assim, a real variabilidade diária no regime de escoamento das vazões pode não ser detectada pelas estações fluviométricas convencionais que estão sob influência da operação dos reservatórios, em decorrência da realização de apenas duas leituras linimétricas ao longo do dia (7 e 17h). Além disso, conforme menciona a ANA e RHA (2011), a qualidade dos dados obtidos nas séries históricas pode estar associada a diversos outros fatores como: a escolha do local de instalação das estações, a precisão dos instrumentos, a responsabilidade dos observadores, erros de digitação das cadernetas de campo e, destacadamente, o ajuste das curvas de calibração (curvas-chave).

Estudos como os de Virões e Cirilo (2019) tem relatado dificuldades na modelagem hidrológica quando se tem séries de vazões influenciadas pela presença de reservatórios, já que a qualidade dos dados nessas estações acaba sendo afetada.

Pode-se perceber, portanto, que a estação Q1, estação fluviométrica convencional localizada a jusante e próxima do reservatório de regularização da UHE Queimado apresentou um percentual expressivo de dados incoerentes de vazão, provavelmente decorrentes da baixa

representatividade das medições quanto ao regime de variação de vazões nas estações convencionais da rede hidrometeorológica nacional.

1.3.3. Desempenho dos modelos de aprendizado de máquinas e de regressão linear para Q2

Na Figura 9 estão apresentados os valores encontrados para os índices estatísticos, tanto no treinamento quanto no teste, dos modelos utilizados para estimativa da vazão da estação fluviométrica Q2, segunda estação fluviométrica convencional localizada a jusante do reservatório da UHE Queimado, considerando a adoção dos critérios associados aos cenários C1 e C4. Os valores apresentados para a RLS se referem ao modelo cuja variável explicativa de melhor desempenho foi atribuída a Qd_1.

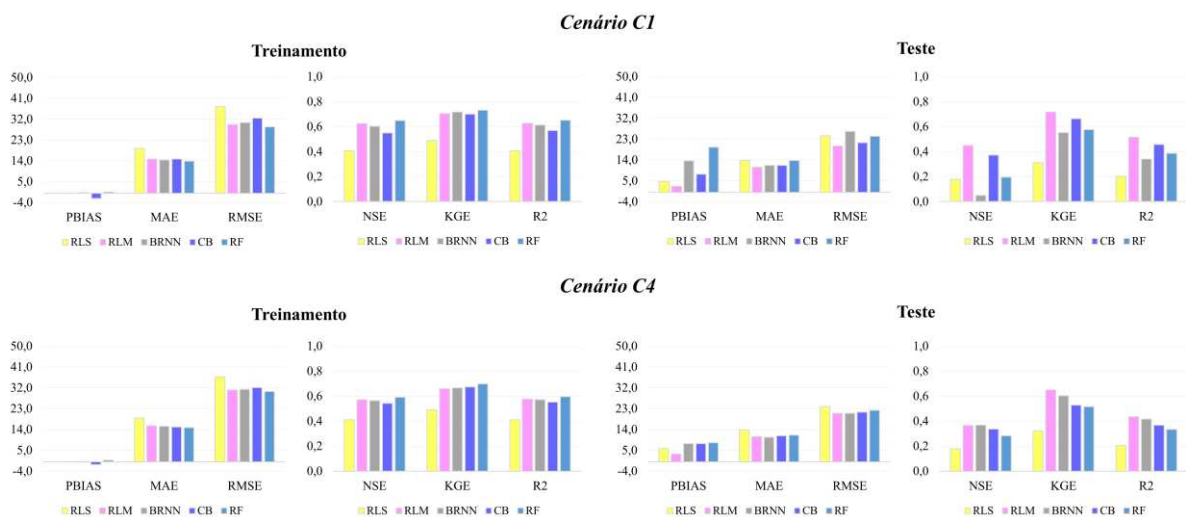


Figura 9 - Índices estatísticos obtidos no treinamento e no teste dos modelos de aprendizado de máquina (BRNN, CB e RF) e métodos de regressão linear (RLS e RLM) para estimativa da variável Q2, considerando os cenários C1 e C4.

Conforme se observa na Figura 9, os valores da maioria das métricas estatísticas demonstram baixo desempenho para todos os modelos, tanto na fase de treinamento quanto na de validação. Nenhum modelo apresentou valores de $KGE > 0,80$ e $R^2 > 0,70$. Além disso, os valores de RMSE encontrados para a maioria dos modelos, especialmente na fase de treinamento, foram maiores que a metade do desvio padrão dos dados observados em Q2 (46,38 m^3/s).

Os resultados indicariam, nesse caso, que as vazões defluentes do reservatório da UHE Queimado apresentariam pouca ou nenhuma influência sobre o escoamento da água nesta seção

da bacia, o que poderia ser justificado pela distância de aproximadamente 94 km entre as seções, seguindo pelo curso d'água, pelo incremento da área de drenagem em pouco mais de 27% ou, ainda, por problemas associados à calibração da curva-chave usada para estimativa das vazões que escoam na estação Q2.

Dessa forma, a aplicação da modelagem levando em consideração apenas a vazão defluente ao reservatório UHE Queimado, Qd, para a predição de vazões na estação Q2, sem a análise prévia e/ou remoção dos dados considerados incoerentes ao comportamento do escoamento não apresentou resultados satisfatórios para nenhum dos modelos testados, não sendo, portanto, recomendada sua utilização.

Na Figura 10 estão apresentados os valores encontrados para os índices estatísticos, tanto no treinamento quanto no teste, dos modelos para utilizados para estimativa da vazão da estação fluviométrica Q2, considerando agora a adoção de critérios associados aos cenários C2 e C5. Os valores apresentados para a RLS se referem ao modelo cuja variável explicativa de melhor desempenho foi atribuída a Qd_1.

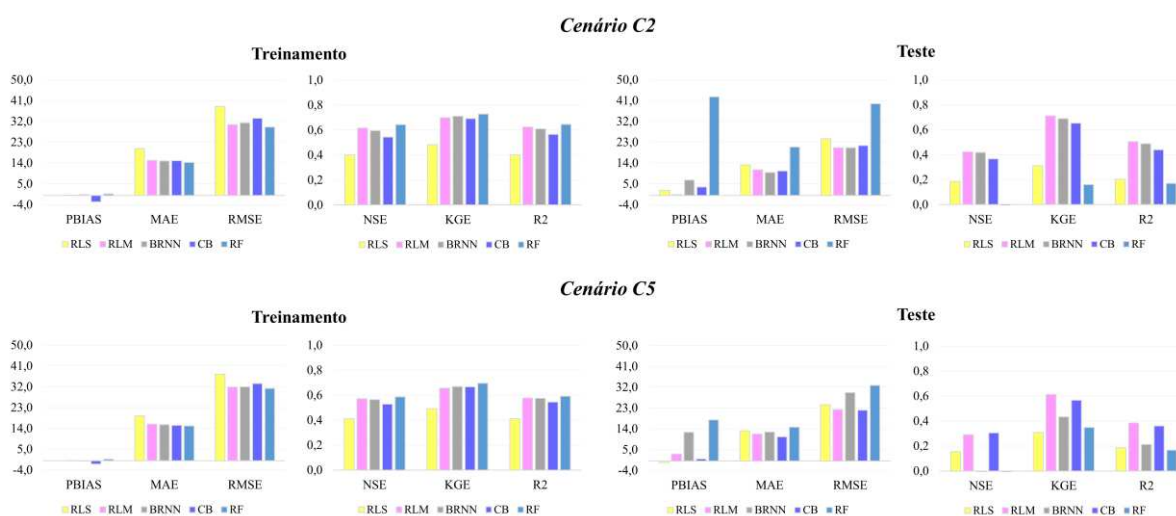


Figura 10 - Índices estatísticos obtidos no treinamento e no teste dos modelos de aprendizado de máquina (BRNN, CB e RF) e métodos de regressão linear (RLS e RLM) para estimativa da variável Q2, considerando os cenários C2 e C5.

Na Figura 10 é possível notar que não houve alteração considerável no desempenho dos modelos, mesmo com a remoção de alguns dados, ou seja, ao serem removidos apenas os dados em que a relação da vazão de Q2 e Qd_1 se apresentou menor ou igual a 90% os desempenhos permaneceram insatisfatórios para a previsão tanto no treinamento quanto na validação para todos os modelos avaliados, não sendo, portanto, recomendados para aplicação.

Por fim, na Figura 11 estão apresentados os valores encontrados para os índices estatísticos, tanto no treinamento quanto no teste, dos modelos utilizados para estimativa da vazão da estação fluviométrica Q2, considerando a adoção de critérios associados aos cenários C3 e C6. Os valores apresentados para a RLS se referem ao modelo cuja variável explicativa de melhor desempenho foi atribuída a Qd_1.

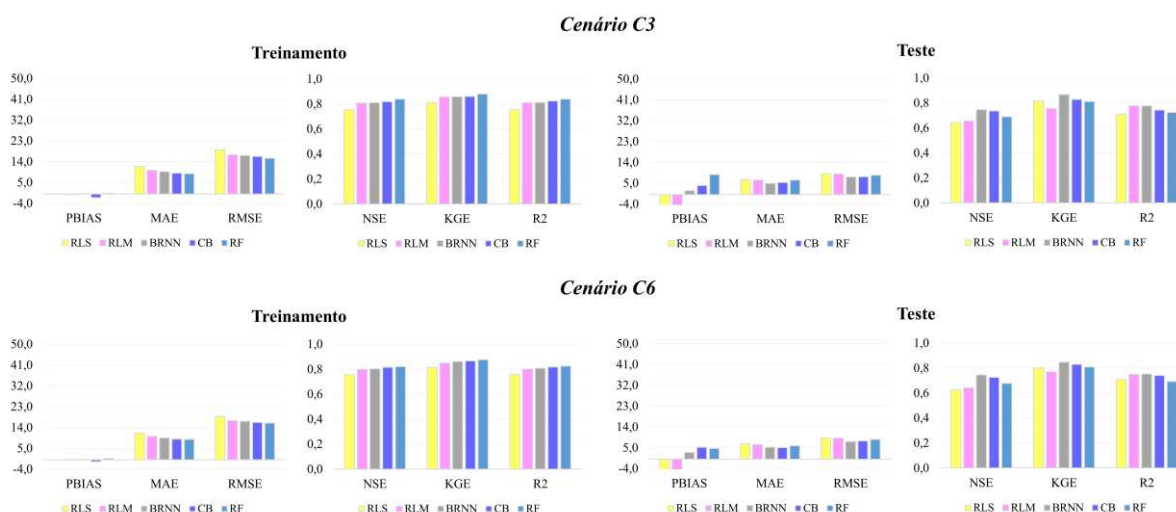


Figura 11 - Índices estatísticos obtidos no treinamento e no teste dos modelos de aprendizado de máquina (BRNN, CB e RF) e métodos de regressão linear (RLS e RLM) para estimativa da variável Q2, considerando os cenários C3 e C6.

Fica evidente, conforme verifica-se na Figura 11, que a remoção de cerca de 15% dos dados, considerados incoerentes ao regime de escoamento, de C1 para C3 e C4 para C6, proporcionou melhora substancial em todos os modelos de predição, destacando o RF na fase de treinamento e o BRNN no teste.

O BRNN apresentou certa superioridade em relação aos demais modelos de aprendizado de máquina e de regressão linear, com NSE de 0,75, KGE de 0,87 e R^2 de 0,78 para C3 e NSE de 0,75, KGE de 0,85 e R^2 de 0,75 para C6.

Apesar disso, todos os outros modelos apresentaram desempenho satisfatório e poderiam ser utilizados no processo de predição de vazões para Q2, sendo que em todos eles merece destaque o alto percentual de importância da variável Qd, conforme se verifica na Figura 12.

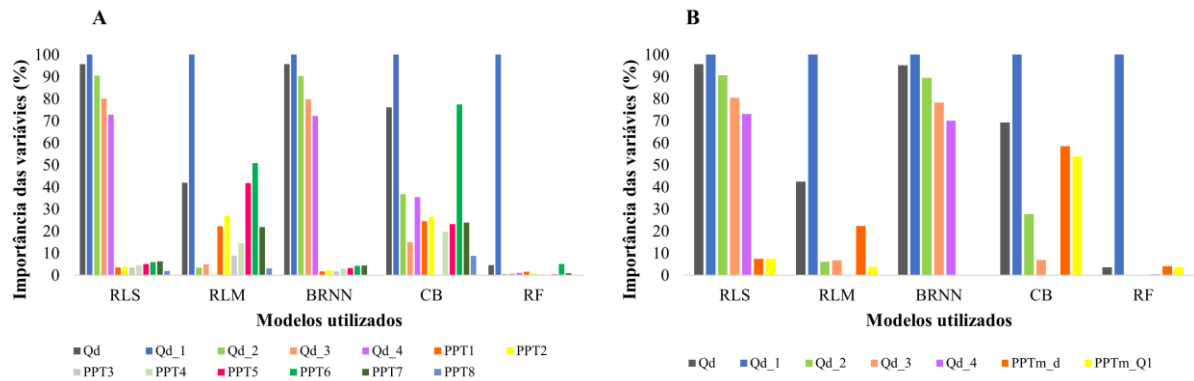


Figura 12 – Importância relativa das covariáveis selecionadas para cada modelo de aprendizado de máquina e de regressão linear utilizado para a predição de vazões na estação Q2, considerando os cenários C3 (A) e C6 (B).

A covariável Qd defasada um dia (Qd_1) foi a que apresentou maior importância para quase todos os modelos de predição utilizados nos cenários C3 e C6, com especial atenção para o BRNN, modelo que apresentou melhor performance. Esse resultado reforça que, assim como Q1, a estação Q2 tem o seu escoamento influenciado pelas atividades relacionadas à operação do reservatório da UHE Queimado, um dia após a sua defluência.

Esses resultados indicam, ainda, que a utilização do PPT_i e PPT_m, diferentemente da estação Q1, apresentaram percentuais variáveis e maiores de importância no desempenho de alguns dos modelos. Entretanto, para o BRNN a importância das precipitações continuou pouco expressiva no C3 e nem mesmo foram detectadas para o C6. De modo geral, ao se comparar a performance dos modelos para os cenários C3 e C6 verificou-se, embora por diferenças muito pequenas, valores mais satisfatórios de desempenho, na sua maioria, para o C3.

Os resultados encontrados reforçam que nenhum modelo, por mais robusto e eficaz que seja, poderá apresentar bons desempenhos se a qualidade do banco de dados a ser utilizado na modelagem for baixa ou se o mesmo for pouco representativo da realidade da bacia hidrográfica. Para o preenchimento de falhas em estações sob a influência de reservatórios de regularização, a utilização de dados de vazão defluente coletados em estações telemétricas instaladas nos reservatórios pode ser mais adequada e segura para a predição de vazões diárias em seções localizadas a jusante.

Dessa forma, fica clara a necessidade de uma análise estrutural com relação à qualidade dos dados que se dispõe, de tal maneira que seja possível, identificar, quando houver, registros de vazões que sejam incompatíveis com o comportamento da bacia. A partir da remoção dessas inconsistências, os dados passam a ser tratados também como falhas e, dessa forma, devem ser preenchidos com base no melhor modelo ajustado visando melhorar a qualidade das informações.

1.3.4. Desempenho dos modelos de aprendizado de máquinas e de regressão linear para Q3

Os índices estatísticos obtidos no treinamento e no teste dos modelos utilizados para a estimativa da vazão da estação fluviométrica Q3, terceira estação fluviométrica convencional localizada a jusante do reservatório da UHE Queimado, considerando a adoção de critérios associados aos cenários C1 e C4 e C2 e C5 são apresentados na Figura 13.

Os resultados da Figura 13 foram obtidos a partir da consideração das covariáveis relacionadas às precipitações (G_{PPTi} e G_{PPTm}) e, exclusivamente, das vazões de defluência do reservatório da UHE Queimado (Q_d) associadas a diferentes defasagens no tempo. Os valores apresentados para a RLS se referem ao modelo cuja variável explicativa de melhor desempenho foi atribuída a Q_d_1 .

A Figura 13 evidencia que, mesmo retirando os dados considerados inconsistentes para essa análise, não houve ajuste adequado de nenhum dos modelos testados para a estação Q3, nem na fase de treinamento nem no teste.

A estação Q3 está a uma distância muito maior, quando comparada a Q2, do reservatório de regularização UHE Queimado, cerca de 150 km seguindo o trajeto do curso de água/. A área de drenagem incremental de Q3 em relação a Q_d é de mais de 60%, apresentando, dessa forma, maiores variabilidades fisiográficas, geológicas e hidrológicas entre as seções, conforme já mencionado neste estudo. Além disso, é verificado um aumento substancial da demanda de água na área de contribuição da estação Q3 pela atividade de irrigação na região, especialmente pelo uso de pivôs centrais. Essas evidências podem, portanto, ajudar a justificar a redução no grau de influência do comportamento das vazões defluentes no regime de escoamento das vazões nessa estação.

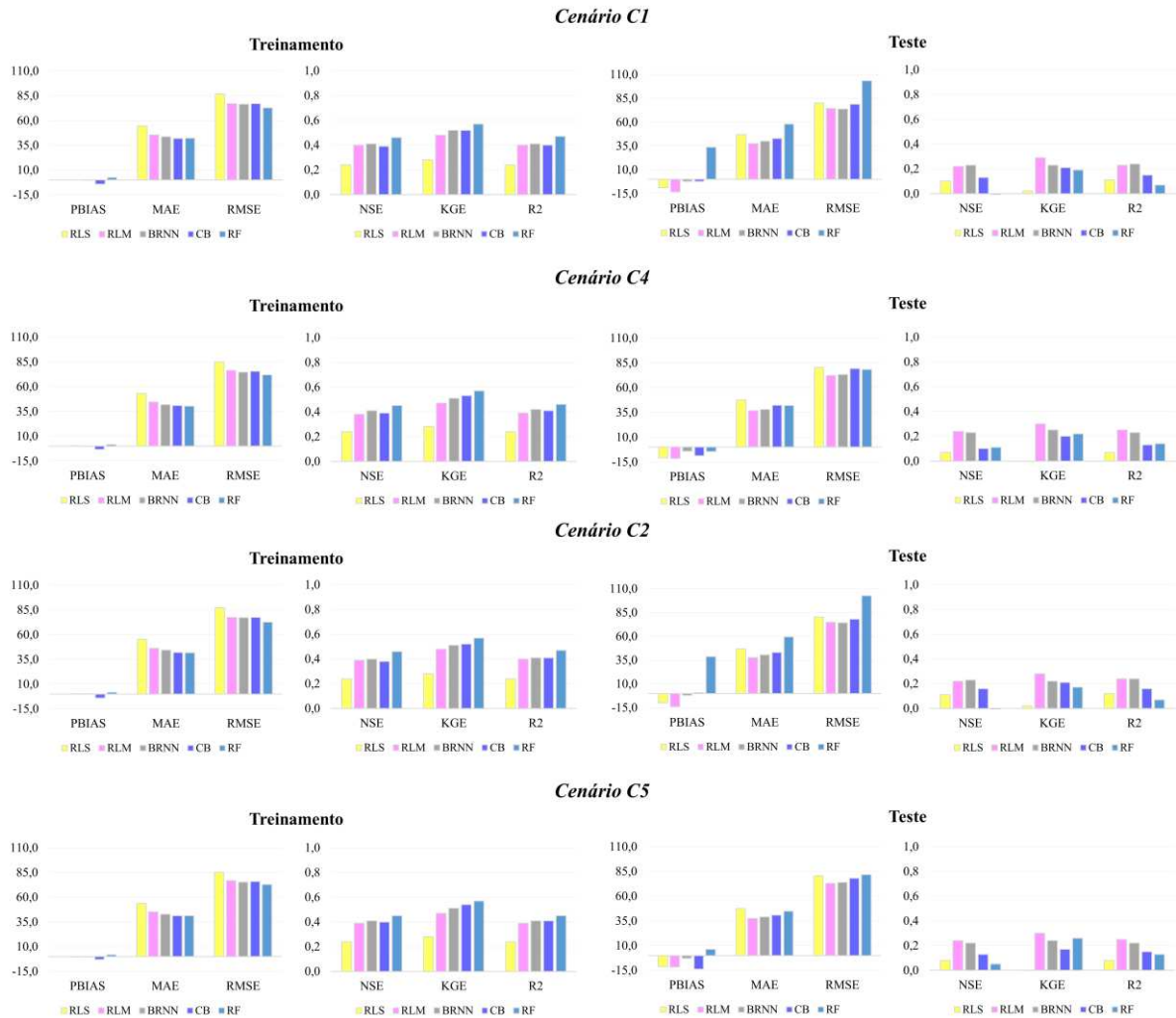


Figura 13 - Índices estatísticos obtidos no treinamento e no teste dos modelos de aprendizado de máquina (BRNN, CB e RF) e métodos de regressão linear (RLS e RLM) para estimativa da variável Q3, considerando os cenários C1 e C4 e C2 e C5, utilizando como variáveis explicativas as precipitações (PPT_i e PPT_m) e as vazões de defluência do reservatório (Qd)

Por esse motivo, para esta estação foi realizada uma nova modelagem utilizando não apenas dados de vazão defluente (Qd) na UHE Queimado, mas também as séries históricas observadas nas estações Q1 e Q2, conjuntamente com as informações relativas às precipitações (PPT_i e PPT_m).

Dessa forma, para os cenários C1 e C4 foram considerados todos os dados disponíveis de Q3 e as covariáveis Q1 e Q2, considerando as defasagens temporais, e PPT_i e PPT_m, variando entre os cenários apenas o tipo de dado associado à precipitação, como nas demais modelagens já realizadas para Q1 e Q2. Para C2 e C5 foram removidos, conforme mostrado na Figura 3, apenas os valores de vazão que apresentaram relação entre Q3/Qd₁ menores ou iguais a 1,0.

Na Figura 14 são apresentados os resultados das métricas estatísticas obtidas no treinamento e no teste dos modelos de aprendizado de máquina (BRNN, CB e RF) e métodos

de regressão linear (RLS e RLM) para estimativa da variável Q3, considerando os cenários C1 e C4 e C2 e C5, utilizando como variáveis explicativas as precipitações (PPTi e PPTm) e as vazões Qd, Q1 e Q2, incluindo as defasagens temporais. Os valores apresentados para a RLS se referem ao modelo cuja variável explicativa de melhor desempenho foi atribuída a Q2_1.

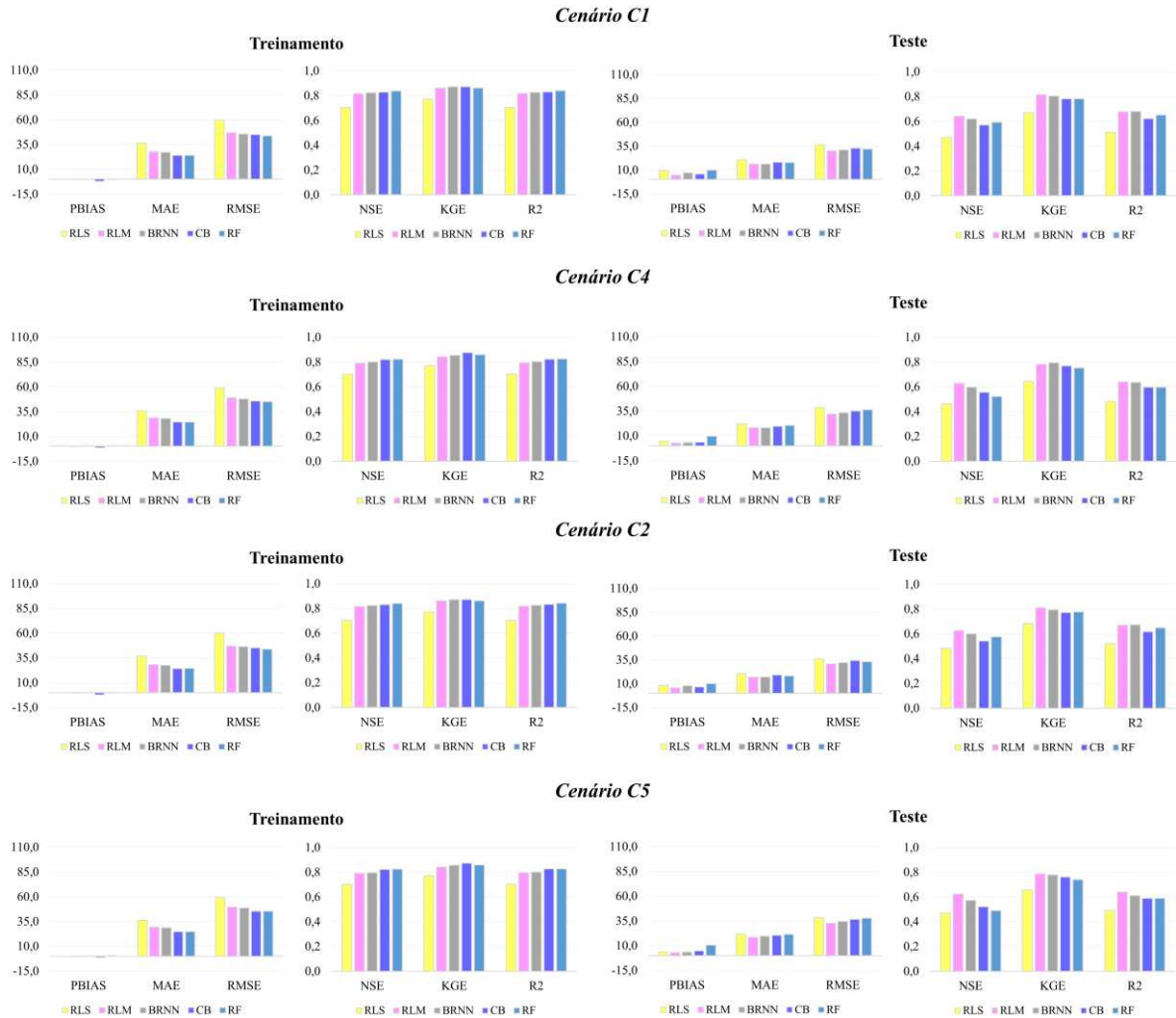


Figura 14 - Índices estatísticos obtidos no treinamento e no teste dos modelos de aprendizado de máquina (BRNN, CB e RF) e métodos de regressão linear (RLS e RLM) para estimativa da variável Q3, considerando os cenários C1 e C4 e C2 e C5, utilizando como variáveis explicativas as precipitações (PPTi e PPTm) e as vazões Qd, Q1 e Q2, incluindo as defasagens temporais

Observando os resultados da Figura 14, verifica-se que na modelagem das vazões na estação Q3 ocorreu uma melhora expressiva nas métricas estatísticas estabelecidas como referência para análise de desempenho de cada um dos modelos testados, quando foram considerados as séries de vazões de outras estações fluviométricas. Para os cenários C1 e C4 é possível visualizar acentuada melhora na fase de treinamento dos modelos, com métricas estatísticas que refletem desempenho satisfatório. Os valores de NSE foram maiores para RF,

seguido de CB, BRNN, RLM e RLS, sendo todos acima ou igual a 0,70. Para o KGE os maiores valores encontrados foram para o modelo CB e para R^2 os melhores desempenhos foram para o modelo RF, que também apresentou o menor RMSE entre os modelos estudados.

Já na fase de teste, todos os modelos apresentaram piora no desempenho, entretanto, os modelos RLM e BRNN proporcionaram valores de NSE e KGE que se mantiveram acima de 0,50 e 0,80, respectivamente, quando se considerou a PPT_i , apesar do valor de R^2 ter se apresentado no C1, mesmo que por uma diferença muito pequena, abaixo de 0,70. Dessa forma, tais modelos foram considerados, pelo estudo, com potencial para utilização em procedimentos associados ao preenchimento de falhas da estação Q3, visto a elevada complexidade e dinâmica do regime de escoamento que a rede hidrográfica da bacia do rio Preto possui.

A Figura 15 reforça os resultados encontrados na Figura 14, mostrando que as covariáveis relacionadas à estação Q2 foram as que apresentaram maiores importâncias para todos os modelos, sendo Q2_1 e Q2_2 as mais expressivas. As variáveis relacionadas a Qd, embora tenham apresentado certa importância para a predição das vazões pelos modelos, foram consideravelmente menos relevantes quando comparadas às vazões em Q2.

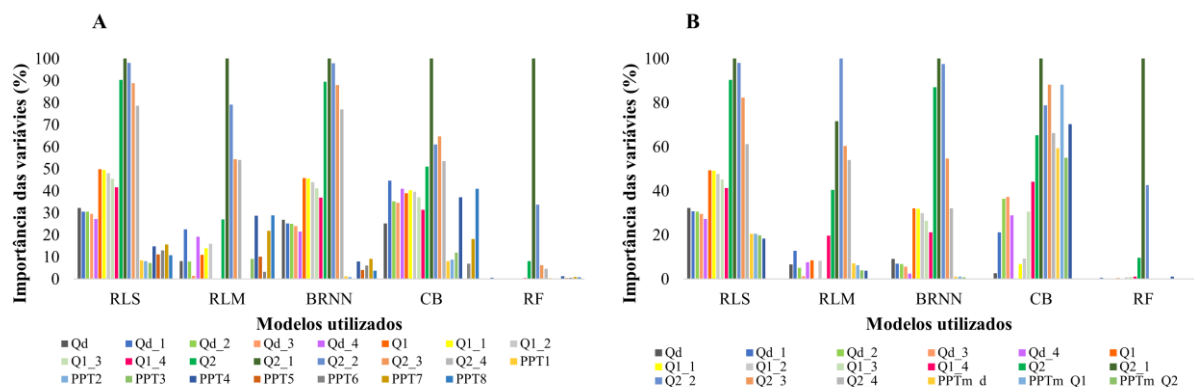


Figura 15 – Importância relativa das covariáveis selecionadas para cada modelo de aprendizado de máquina e de regressão linear utilizado para a predição de vazões na estação Q3, considerando os cenários C1 (A) e C4 (B).

Avaliando os cenários C2 e C5, verificou-se que não houve melhora nas estimativas, pelo contrário, foi possível perceber uma pequena perda de desempenho em relação a C1 e C4 para algumas métricas. Em C2 e C5, dados em que a relação entre Q3 e Qd_1 se apresentaram menores ou iguais a 1,0 foram removidos, significando que onde essas relações se estabeleceram, todos os dados das demais covariáveis daquele dia específico foram eliminadas do banco de dados, o que pode, de alguma maneira, ter promovido a retirada de dados de boa

qualidade relacionados às demais covariáveis utilizadas, como é o caso de Q1 e Q2, por exemplo, proporcionando alguma pequena piora no desempenho dos modelos.

Sendo assim, embora os modelos de aprendizado de máquina sejam capazes de aprender sistemas não lineares complexos a partir de um conjunto de observações, a obtenção dos resultados apresentados demonstra que a qualidade do conjunto de dados e o nível de dependência entre as variáveis podem viabilizar a implementação de modelos mais simplificados e com bom nível de ajuste, como no caso do RLS.

Neste contexto, quanto melhor for a qualidade dos dados das séries históricas de vazão, especialmente dos conjuntos de dados influenciados pela operação de reservatórios de regularização, melhores serão os resultados em quaisquer análises hidrológicas.

Contudo não é possível definir especificamente quais as melhores variáveis explicativas para a realização de modelagens de previsão de vazões diárias em estações fluviométricas convencionais, embora tenha se observado que a utilização de uma estação telemétrica de suporte a montante apresenta grandes chances de ser um bom preditor, especialmente em estações mais próximas, atentando-se ao fato de que se deve estabelecer o melhor período de defasagem de uma vazão em relação a outra.

1.4. CONCLUSÕES

Com base nos resultados obtidos no presente trabalho, conclui-se que:

- As covariáveis de maior importância para explicar o comportamento das vazões nos modelos de previsão para as estações fluviométricas convencionais Q1 e Q2 foram a vazão defluente ao reservatório UHE Queimado (Qd) e a vazão defluente com um dia de defasagem (Qd_1) respectivamente.
- Para a estação mais distante do reservatório, Q3, a covariável vazão defluente (Qd) teve reduzida influência na modelagem das vazões diárias, levando, conseqüentemente, à necessidade de inserção de outras variáveis explicativas nos modelos de previsão, destacando, neste caso, a Q2 como a de maior importância.
- As variáveis associadas à precipitação apresentaram pouca ou nenhuma influência na performance dos modelos de previsão de vazão para a estação Q1. Para Q2 e Q3 a utilização

das séries de precipitação média diária provenientes das estações pluviométricas localizadas dentro da bacia (PPT_i) demonstrou certa superioridade em relação à precipitação média na área de drenagem das estações fluviométricas (PPT_m).

- O modelo baseado em redes neurais (BRNN) apresentou a melhor performance para as predições de vazões diárias para as estações Q1 e Q2, localizadas a jusante e próximas do reservatório da UHE Queimado, quando se considerou o pré-tratamento dos dados.

- O modelo de Regressão Linear Múltipla (RLM) apresentou o melhor desempenho para predição das vazões na estação Q3, a mais distante do reservatório da UHE Queimado, sem que fosse necessário a realização do pré-tratamento dos dados.

1.5. REFERÊNCIAS

- Althoff, D., Rodrigues, L. N. 2019. The expansion of center-pivot irrigation in the Cerrado biome. Embrapa Cerrados-Nota Técnica/Nota Científica (ALICE).
- Althoff, D., Rodrigues, L. N., Silva, D. D. da. 2020a. Impacts of climate change on the evaporation and availability of water in small reservoirs in the Brazilian savannah. *Climatic Change*, v. 159, n. 2, p. 215-232.
- Althoff, D., Filgueiras, R., Rodrigues, L. N. 2020b. Estimating small reservoir evaporation using machine learning model for the Brazilian Savannah. *Journal Hydrologic Engineering*, n. 25, v. 8.
- Agência Nacional de Águas (ANA). 2017. Atlas irrigação: uso da água na agricultura irrigada. Brasília. 86p.
- Agência Nacional de Águas (ANA) e RHA Engenharia e Consultoria SS Ltda. 2011. Qualificação de dados hidrológicos e reconstituição de vazões naturais no país. Resumo Executivo. Brasília. 66p.
- Agência Nacional de Águas (ANA). UHE Queimado (GO/MG) terá redução da vazão mínima defluente até 31 de outubro. 2019. Disponível em: <<https://www.ana.gov.br/noticias-antigas/uhe-queimado-go-mg-tera-reducao-da-vazao-minima-2019-03-15.2354150460>>. Acesso em: 21 out 2020.
- Asfora, M. C., Cirilo, J.A. 2005. Reservatórios de regularização: alocação de água para usos múltiplos com diferentes garantias. *Revista de Gestão de Água da América Latina*, v.2, p 27-38.
- Bozzini, P.L., Mello Junior, A.V. 2020. Previsões de precipitação de modelos atmosféricos como subsídio à operação de sistemas de reservatório. *Revista Brasileira de Meteorologia*, v. 35, n. 1, p. 99-109.

- Breiman, L. 2001. Random Forests. *Machine Learning*, v. 45, n. 1, p. 5-32.
- Bui, D. T., Khosravi, K., Tiefenbacher, J., Nguyen, H., Kazakis, N. 2020. Improving prediction of water quality indices using novel hybrid machine-learning algorithms. *Science of The Total Environment*, v. 721.
- Carbonneau, R., Laframboise, K., Vahidov, R. 2008. Application of machine learning techniques for supply chain demand forecasting. *European Journal of Operational Research* 184. doi:10.1016/j.ejor.2006.12.004.
- CEMIG – Companhia Energética de Minas Gerais. UHE Queimado. Disponível em: <<https://novoportal.cemig.com.br/usina/uhe-queimado/>>. Acesso em: 21 out 2020.
- CEMIG – Companhia Energética de Minas Gerais. 2020. Alterações hidrológicas a jusante de uma hidrelétrica (livro eletrônico): efeitos ecológicos e proposições. Belo Horizonte.
- Corato, G., Moramarco, T., Tucciarelli, T. 2011. Discharge estimation combining flow routing and occasional measurements of velocity. *Hydrology and Earth System Sciences*, v. 15, n. 9, p. 2979-2994.
- Cutler, D. R., Edwards, T. C., Beard, K. H., Cutler, A., Hess, K. T., Gibson, J., Lawler, J. J. 2007. Random Forests for classification in ecology. *Ecology*, 88(11), 2783–2792. doi:10.1890/07-0539.1.
- Dhingra, 2017. Streamflow Hydrology Estimate Using Machine Learning (SHEM). *Journal of the American Water Resources Association (JAWRA)* 1-14. doi:10.1111/1752-1688.12555.
- Elshorbagy, Amin A., Panu, U. S., Simonovic, S. P., 2000. Group-based estimation of missing hydrological data: I. Approach and general methodology. *Hydrological Sciences Journal*, [s.l.], 45(6): 849-866, doi: <http://dx.doi.org/10.1080/02626660009492388>.
- Embrapa – Empresa Brasileira de Pesquisa Agropecuária. 2005. Excursão caracteriza relevo e uso da bacia do rio Preto. Disponível em: < <https://www.embrapa.br/busca-de-noticias/-/noticia/17981159/excursao-caracteriza-relevo-e-uso-da-bacia-do-rio-preto>>. Acesso em 10 de fevereiro de 2021.
- Freitas, M. C., Silva, G. B. L da. 2011. Estimativa de vazões diárias em rios a partir de vazões a montante usando redes neurais artificiais. XIX Simpósio Brasileiro de Recursos Hídricos. Maceió – AL.
- Garson, G.D. 1991. Interpreting neural network connection weights. *Artificial Intelligence Expert* 6, 47/51.
- Granemann, A. R. B., Mine, M. R. M., Kaviski, E. 2018. Frequency analysis of minimum flows. RBRH. Porto Alegre.
- Gianfagna, C. C., Johnson, C. E., Chandler, D. G., Hofmann, C. 2015. Watershed area ratio accurately predicts daily streamflow in nested catchments in the Catskills, New York. *Journal of Hydrology: Regional Studies* 4, 583-594. doi: 10.1016/j.ejrh.2015.09.002.

- Gorbani, M. A., Deo, R. C., Km, S., Kashani, M. H., Karimi, V., Izadkhah, M. 2020. Development and evaluation of the cascade correlation neural network and the random forest models for river stage and river flow prediction in Australia. *Soft Computing*. doi:10.1007/s00500-019-04648-2.
- Guimarães, L. M. S., Meireles, M. R. G., Almeida, P. E. M de. 2019. Avaliação das etapas de pré-processamento e de treinamento em algoritmos de classificação de textos no contexto da recuperação da informação. *Perspectivas em Ciência de Informação*. v. 24, n. 1, Belo Horizonte.
- Kamwaga, S., Mulungu, D. M. M., Valimba, P. 2018. Assessment of empirical and regression methods for infilling missing streamflow data in Little Ruaha catchment Tanzania, *Physics and Chemistry of the Earth, Parts A/B/C*, v. 106, 17-28, ISSN 1474-7065, <https://doi.org/10.1016/j.pce.2018.05.008>.
- Kingston, G.B., Maier, H.R., Dandy, G.C. 2018, Review of Artificial Intelligence Techniques and their Applications to 25 Hydrological Modeling and Water Resources Management. Part 1 – Simulation, disponível em: <https://www.researchgate.net/publication/277005048_Review_of_Artificial_Intelligence_Techniques_and_their_Applications_to_Hydrological_Modeling_and_Water_Resources_Management_Part_1_-_Simulation>, acesso em: 27 de janeiro de 2021.
- Kuhn, M. Weston, S., Keefer, C., Coutler, N. 2012. Cubist models for regression. Acesso: abril de 2021.
- Kuhn, M., et al. 2018. Caret: Classification and regression training. Acesso: novembro de 2020.
- Li, X. Sha, J., Wang, Z-L. 2019. Comparison of daily streamflow forecasts using extreme learning machines and the random forest method. *Hydrological Sciences Journal*, doi: 10.1080/02626667.2019.1680846.
- Liaw, A., M. Wiener. 2002. “Classification and regression by random- Forest.” *R News* 2 (3): 18–22.
- Lu, G. Y., Wong, D. W. 2008. An adaptive inverse-distance weighting spatial interpolation technique. *Computers and Geosciences*, v. 34, n. 9, p. 1044- 1055, doi: 10.1016/j.cageo.2007.07.010.
- Ma, L., Ahuja, L., Nolan, B., Malone, R., Trout, T., Qi, Z., 2012. Root Zone Water Quality Model (RZWQM 2): Model use, calibration, and validation. *Trans. ASABE*, 55(4), 1425- 1446. <http://dx.doi.org/10.13031/2013.42252>.
- Maillard, P., Pereira, E. O. 2017. Uma ferramenta para processar dados de altimetria por satélites para aplicações em hidrologia. *ABRH - XXII Simpósio Brasileiro de Recursos Hídricos*. Florianópolis.
- Marques, Y. B. 2015. Mirnacle: Aprendizagem de máquina utilizando smote e Random Forest para prover aumento da seletividade na predição ab initio de pré-miRNAs. Dissertação

de Mestrado – Pós-Graduação em Ciência da Computação, Universidade Federal de Viçosa.

- Mello, C. R.; Silva, A. M. 2013. Hidrologia: princípios e aplicações em sistemas agrícolas. Lavras: Ed. UFLA, 455p.
- Mewes, B., Opiel, H, Hartmann, A. J., Schumann, A. H. 2018. Looking Beyond Data: Information-theory based criteria to evaluate different machine-learning approaches for hydrological problems. American Geophysical Union – 146A.
- Mishra, P. 2019. PyTorch Recipes: A problem-solution approach, eletronic. ed, PyTorch Recipes. Apress, Bangalore.
- Molina, Z. B. de la C. 2016. Avaliação da acurácia do modelo de simulação a usinas individualizadas baseadas no pós operativo. Dissertação de Mestrado – Pós-Graduação em Engenharia Elétrica, Universidade Estadual de Campinas.
- Mosavi, A., Ozturk, P., Chau, K-W. 2018. Flood prediction using machine Learning models: literature review. Water.
- Naghetini, M., Pinto, E. J. A. 2007. Hidrologia Estatística. 1. ed. Belo Horizonte: CPRM – Companhia de Pesquisa de Recursos Minerais – Superintendência Regional de Belo Horizonte.
- Noi, P. T., Degener, J., Kappas, M. 2017. Comparison of multiple linear regression, cubist regression, and random forest algorithms to estimate daily air surface temperature from dynamic combinations of MODIS LST data. Remote Sensing.
- Niu, W., Feng, Z. 2021. Evaluating the performances of several artificial intelligence methods in forecasting daily streamflow time series for sustainable water resources management. Sustainable Cities and Society, 64, 102562. doi:10.1016/j.scs.2020.102562.
- Obringer, R., Nateghi, R. 2018. Prediction urban reservoir levels using statistical learning techniques. Scientific Reports.
- Oliveira, F. A., Pereira, T. S. R., Soares, A. K., Formiga, K. T. M. 2016. Uso de modelo hidrodinâmico para determinação da vazão a partir de medições de nível. RBRH, v. 21, n. 4, Porto Alegre.
- Orellana-Alvear, J., Célleri, R., Rollenceck, R. Muñoz, P., Contreras, P., Bendix, J. 2020. Assessment of Native Radar Reflectivity and Radar Rainfall Estimates for Discharge Forecasting in Mountain Catchments with a Random Forest Model. Remote Sensing, 12 doi:10.3390/rs12121986.
- OMM - Organização Meteorológica Mundial. 1994. Guia para práticas hidrológicas: Aquisição e processamento de dados, análises, estimativas e outras aplicações, 15.ed., 735p.
- Pal, M., Deswal, S. 2009. M5 model tree based modelling of reference evapotranspiration. Hydrol. Process. 23 (10), p. 1437–1443.

- Pini, M., Scalvini, A., Liaqat, M. U., Ranzi, R., Serina, I., Mehmood, T. 2020. Evaluation of machine learning techniques for inflow prediction in lake Como, Italy. *Procedia Computer Science*. 176, p. 918-927. doi: <https://doi.org/10.1016/j.procs.2020.09.087>.
- Quinlan. J. R. 1992. Learning with continuous classes. *Proceedings of the 5th Australian Joint Conference On Artificial Intelligence*. pp. 343-348.
- Raouli, K., Hsieh, W. W., Cannon, A. L. 2011. Daily streamflow forecasting by machine learning methods with weather and climate inputs. *Journal of Hydrology* 414-415. doi: 10.1016/j.jhydrol.2011.10.039.
- Reis, G. B., Silva, D. D. da., Fernandes Filho, E. I., Moreira, M. C., Veloso, G. V., Fraga, M. de S., Pinheiro, S. A. R. 2021. Effect of environmental covariable selection in the hydrological modeling using machine learning models to predict daily streamflow. *Journal of Environmental Management*. doi:10.1016/j.jenvman.2021.112625.
- Richter, K., Hank, T. B., Atzberger, C., Mauser, W. 2011. Goodness-of-fit measures: What do they tell about vegetation variable retrieval performance from Earth observation data. *Remote sensing for agriculture, ecosystems, and hydrology XIII*, 81740R. Prague, Czech Republic: International Society for Optics and Photonics.
- Rodriguez, P. P., Gianola, D. 2018. BRNN: Bayesian regularization for feed-forward neural networks. Disponível em: <https://cran.r-project.org/web/packages/brnn/> Acesso em dezembro de 2020.
- Saravanan, M., Nagoor Meeran, A. R. 2020. Improving the analysis rate of ground level water prediction using different machine learning models. *Journal of Critical Reviews*, v.7, n. 4.
- Santos, H. G. dos., Carvalho Junior, W. de., Dart, R. de O., Aglio, M. L. D., Sousa, J. S. de., Pares, J. G., Fontana, A., Martins, A. L. da S., Oliveira, A. P. de. 2011. O novo mapa de solos do Brasil: legenda atualizada. Rio de Janeiro: Embrapa Solos, 67 p. (Embrapa Solos. Documentos, 130.) 1 mapa, color. Escala 1:5.000.000. Disponível em: <<https://ainfo.cnptia.embrapa.br/digital/bitstream/item/123772/1/DOC-130-O-novo-mapa-de-solos-do-Brasil.pdf>>, acesso em abril de 2021.
- SEMAD – Secretária do Estado de Meio Ambiente e Desenvolvimento Sustentável. Licenciamento Ambiental, PARECER ÚNICO N° 1137215/2017 (SIAM). Disponível em: <sistemas.meioambiente.mg.gov.br > reunioes > uploadsUnai Baixo – Semad>. Acesso em: 21 out 2020.
- Shortridge, J., Guikema, S. D., Zaitchik, B. F. 2016. Machine learning methods for empirical streamflow simulation: a comparison of model accuracy, interpretability, and uncertainty in seasonal watersheds. *Hydrol. Earth. Syst. Sci*, 20, p. 2611-2628. Doi: 10.5194/hess-20-2611-2016.
- Souza, L. F. T. de. 2017. Modelagem de processos erosivos em área de contatos geológicos no Alto Paranaíba. 133 f. Dissertação (Mestrado em Solos e Nutrição de Plantas) - Universidade Federal de Viçosa, Viçosa. 2017.

- Ticknor, J. L. 2013. A Bayesian regularized artificial neural network for stock market forecasting. *Expert Systems with Applications*, 40(14):5501-5506.
- Tyralis, H., Papacharalampous, G., Langousis, A. 2019. A brief review of Random Forests for water scientists and practitioners and their recent history in water resources. *Water*, 11(5).
- Virões, M. V., Cirilo, J. A. 2019. Regionalization of hydrological model parameters for the semi-arid region of the northeast Brazil. *Brazilian Journal of Water Resources*, v.29, e49, Porto Alegre. doi:10.1590/2318-0331.241920180114.
- Willmott, C. J., Matsuura, K. 2005. Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance. *Climate Research*, v. 30, n. 1, p. 79–82.
- Yaseen, Z. M., Allawi, M. F., Yousif, A. A., Jaafar, O., Hamzah, F. M., and El-Shafie, A. 2016. Non-tuned machine learning approach for hydrological time series forecasting. *Neural Computing and Applications*, 30(5), 1479–1491. doi:10.1007/s00521-016-2763-0.
- Xavier, A. C., King, C. W., Scanlon, B. R. 2015. Daily gridded meteorological variables in Brazil (1980-2013). *International Journal of Climatology*, 36(6), 2644-2659. doi: <https://doi.org/10.1002/joc.4518>.

CAPÍTULO 2:

Análise do comportamento de vazões sob a influência de reservatório com regularização de vazões

RESUMO: O conhecimento do gerenciamento/operacionalização dos reservatórios com regularização de vazões formados pelas usinas hidrelétricas é essencial para a avaliação da disponibilidade hídrica e, conseqüentemente, para o planejamento de uso dos recursos hídricos, especialmente nas posições de jusante. A operacionalização dos reservatórios pode ser influenciada tanto pelas características climáticas quanto pela forma com que as atividades humanas se desenvolvem na bacia. Sendo assim, objetivou-se com esse trabalho avaliar as relações existentes entre as vazões afluentes e defluentes à um reservatório, bem como com as estações fluviométricas convencionais localizadas a jusante do barramento. Além disso, buscou-se também avaliar as tendências do comportamento das vazões mínimas, médias e máximas, no período pós-operação, levando em consideração as características da precipitação e da irrigação na região. Os resultados indicaram que a operacionalização do reservatório está fortemente relacionada ao comportamento das vazões afluentes, especialmente ao se considerar análises em períodos trimestrais, afetando também o escoamento nas estações fluviométricas a jusante. Foi verificado ainda redução em todas as variáveis analisadas relacionadas as vazões afluentes e defluentes, bem como das estações a jusante do barramento, a exceção da vazão máxima na estação mais distante ao reservatório, que apresentou comportamento estacionário. As reduções observadas nas vazões podem estar relacionadas ao aumento em quase três vezes da área irrigada por pivô-central na bacia, contudo, o mesmo não se pode dizer em relação regime pluviométrico anual da região, já que este apresentou comportamento estacionário para a grande maioria de estações avaliadas.

Palavras-chave: tendência, operacionalização, barramento, correlação.

ABSTRACT: Understanding the behavior of reservoirs with regularization of flows formed by hydroelectric power plants is essential for the assessment of water availability and, consequently, for planning the use of water resources, especially in the downstream positions. The operationalization of the reservoirs can be influenced both by climatic characteristics and by the consequences resulting from human actions in the basin. Therefore, the objective of this work was to evaluate the existing relationships between the inflow and outflow to a reservoir,

as well as with the conventional fluviometric stations downstream of the dam. In addition, it was also sought to evaluate trends in the behavior of minimum, average and maximum flows, in the post-operation period, taking into account the characteristics of rainfall and irrigation in the region. The results indicated that the operationalization of the reservoir is strongly related to the behavior of the inflows, especially when considering analyzes in quarterly periods, also affecting the flow in the fluviometric stations downstream. It was also verified a reduction in all variables analyzed related to inflows and outflows, as well as in the stations downstream of the dam, with the exception of the maximum flow in the station farthest from the reservoir, which presented a stationary behavior. The reductions observed in the flows may be related to the increase in almost three times of the area irrigated by central pivot in the basin, however, the same cannot be said in relation to the annual rainfall regime of the region, since it presented a stationary behavior for the vast majority. of stations evaluated.

Keywords: tend, operationalization, dam, correlation

2.1. INTRODUÇÃO

No Brasil, a capacidade instalada para a geração de energia elétrica era de 162 GW no final de 2019, dos quais 101,9 GW (62,9%) provenientes de usinas hidroelétricas, 22,9 GW (14,1%) de termoelétricas convencionais e nucleares e 37,3 GW (23,0%) de Pequenas Centrais Hidroelétricas (PCHs), usinas a biomassa, eólica e solares, conforme dados disponibilizados pelo Operador Nacional do Sistema Elétrico (ONS, 2020). Isso significa que a hidroeletricidade continua sendo a principal fonte de geração de energia do país.

As usinas hidroelétricas são compostas por grandes reservatórios de água que, normalmente, objetivam a formação de reservas nos períodos de maior disponibilidade hídrica, as quais podem ser utilizadas para atenuação da escassez nos períodos de estiagem (Setti et al. 2000). Assim, a construção de reservatórios causa alterações nos regimes hidrológicos naturais, armazenando água e liberando-a conforme o padrão temporal exigido pela demanda.

Em termos globais, a compreensão da dinâmica quanto ao armazenamento da água em lagos e reservatórios precisa ser melhorada (Busker et al., 2018). Segundo Nourani et al. (2020), entre os parâmetros hidrológicos e meteorológicos que afetam a quantidade de água armazenada em um reservatório podem-se incluir a precipitação, o volume de água liberado para atender as necessidades a jusante e também, com especial destaque, as vazões afluentes.

Nesse contexto, essas estruturas tornam-se vulneráveis às mudanças climáticas (Schaefli, 2015) e, ainda, a outros fatores como mudanças no uso e ocupação da terra e demanda por irrigação a montante do reservatório (Splading-Fecher et al., 2014), que podem influenciar diretamente na disponibilidade de água para a geração de energia elétrica. A análise do comportamento dessas variáveis ao longo dos anos tem, inclusive, se tornado cada vez mais relevante, devido a possibilidade de identificar tendências que auxiliem na gestão da água dentro de um território ou bacia hidrográfica.

Pereira et al. (2018), analisando a variabilidade da precipitação em todo o território do estado de Minas Gerais, no decorrer dos anos de 1891 a 2017, identificaram tendência de redução da estação chuvosa para a região norte, nenhuma alteração de comportamento para a região central e tendência de aumento da estação chuvosa para a região sul e extremo oeste do estado evidenciando, desta forma, a existência de mudanças do comportamento na distribuição das chuvas no estado.

Outros estudos como o de Abeysingha et al (2016), além de analisarem a tendência da precipitação também avaliaram o comportamento dos fatores relacionados às atividades humanas, objetivando compreender as tendências decrescentes verificadas para as vazões do rio Gomti, na Índia, identificando que o crescimento da irrigação e da população foram os maiores responsáveis pela retirada de água na bacia. Nesse sentido, Silva et al (2020) buscando avaliar o comportamento dos dados de vazão, precipitação e uso do solo, na bacia do Paracatu, localizada no estado de Minas Gerais, no Brasil, também evidenciaram tendências de redução das vazões em todas as estações fluviométricas consideradas, e embora não tenham encontrado alterações de comportamento das precipitações, destacou que as atividades associadas a agricultura irrigada podem estar afetando negativamente a sustentabilidade dos recursos hídricos.

No Brasil, a Agência Nacional de Águas e Saneamento Básico destacou em seu Atlas de Irrigação que o crescimento da irrigação se intensificou a partir das décadas de 1970 e 1980, com aumentos mais expressivos em São Paulo, Minas Gerais, Tocantins, Bahia, Rio Grande do Sul e, mais recentemente, em Goiás. O estudo identifica, ainda, que em 2019 havia 1,556 Mha irrigados por pivôs-centrais (sistema que mais cresce atualmente), sendo essa área 50 vezes maior do que a área mapeada no ano de 1985. Ressaltaram também que os maiores polos estão localizados em Minas Gerais (28,8%) e Goiás (17,3%), embora todos os estados brasileiros apresentaram crescimento em médio e curto prazos (ANA, 2021).

Sendo assim, a ampliação de áreas irrigadas tem levado a um aumento expressivo no uso da água superficial e, também, da extração da água subterrânea, acarretando riscos em

relação a manutenção da disponibilidade hídrica nos cursos d'água. Esse tipo de situação pode contribuir para o deplecionamento da quantidade de água armazenada nos reservatórios, devendo essas análises serem sempre realizadas de forma associada ao comportamento dos fatores relacionados ao clima, como a precipitação. Vale mencionar que, além da irrigação, a existência de pequenos barramentos a montante dos reservatórios também pode influenciar no quantitativo de vazões afluentes disponíveis para reservação.

Um estudo realizado por Shadkam et al. (2016) mostrou que embora as diminuições das vazões afluentes no lago Urmia tenham seguido as reduções observadas na precipitação, a variabilidade das vazões afluentes foi mais pronunciada do que a variabilidade na precipitação, sugerindo que o uso da água para irrigação aumentou a pressão sobre a disponibilidade de água na bacia e provocou uma diminuição das vazões. Além disso, o estudo confirmou que a presença de reservatórios nos afluentes contribuiu positivamente para a disponibilidade de água durante anos relativamente secos, recomendando que para se evitar maiores impactos existe a necessidade de desenvolvimento de planos nacionais que reduzam o uso da irrigação e planos internacionais para enfrentar as mudanças climáticas.

Sendo assim, o estudo das relações entre as vazões afluentes e defluentes de um reservatório pode ser fundamental para determinar e/ou planejar regimes operacionais eficientes, como no trabalho de Chang e Chang (2001), em que foram determinados cronogramas de operações ideais para um reservatório a partir da utilização de séries temporais de vazões afluentes.

Neste contexto, o presente trabalho foi desenvolvido objetivando: 1) avaliar as relações existentes entre as vazões afluentes e defluentes ao reservatório da UHE Queimado; 2) as relações entre as vazões defluentes e as vazões das estações fluviométricas convencionais localizadas a jusante do reservatório da UHE Queimado, destacadamente vazões mínimas, médias e máximas anuais; e 3) avaliar as tendências do comportamento das vazões mínimas, médias e máximas e sua relação com o regime pluviométrico anual e avanço da irrigação com pivô central na região de estudo no período pós-operação do reservatório da UHE Queimado.

2.2. MATERIAL E MÉTODOS

2.2.1. Área de estudo

Para o estudo foi selecionado o reservatório da Usina Hidrelétrica de Queimado (UHE Queimado), localizado na bacia do rio Preto (10.325 km²), sub-bacia do rio Paracatu, inserida,

por sua vez, na bacia hidrográfica do rio São Francisco. A bacia do rio Preto abrange parte do Distrito Federal (13%) e dos estados de Goiás (22%) e Minas Gerais (65%).

A UHE Queimado (Figura 1), de propriedade do Consórcio entre a Companhia Energética de Minas Gerais (CEMIG) e a Companhia Energética de Brasília (CEB), está localizada em Palmital de Minas, distrito pertencente ao município mineiro de Cabeceira Grande.

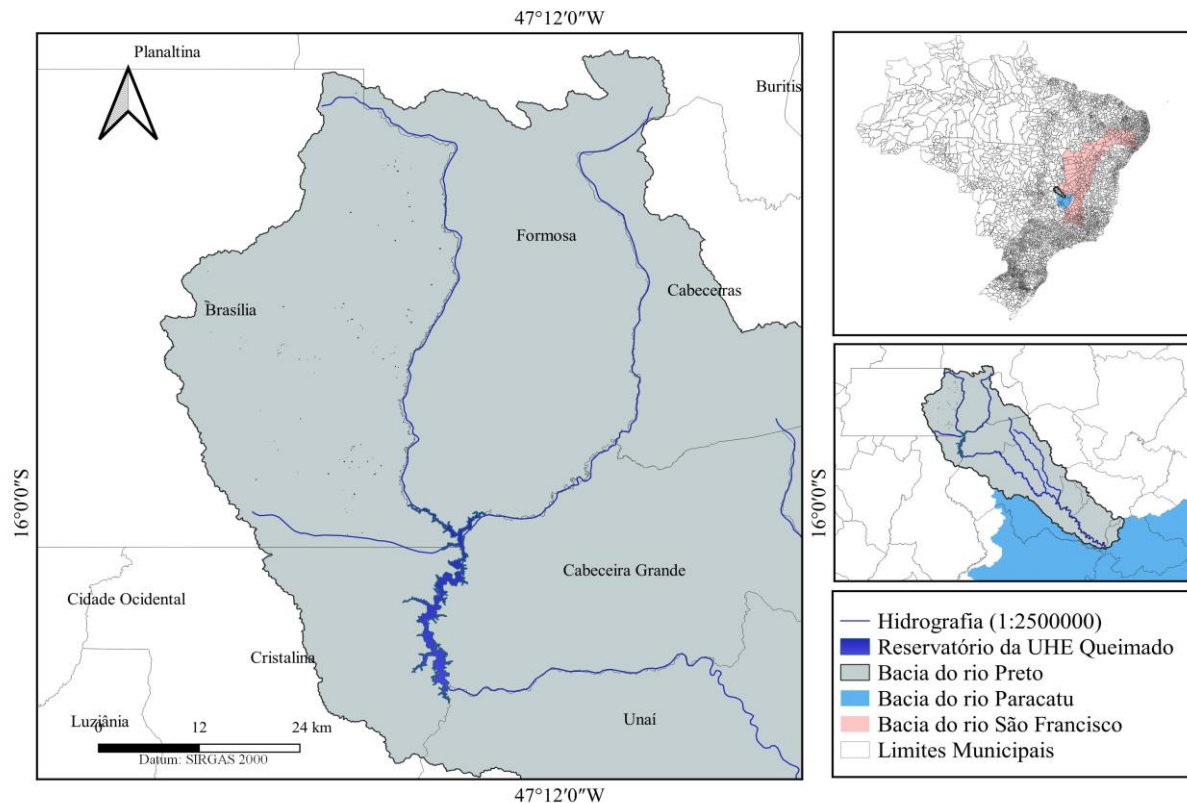


Figura 1 – Localização do reservatório UHE Queimado

As obras para a sua construção foram iniciadas no ano 2000 e concluídas em 2004, quando entrou em plena operação comercial. Com potência instalada de 105 MW, a UHE Queimado gera energia suficiente para atender 300 mil consumidores. A altura máxima da barragem é de 70 metros, com cerca de 36,26 km² de área inundada. Os valores máximo e mínimo de cotas do nível de água no reservatório são 830 e 811 metros, respectivamente.

O reservatório formado pela UHE Queimado possui volume útil de 389,46 hm³ com volume máximo de 477,97 hm³ e volume mínimo de 88,51 hm³, segundo dados do boletim diário disponibilizado em 17/05/2021 pela ANA.

As condições para sua operação devem respeitar os usos múltiplos, conforme estabelece a ANA, através da Resolução nº 147, de 2 de março de 2015, com descarga mínima

de 8,8 m³/s no período úmido (novembro a abril) e 17 m³/s no período de estiagem (maio a outubro). De acordo com o Operador Nacional do Sistema Elétrico (ONS), o reservatório possui restrição de defluência que, ao se combinar com a vazão incremental no trecho Queimado/Unaí, não poderá ultrapassar a vazão de 300 m³/s no município de Unaí, auxiliando assim no controle de cheias (ONS, 2020).

2.2.2. Dados hidrológicos

A partir do inventário disponibilizado pela ANA foram identificadas 76 estações fluviométricas na bacia hidrográfica do rio Preto. A seleção das estações a serem utilizadas no presente estudo foi baseada nos seguintes critérios: i) estar localizada na bacia do rio Preto; ii) a jusante do reservatório da UHE Queimado; e iii) possuir, no mínimo, 10 anos de dados, considerando o período base de 2005 a 2019, posterior a construção e estabilização das atividades operacionais da UHE Queimando.

Utilizando os critérios supracitados, verificou-se que, das 76 estações existentes na bacia do rio Preto, apenas 22 estações encontram-se à jusante do reservatório UHE Queimado. No entanto, ao se verificar a disponibilidade de dados observados para essas estações, apenas 3 delas atenderam o critério iii, as quais foram as selecionadas para o estudo: Fazenda Limeira (42460000), Unaí (42490000) e Porto dos Poções (42600000), denominadas na Figura 2 como Q1, Q2 e Q3, respectivamente. As séries históricas dessas estações foram obtidas na plataforma Hidroweb (<https://www.snirh.gov.br/hidroweb/serieshistoricas>), que inclui dados da Rede Hidrometeorológica Nacional (RHN), de responsabilidade da ANA.

Os registros históricos de vazões afluente (Qa) e defluente (Qd), associados à operacionalização do reservatório da UHE Queimado, também foram obtidos na ANA, através da plataforma do Sistema de Acompanhamento de Reservatórios – SAR (<https://www.ana.gov.br/sar/>), que fornece, entre outras informações, as vazões diárias afluentes do período de 1993 a 2019 e defluentes do período de 2003 a 2019.

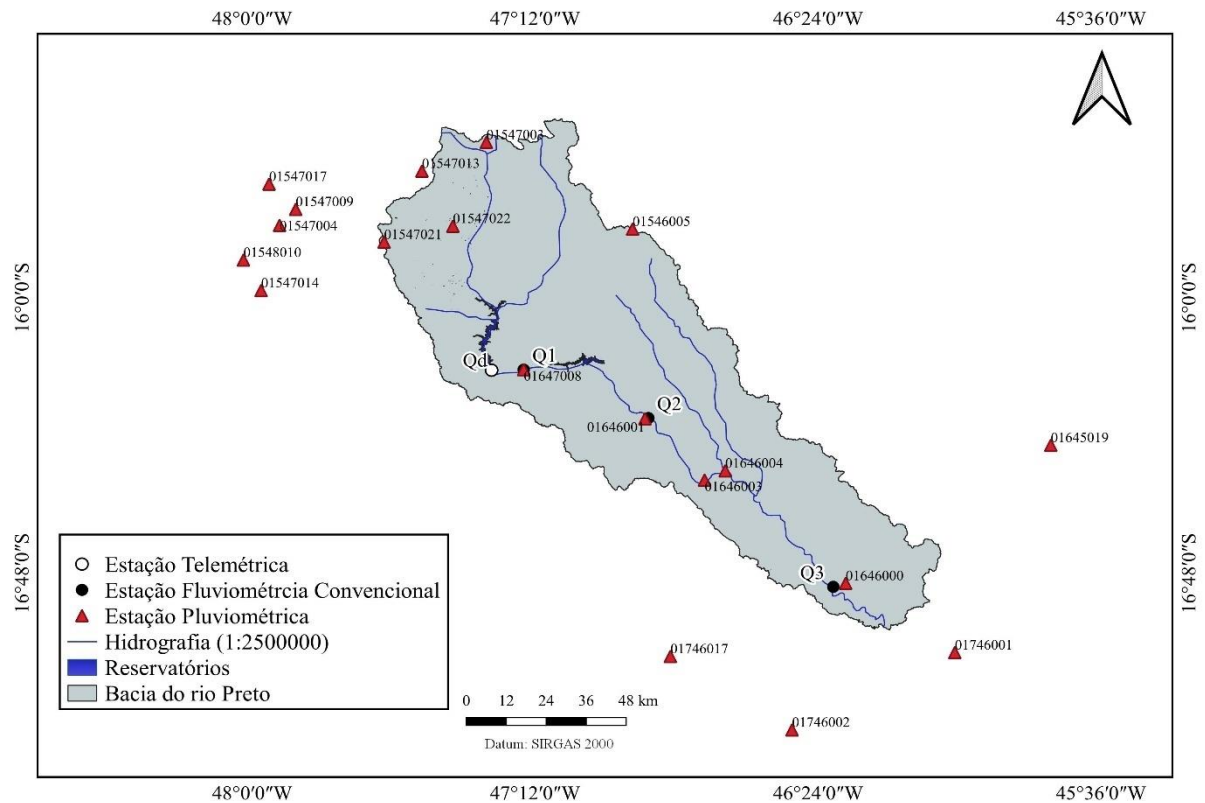


Figura 2 – Localização das estações fluviométricas e pluviométricas da rede hidrometeorológica da ANA selecionadas para o estudo na bacia hidrográfica do rio Preto e da estação telemétrica de vazões defluentes do reservatório UHE Queimado.

Segundo informações coletadas junto à CEMIG, as vazões defluentes são calculadas a partir de informações de rendimento do conjunto turbina-gerador, função de queda bruta ou líquida pela potência e da curva de descarga do vertedouro, obtida por modelos de escala reduzida ou modelagem matemática. Para as vazões defluentes há um ponto de controle para referência, considerado pela ANA e denominado de Estação Telemétrica UHE Queimado Barramento (código 42459080), de onde se obtém a vazão média diária a partir do dados horários de vazão defluente (<http://www.snirh.gov.br/hidrotelemetria/Mapa.aspx>).

Já as vazões afluentes são obtidas pela equação do balanço hídrico, considerando a variação do volume estocado no tempo e o volume que saiu do reservatório nesse mesmo intervalo de tempo. Dessa forma, as estimativas de afluências não são obtidas a partir de uma estação ou ponto de controle específico e, por isso, não existe na Figura 2 a localização das medições de Q_a .

As principais informações para a estação fluviométrica telemétrica que coleta os dados de Q_d e para as outras três estações fluviométricas convencionais (Q1, Q2 e Q3) utilizadas no estudo estão apresentadas na Tabela 1.

Tabela 1 – Características das estações fluviométricas convencionais e da estação fluviométrica telemétrica localizada na UHE Queimado

Código	Nome	Estações	Latitude	Longitude	Área (km ²)	Tipo
42460000	Fazenda Limeira	Q1	-16,2089	-47,2325	3.890	Convencional
42490000	Unai	Q2	-16,3494	-46,8800	5.360	Convencional
42600000	Porto dos Poções	Q3	-16,8397	-46,3572	9.400	Convencional
42459080	UHE Queimado Barramento	Qd	-16,2106	-47,3233	3.657	Telemétrica

Para verificar o comportamento pluviométrico na região, no período de funcionamento (operação) do reservatório, foram selecionadas 19 estações pluviométricas (Figura 2) localizadas dentro da bacia do rio Preto e no seu entorno (buffer de 50 km), com disponibilidade mínima de 10 anos de dados dentro do período base de 2005 a 2019, buscando-se avaliar a existência de possíveis tendências do regime pluviométrico da região de interesse.

Os dados históricos de precipitação também foram obtidos na plataforma Hidroweb da ANA, sendo selecionadas apenas aquelas estações que satisfizeram aos mesmos critérios estabelecidos em iii para as estações fluviométricas. Foram, portanto, utilizadas informações relativas a 19 estações pluviométricas no total, sendo 8 delas localizadas dentro da bacia do rio Preto (Figura 2).

Após a sistematização dos dados, foram obtidos para cada uma das séries históricas de vazão (Qa, Qd, Q1, Q2 e Q3) os valores anuais de vazão média ($Q_{méd}$), vazão mínima média de 7 dias de duração (Q_7) e vazão máxima diária ($Q_{máx}$), tendo como base os anos hidrológicos de 2005/2006 a 2018/2019. Com base na análise dos dados pluviométricos e fluviométricos das estações localizadas na bacia do rio Preto, o ano hidrológico para a região de estudo tem início em novembro e término em outubro.

Foram estimados, ainda, os valores de $Q_{méd}$ mensais para serem avaliados sob duas condições: na escala anual, considerando os valores de vazão média obtidos para cada mês, de cada ano, dentro do período considerado, e trimestral, considerando a vazão média obtida para cada mês dentro do trimestre selecionado. O estabelecimento dos meses para cada trimestre foi baseado na definição do ano hidrológico da região, sendo que o primeiro (1º) trimestre foi composto pelos meses de novembro, dezembro e janeiro; o segundo (2º) foi composto pelos meses de fevereiro, março e abril; o terceiro (3º) composto pelos meses de maio, junho e julho e o quarto (4º) composto pelos meses de agosto, setembro e outubro, obedecendo a definição do ano hidrológico da região, buscando assim avaliar a representatividade dos dados em função da sazonalidade da região.

Ressalta-se que os dados de vazões diárias, provenientes das estações convencionais da ANA (Q1, Q2 e Q3), utilizadas para as estimativas de $Q_{\text{méd}}$, Q_7 e $Q_{\text{máx}}$, foram submetidos a um processo de tratamento/consistência de dados e posterior preenchimento de falhas aplicando modelagem com o uso de algoritmos de aprendizado de máquinas. As séries de vazões Q_a e Q_d não apresentaram falhas dentro do período em análise.

Para as séries históricas de precipitação foram obtidos, para o mesmo período-base das vazões, os valores de precipitação total anual para cada estação pluviométrica selecionada no estudo.

Dadas as condições para as quais este estudo foi realizado, buscou-se verificar as variabilidades hidrológicas nas séries de vazões decorrentes da influência da operação do reservatório formado pela UHE Queimado, cuja operação foi iniciada em 2004. Em função disso, adotou-se o período-base para as análises de 2005/2006 a 2018/2019, contemplando uma série histórica com 14 anos de registros, de modo a evitar que fossem avaliados de forma conjunta períodos anteriores e posteriores à construção do reservatório de regularização de vazões.

Apesar da Organização Meteorológica Mundial (OMM) indicar que o clima de uma região deve ser caracterizado com base no período mínimo de 30 anos (OMM, 1975), o uso de série longas pode não ser representativo quando há influência de efeitos antrópicos nos dados, decorrentes, por exemplo, da construção de reservatórios (Tucci, 2003), demonstrando assim que o período-base adotado nesse estudo, para análises estatísticas de tendência, é coerente com a situação apresentada.

2.2.3. Análise exploratória da relação entre as vazões

Foi realizada a análise exploratória do comportamento das vazões sob influência do reservatório UHE Queimado utilizando, para isso, a análise de correlação de Pearson, Equação 1, (Wu et al., 2018). Para isto, foi considerado duas escalas temporais (anual e trimestral) para avaliar o grau de associação entre as séries de vazões médias afluentes, defluentes, Q1, Q2 e Q3, dentro do período-base considerado.

A correlação de Pearson (r) é a medida de correlação mais comumente usada e às vezes chamada de coeficiente de correlação linear porque mede a associação linear entre duas variáveis. Se os dados estiverem sobre uma linha reta com uma inclinação positiva, então $r = 1$, caso a linha tenha inclinação negativa, então $r = -1$ (Helsel et al., 2020).

$$r_{xy} = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^n (y_i - \bar{y})^2}} \quad \text{Equação 1}$$

Em que, r é a correlação de Pearson; n é o número de observações x_i e y_i (tamanho da amostra), x_i e y_i são as vazões observadas e \bar{x} e \bar{y} representam as médias das vazões observadas a serem consideradas.

Assim, foi avaliada a correlação das vazões anuais defluentes (variável dependente) com as vazões anuais afluentes (variável independente) e, posteriormente, a correlação entre as vazões defluentes, agora considerada como variável independente, e as vazões médias (escalas anuais e trimestrais) das estações fluviométricas convencionais localizadas a jusante do barramento (Q1, Q2 e Q3), tomadas como variáveis dependentes.

A classificação da correlação entre as variáveis foi definida com base na proposição apresentada por Bozzoni et al. (2020), segundo a qual a correlação é moderada para valores de r acima de 0,5; forte para valores acima de 0,7 e muito forte para valores acima de 0,9.

2.2.4. Análises de tendência

2.2.4.1. Método de Mann-Kendall e Mann-Kendall Modificado

Conforme recomendado por OMM (1988), para a análise das séries temporais de vazão e precipitação utilizadas neste trabalho foi adotado o teste não paramétrico proposto por Mann (1945) e Kendall (1975). O teste foi utilizado por não requerer distribuição normal dos dados (Yue et al, 2002), por ser um método robusto (Helsel e Hitch, 2002) e, ainda por lidar bem com falhas nas séries com dados abaixo do limite de detecção.

O teste de Mann-Kendall consiste em comparar cada valor da série com os demais, em uma ordem sequencial, contando a quantidade de vezes que os demais valores são maiores que o dado observado (Back, 2001). A estatística do teste é baseada na Equação 2.

$$S = \sum_{i=2}^n \sum_{j=1}^{i-1} \text{signal}(x_i - x_j) \quad \text{Equação 2}$$

Em que, x_i e x_j são os valores dos dados em sequência e n é o tamanho da série de dados, e:

$$\text{senal}(x_i - x_j) = \begin{cases} +1 & \text{se } (x_i - x_j) > 0 \\ 0 & \text{se } (x_i - x_j) = 0 \\ -1 & \text{se } (x_i - x_j) < 0 \end{cases} \quad \text{Equação 3}$$

Conforme explicam Chauluka et al. (2021), quando $n > 10$ assume-se que a estatística de teste S é normalmente distribuída com:

$$E[S] = 0 \quad \text{Equação 4}$$

$$V(S) = \frac{n(n+1)(2n+5) - \sum_{i=1}^q t_p(t_p-1)(2t_p+5)}{18} \quad \text{Equação 5}$$

Em que, $E[S]$ representa a média do sinal, $V(S)$ a variância do sinal, t_p representa o número de dados com valores iguais em certo grupo (o p -ésimo), q é o número de grupos iguais na série e n é o tamanho da série.

Com base nos valores de S foi verificado se a tendência, quando existente, foi crescente (valores positivos) ou decrescente (valores negativos). Outro parâmetro estatístico necessário para aplicação do teste de Mann Kendall é o valor de Z , determinado a partir da Equação 6.

$$Z_{mk} = \begin{cases} \frac{S-1}{\sqrt{\text{Var}[S]}} & \text{se } S > 0 \\ 0 & \text{se } S = 0 \\ \frac{S+1}{\sqrt{\text{Var}[S]}} & \text{se } S < 0 \end{cases} \quad \text{Equação 6}$$

O teste avalia as seguintes hipóteses: H_0 - os dados são independentes e igualmente distribuídos (não há tendência) e H_1 - os dados possuem tendência monotônica no tempo (há tendência). Em um teste bilateral para tendência, H_0 deve ser aceito para um dado nível de significância α , se para o quantil $Z_{\alpha/2}$ de uma distribuição normal padrão $|Z| \leq Z_{\alpha/2}$.

O estabelecimento do nível de significância divide o domínio da variável de teste em duas regiões. Uma destas regiões é constituída de valores da variável teste que levam à aceitação de H_0 . Esta região é denominada região de aceitação. A outra região é constituída do conjunto de valores da variável de teste que levam à rejeição de H_0 , denominada região crítica. O nível

de significância adotado neste trabalho foi de 5%, estando associado a um valor de $Z_{\alpha/2} = Z_{0,025} = 1,96$.

Para a aplicação do teste de Mann Kendall é necessário que os dados sejam aleatórios e independentes (Neeti e Eastman, 2011) e, em função disso, essas condições foram analisadas previamente através da aplicação do teste de sequência (Run Test) e do teste de autocorrelação serial, respectivamente (Salviano et al., 2015). Este último, compara os valores da série, em um determinado período de tempo, e os valores da mesma série em períodos defasados de tempo. Para essas defasagens, denominadas *lags*, os resultados são avaliados através de correlogramas. Neste estudo considerou-se autocorrelação significativa avaliando os dados encontrados para a *lag-1*, conforme recomendado por Abeysingha et al. (2016).

Quando as séries apresentaram características de aleatoriedade e independência entre si foi aplicado o teste de Mann-Kendall. Entretanto, para aquelas séries que apresentaram comportamento não aleatório e dependentes, foi realizada a aplicação do teste de Mann-Kendall Modificado, uma vez que é necessário um ajuste no método para que a correlação seja levada em consideração (Hamed e Rao, 1998).

O ajuste mencionado é feito através de uma correção no cálculo da variância, conforme apresentado na Equação 7.

$$V^*(S) = V(S) \frac{n}{n_S^*} \quad \text{Equação 7}$$

onde,

$$\frac{n}{n_S^*} = 1 + \frac{2}{n(n-1)(n-2)} \times \sum_{k=1}^{n-1} (n-k)(n-k-1)(n-k-2)p_k \quad \text{Equação 8}$$

Em que, n é o número atual de observações, n_S^* é considerado como o número efetivo de observações para explicar a autocorrelação nos dados e p_k é o coeficiente de autocorrelação estimado para o *lag-k* a partir das séries temporais.

As análises para verificação de aleatoriedade e independência entre as séries, bem como a aplicação dos testes de tendências de Mann-Kendall ou, quando necessário, de Mann-Kendall modificado, foram realizadas utilizando o software R.

2.2.4.2. Método de Sen Slope

Em decorrência do teste de Mann-Kendall não fornecer a magnitude das tendências, quando detectadas, aplicou-se o método de Sen Slope. Este método foi proposto por Sen (1968) e aprimorado por Hirsch et al. (1982), e pode prover uma medida realística das tendências, quando verificadas, em uma série histórica (Ferrari, 2012).

O Sen Slope é amplamente utilizado em estudos quanto ao comportamento temporal de séries hidrometeorológicas (Silva et al., 2015; Dubey et al. 2020) e é estimado através da aplicação da estatística Q_{ij} , que pode ser calculada conforme a Equação 9.

$$Q_{ij} = \frac{x_j - x_i}{j - i}, \text{ com } i < j \quad \text{Equação 9}$$

Em que, x_j, x_i são os valores das variáveis em estudo nos anos j e i , respectivamente. Se houver n valores na série analisada, então o número de pares estimados de Q é dado por $N < n(n - 1)/2$. Os n valores de Q_{ij} são classificados do menor para o maior e a mediana da estimativa de Sen Slope é calculada como mostra a Equação 10.

$$Q_m = \left\{ \begin{array}{l} Q_{[(n+1)/2]}, \text{ se } n \text{ for impar} \\ \frac{Q_{[n/2]} + Q_{[(n+2)/2]}}{2}, \text{ se } n \text{ for par} \end{array} \right\} \quad \text{Equação 10}$$

O sinal de Q_m reflete a tendência (aumento ou redução) dos dados, enquanto o seu valor Q_m indica a magnitude da tendência. Para determinar se a inclinação mediana é estatisticamente diferente de zero obtém-se o intervalo de confiança de Q_m na probabilidade específica.

O intervalo de confiança de Sen Slope pode ser calculado aplicando-se a Equação 11.

$$C_\alpha = Z_{1-\alpha/2} \sqrt{V(S)} \quad \text{Equação 11}$$

Em que, $V(S)$ é calculado conforme a Equação 5 ou 7, variando conforme a avaliação preliminar dos dados em relação a aplicabilidade dos métodos de Mann-Kendall ou Mann-Kendall Modificado, respectivamente; $Z_{1-\alpha/2}$ é obtido a partir da tabela de distribuição normal padrão, sendo que neste estudo foi calculado para o nível de significância de 5% ($\alpha = 0,05$).

Então, $M_1 = (n - C_\alpha)/2$ e $M_2 = (n + C_\alpha)/2$ são calculados, representando os limites inferior e superior do intervalo de confiança.

2.3. RESULTADOS E DISCUSSÃO

2.3.1. Análise exploratória das vazões afluentes e defluentes

Na Figura 3 é apresentado a variação das vazões afluentes (Figura 3A) e defluentes (Figura 3B) diárias, considerando cada um dos anos de estudo adotados nesse trabalho (2005 a 2019), os quais são representados pelos pontos azuis. Dessa forma, para um mesmo dia de um determinado mês, tem-se representado 14 valores de vazão.

Já os pontos vermelhos representam os valores médios diários obtidos entre os anos, isto é, para um determinado dia do mês tem-se a representação de um único valor de vazão, calculado a partir da média dos 14 anos analisados.

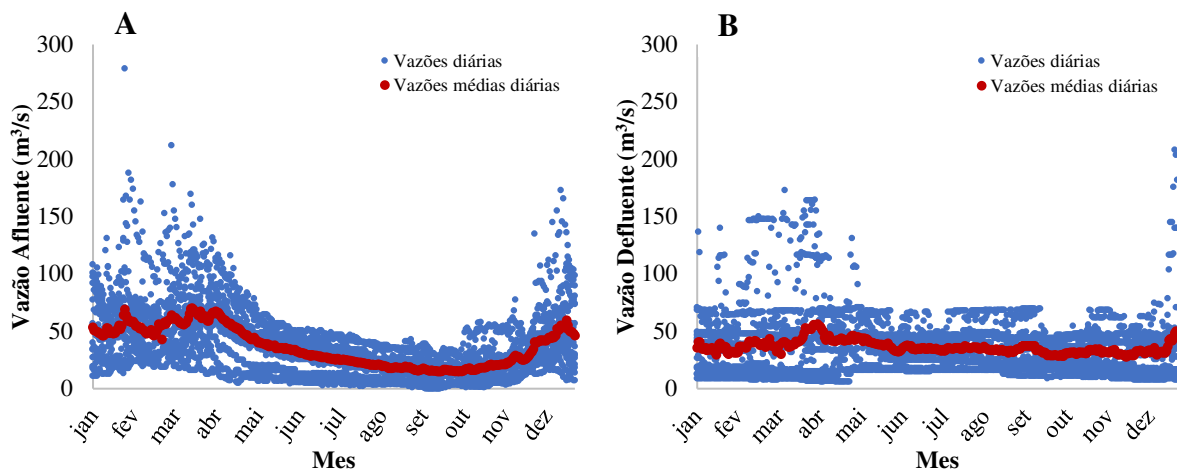


Figura 3 – Vazões afluentes (A) e defluentes (B) diárias para o reservatório UHE Queimado, entre o período de 2005 a 2019.

É possível notar na Figura 3A que as vazões afluentes possuem um comportamento sazonal evidente, com vazões maiores entre os meses de novembro e abril, representativo do período chuvoso da região, e vazões menores entre maio e outubro, meses caracterizados pela reduzida incidência de chuvas.

Fica claro, ainda, que as vazões afluentes, especialmente para o período chuvoso, apresentam elevada amplitude, como no mês de fevereiro, em que as vazões diárias variaram de 13,6 a 280 m³/s ao longo da série histórica considerada. Já para o período seco, a amplitude da variabilidade é menor, como, por exemplo, no mês de julho, em que as vazões variaram entre 3,5 e 51 m³/s.

Já na Figura 3B, que representa o comportamento das vazões defluentes, percebe-se que em alguns poucos momentos, do período chuvoso, especialmente no mês de dezembro, a

magnitude das vazões superou o próprio valor das vazões afluentes, o que normalmente não é esperado. Esses valores, porém, foram registrados entre 2005 e 2006, período marcado pelo início das atividades do reservatório, o que pode significar que ajustes ainda estavam sendo realizados para que a sua correta operacionalização.

De maneira geral, para as vazões defluentes, o comportamento do escoamento apresenta uma maior constância ao longo de todo o ano, evidenciando a regularização de vazões proporcionada pelo reservatório. Apesar disso, ainda se observa diferenças de vazões defluentes para diferentes anos, considerando o mesmo dia, embora em proporções bem menores que as vazões afluentes.

A Figura 4 representa o diagrama de massas, conforme proposto por Rippl (1883), construído a partir das séries históricas de vazões afluentes e defluentes ao reservatório, de maneira que se pode observar alterações menos abruptas das declividades do diagrama para as vazões defluentes, comparativamente com as alterações observadas para as vazões afluentes.

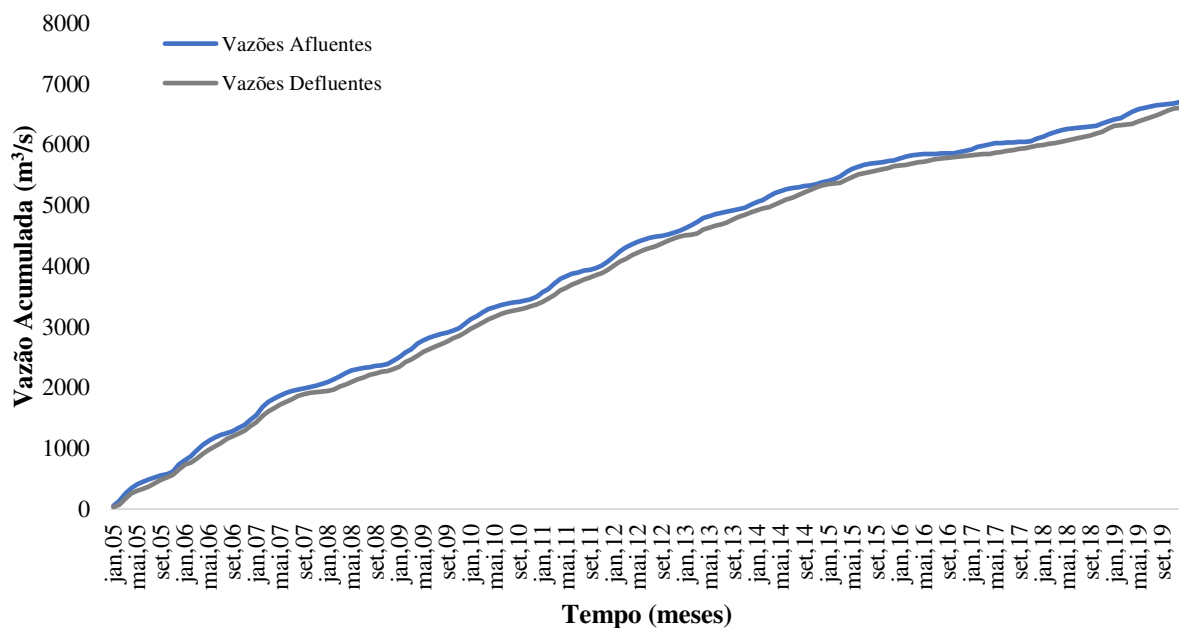


Figura 4 – Diagrama de massas das vazões afluentes e defluentes ao reservatório UHE Queimado.

A análise do comportamento das vazões evidencia que a vazão afluente apresenta maior variabilidade em função da sazonalidade do regime de precipitações, motivo pelo qual ocorre crescimento mais acentuado do acumulado das vazões em alguns períodos do ano (novembro a abril). Em contrapartida, a vazão defluente tem comportamento mais regularizado ao longo de todos os anos dentro do período, influenciado justamente pela operação do

reservatório UHE Queimado, que equilibra o fluxo das vazões ao longo da hidrografia, minimizando eventos de seca e de cheia ao longo dos períodos seco e chuvoso, respectivamente.

Os resultados encontrados da análise de correlação entre as vazões afluentes e defluentes estão apresentados na Figura 5, na escala anual e trimestral.

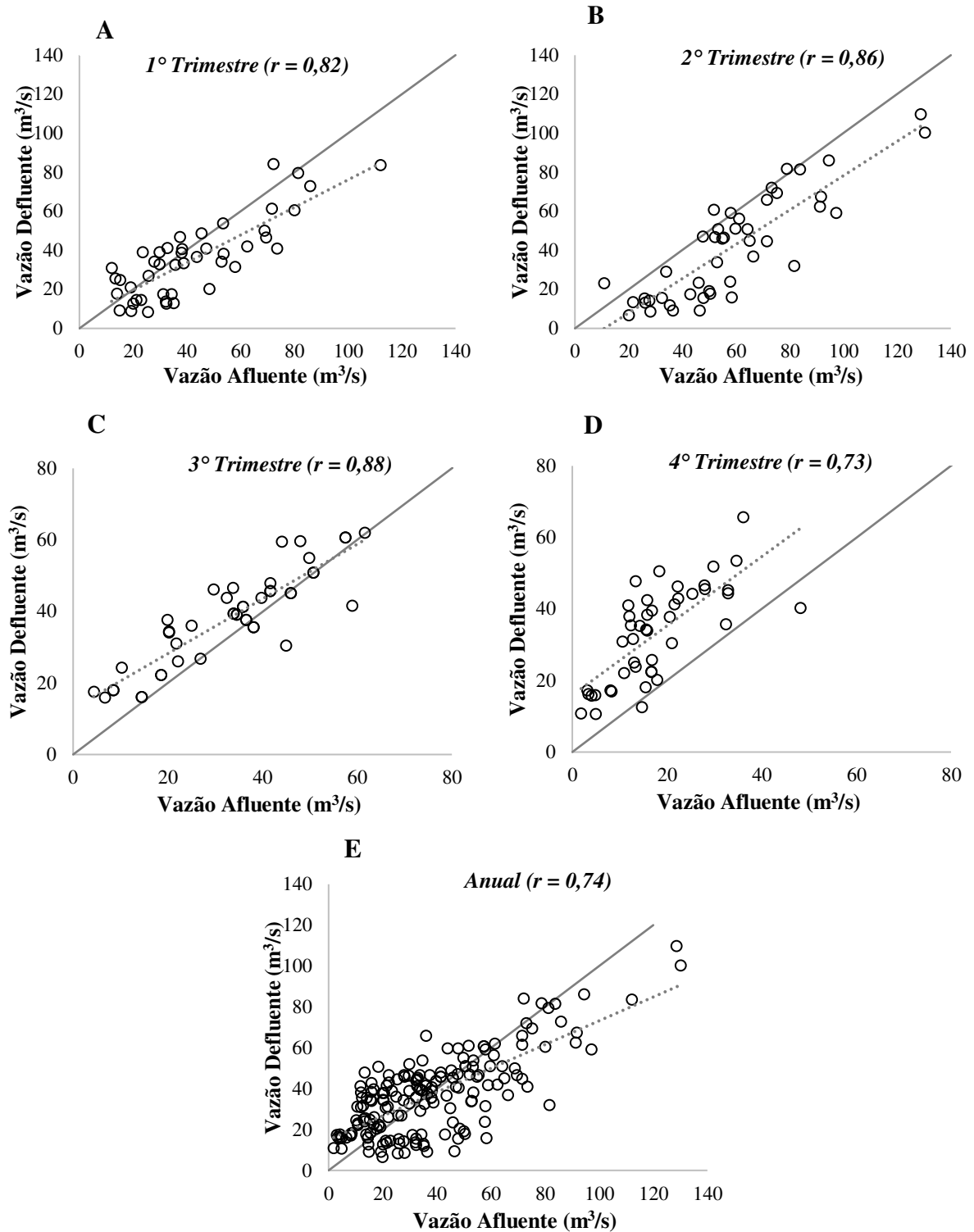


Figura 5 – Correlações obtidas entre as vazões defluentes e afluentes ao reservatório UHE Queimado nas escalas trimestral (A, B, C e D) e anual (E), ao longo do período de 2005 a 2019.

A correlação entre as vazões defluentes e afluentes considerando todos os meses do ano, ao longo do período considerado, foi igual a 0,74, classificada por Bozzoni (2020) como uma correlação forte. Verifica-se, portanto, um resultado satisfatório, demonstrando que os valores das vazões afluentes são determinantes para a operacionalização do reservatório.

Quando se analisa os dados em escala trimestral obtém-se correlações ainda mais elevadas, especialmente para o primeiro (0,82), segundo (0,86) e terceiro (0,88) trimestres, destacando assim a maior dependência das vazões defluentes quanto aos padrões sazonais característicos das vazões afluentes, interferindo assim nas condições de operação do reservatório UHE Queimado.

Para o quarto trimestre, o valor de correlação encontrado foi de 0,73, se aproximando bastante do estimado para a correlação anual. A redução dessa correlação provavelmente ocorre em função da maior atividade de regularização do reservatório nesse período, que deve atender as demandas a jusante mesmo que as vazões afluentes apresentem reduções significativas, ainda que o armazenamento de água nessa estrutura possa estar caracterizado para altos níveis.

Análises dessa natureza são corroboradas em estudos como o de Passaia e Paiva (2019), que avaliaram quais variáveis (vazões afluentes, cota, volume e dia do ano) regem o comportamento da vazão defluente de 150 reservatórios localizados no Brasil e controlados pelo ONS. Segundo os autores, a variável que melhor explicou o comportamento da vazão defluente foi a vazão afluente, sendo encontrado um valor de correlação médio igual a 0,79, o que demonstra a importância explicativa que a vazão afluente representa na determinação do processo de operacionalização dos reservatórios.

Wu et al. (2018) também utilizaram o coeficiente de correlação para quantificar os impactos de um grande reservatório nas relações entre seca hidrológica (HD), que se refere aos níveis dos reservatórios abaixo do normal, e seca meteorológica (MD), associada a precipitação abaixo das normais esperadas. Os autores também obtiveram resultados consideráveis ao utilizar o coeficiente de correlação, de modo que ficou evidenciado que a operação do reservatório também pode estar associada a outras características, como, no caso do estudo, ao HD (reduzindo a duração e amenizando a magnitude da seca) em escalas de curto prazo (intra-anual), e ao MD (aumentando o tempo de resposta das variáveis hidrológicas associadas a precipitação, durante o período pós-reservatório).

Sendo assim, os resultados deste estudo evidenciam que o entendimento quanto ao comportamento das vazões afluentes se torna essencial para o gerenciamento e o planejamento operacional aplicado ao reservatório UHE Queimado. A existência de variabilidades climáticas e/ou antrópicas pode causar interferência no que diz respeito à disponibilidade e

armazenamento de água no reservatório, impactando tanto a geração de energia elétrica quanto as demais atividades implementadas na bacia hidrográfica, a jusante do reservatório, que são usuárias da água.

Essa preocupação fica mais evidente quando se observa, na Figura 4, uma mudança de declividade para as duas linhas representativas do comportamento das vazões afluentes e defluentes, ao final do período estudado, especialmente a partir do ano de 2016, o que significa que o acumulado das vazões ao longo dos anos, nessas posições da hidrografia, está reduzindo e refletindo em menor disponibilidade hídrica na bacia.

Na Tabela 2 são apresentados os resultados da análise de correlação entre as vazões médias defluentes e as vazões médias obtidas a partir das estações fluviométricas convencionais Q1, Q2 e Q3.

Tabela 2. Correlações obtidas entre as vazões defluentes do reservatório UHE Queimado com as estações fluviométricas convencionais Q1, Q2 e Q3 nas escalas anual e trimestral, ao longo do período do estudo.

Período	Coeficiente de Correlação (r)		
	Qd e Q1	Qd e Q2	Qd e Q3
Anual	0,97	0,90	0,62
1º Trimestre	0,92	0,90	0,67
2º Trimestre	0,99	0,96	0,74
3º Trimestre	0,98	0,91	0,88
4º Trimestre	0,99	0,98	0,84

Os resultados demonstram que as correlações entre as vazões defluentes e as vazões na estação Q1, para todos os períodos analisados, foram caracterizadas como muito fortes, indicando um elevado grau de associação entre os dados de vazão da estação fluviométrica convencional a jusante e as vazões defluentes.

Embora todas as correlações encontradas tenham sido muito fortes ($r > 0,90$), os melhores coeficientes foram observados no segundo ($r = 0,99$), terceiro ($r = 0,98$) e quarto ($r = 0,99$) trimestres. Entretanto, o pior coeficiente foi observado para o primeiro ($r = 0,92$) trimestre, mesmo quando comparado com o período anual ($r = 0,97$).

As correlações encontradas entre as vazões defluentes e da estação fluviométrica convencional Q2 apresentaram, como esperado, valores pouco menores, em relação às correlações encontradas para Q1, embora em todos os períodos também tenham sido obtidas relações muito fortes, indicando que os valores médios de vazão defluentes também apresentaram um elevado grau de associação com os dados de vazão da estação fluviométrica convencional a jusante. Para essa situação, os melhores ajustes encontrados foram obtidos para

o segundo ($r = 0,96$) e quarto ($r = 0,98$) trimestres e os menores para o primeiro ($r = 0,90$) e terceiro ($r = 0,91$) trimestres, estando próximo ou igual ao encontrado para o período anual ($r = 0,90$).

Essas análises reforçam a seleção do período-base para a realização dos estudos (2005 a 2019) e a indicação de que os dados de Q1 e Q2 anteriores à entrada em operação do reservatório da UHE Queimado não seriam adequadamente utilizados, já que a não consideração das alterações realizadas no regime de escoamento em função das interferências causadas pela construção do barramento pode conduzir a estimativas de disponibilidade hídrica de baixa qualidade, especialmente pela proximidade entre os pontos analisados.

Já as correlações encontradas entre as vazões defluentes e da estação fluviométrica convencional Q3 foram expressivamente menores quando comparadas com Q1 e Q2, particularmente quando se considerou o período anual ($r = 0,62$), sendo uma correlação moderada conforme a classificação de Bozzoni (2020). O primeiro trimestre também apresentou a mesma classificação, com correlação de 0,67.

A redução observada entre o grau de associação das vazões defluentes e as vazões em Q3 demonstra a perda de influência da operação do reservatório da UHE Queimado no regime de variação das vazões médias nessa estação, justificada em função do aumento da distância entre os pontos analisados. Foi possível observar ainda que as maiores reduções na correlação com Qd se deu para o primeiro ($r = 0,67$) e para o segundo ($r = 0,74$) trimestres, períodos estes caracterizados pelo aumento do regime de chuvas na região, indicando maiores influências das contribuições provenientes das áreas de drenagens a jusante do reservatório, do que efetivamente da operação do reservatório.

Entretanto, para o período seco, em que a contribuição do escoamento superficial se reduz bastante, os valores das correlações para o terceiro ($r = 0,88$) e quarto ($r = 0,84$) trimestres, entre as vazões médias defluentes e Q3, aumentaram, embora, não se possa dizer que a relação entre essas variáveis seja tão forte quanto as encontradas para as outras estações.

Sendo assim, com base na queda evidenciada para as correlações entre Qd e Q3, avaliou-se também a correlação entre as vazões médias afluentes e Q3 (Tabela 3). Nessa avaliação fica evidente que as correlações observadas entre as vazões de Q3 e afluente ao reservatório são expressivamente melhores (forte ou muito forte) do que aquelas observadas com as vazões defluentes.

As correlações obtidas, portanto, podem ser classificadas como fortes para todos os períodos ($r \geq 0,80$), excetuando o 3º trimestre ($r = 0,94$), cuja classificação é ainda melhor (muito forte).

Tabela 3. Correlações obtidas entre as vazões afluentes ao reservatório UHE Queimado com a estação fluviométrica convencional Q3, nas escalas anual e trimestral, ao longo do período do estudo.

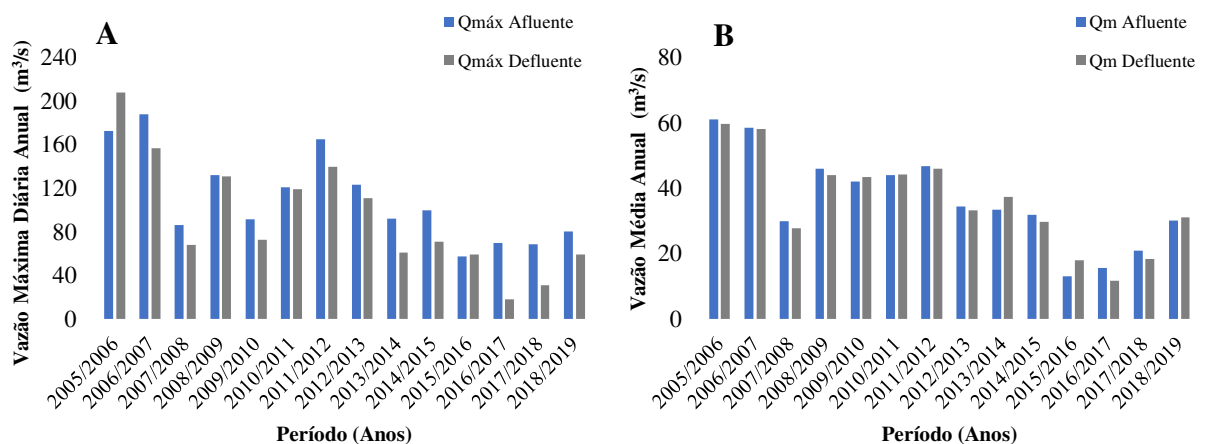
Período	Coefficiente de Correlação (r)
Anual	0,85
1º Trimestre	0,80
2º Trimestre	0,88
3º Trimestre	0,94
4º Trimestre	0,84

Portanto, esses resultados indicam que a variabilidade das vazões nessa posição da hidrografia é influenciada muito mais pelas características hidrológicas e de uso e ocupação da bacia do que pela própria atividade do reservatório.

2.3.2. Análises de tendência

A Figura 6 apresenta os valores da vazão mínima anual com sete dias de duração (Q_7), média anual (Q_m) e máxima anual ($Q_{m\acute{a}x}$) obtidas ao longo dos anos hidrológicos de 2005/2006 a 2018/2019, para as vazões afluentes e defluentes ao reservatório UHE Queimado.

Analisando a Figura 6 observa-se, de maneira geral, a magnitude das vazões máximas afluentes é maior que as máximas defluentes. Por outro lado, as vazões mínimas afluentes apresentam magnitude inferior as mínimas defluentes, o que é justificado pela atividade regulatória do reservatório.



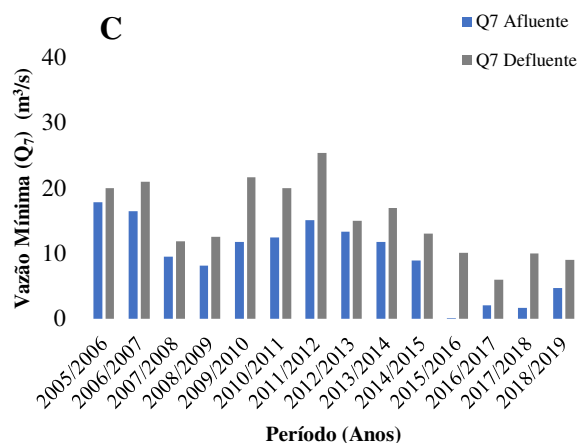


Figura 6 – Valores de vazão máxima diária anual ($Q_{m\acute{a}x}$) (A), média anual (Q_m) (B) e mínima com sete dias de duração (Q_7) (C), afluentes e defluentes do reservatório UHE Queimado, no período de 2005/2006 a 2018/2019.

É possível verificar, ainda, que as vazões, especialmente para os últimos anos, tanto afluentes quanto defluentes, apresentam valores menores para todas as variáveis quando comparados aos anos iniciais da série. Esse comportamento pôde ser confirmado após a aplicação dos testes de tendência a cada uma das séries, em que foi verificado tendência de decréscimo em todas as vazões (Tabela 4).

Tabela 4 – Resultados dos testes de tendência de Mann-Kendall (MK)/Mann-Kendall Modificado (MKM) e Sen Slope para as vazões Q_7 , Q_m e $Q_{m\acute{a}x}$ afluentes e defluentes do reservatório UHE Queimado

Séries Históricas	Variáveis	Mann-Kendall / Mann-Kendall Modificado (5%)				Sen Slope (5%) ($m^3/s/ano$)		
		p-valor	Z-valor	Tau de Kendall	Tendência*	Slope	IC Superior	IC Inferior
Afluentes	Q_7	0,001	-3,235	-0,516	↓	-1,009	-1,569	-0,431
	Q_m	0,002	3,062	-0,626	↓	-2,993	-4,249	-1,7738
	$Q_{m\acute{a}x}$	0,012	2,518	-0,517	↓	-7,125	-11,533	-1,866
Defluentes	Q_7	0,015	2,442	-0,508	↓	-1,321	-2,314	-0,333
	Q_m	0,001	3,389	-0,692	↓	-2,976	-4,132	-1,837
	$Q_{m\acute{a}x}$	0,003	3,058	-0,604	↓	-10,286	-14,200	-3,400

* (↓) tendência de redução; (↑) tendência de elevação e (-) sem tendência.

Os resultados apresentados na Tabela 4 evidencia a redução significativa para todas as vazões afluentes ao reservatório, confirmando a tendência observada na Figura 6. Essa tendência de redução é um indicativo de problemas para fins de geração de energia elétrica na UHE Queimado assim como para os demais usos a jusante do barramento. Por esse motivo, é

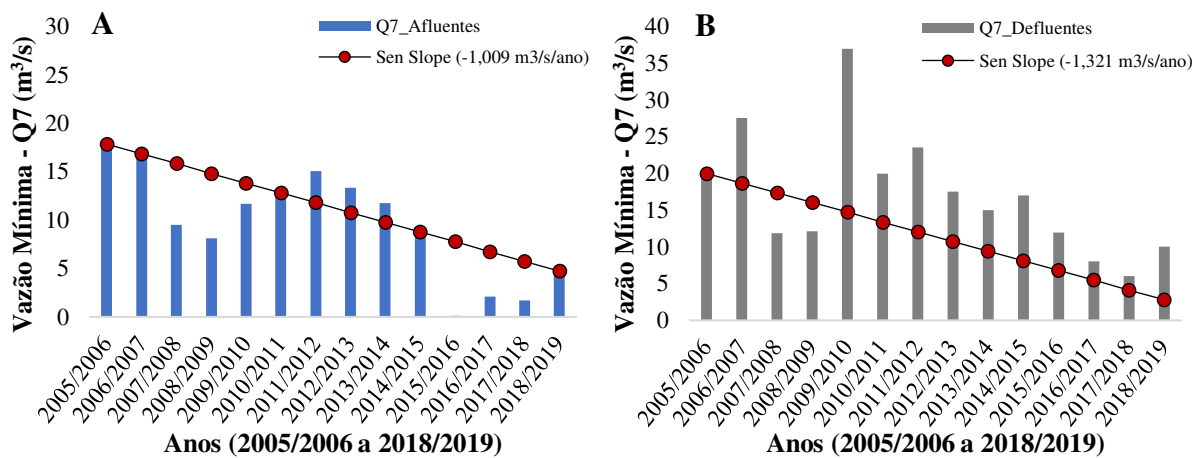
importante quantificar a proporção desses decréscimos ao longo dos anos, visando contribuir com a gestão da disponibilidade hídrica atual e futura.

A aplicação do estimador de Sen Slope forneceu a magnitude desses decréscimos ao longo dos anos para todas as variáveis consideradas (Figura 7). Para a vazão mínima, representada pela Q₇, o escoamento vem apresentando reduções de 1,009 m³/s por ano para as vazões afluentes e 1,321 m³/s por ano para as vazões defluentes.

Esses resultados indicam que mesmo o reservatório realizando a função de regularização das vazões a jusante do barramento, os valores de vazões defluentes são diretamente influenciados pelas características das vazões afluentes, de maneira que as reduções mais significativas encontradas nas vazões defluentes podem estar associadas a maior necessidade de conservação da água no reservatório a fim de garantir o suprimento no período de secas.

Para as vazões médias se verificou reduções anuais também bastante semelhantes, da ordem de 2,993 m³/s e 2,976 m³/s para as vazões afluentes e defluentes, respectivamente, indicando que a disponibilidade hídrica da bacia, bem como o seu potencial para geração de energia elétrica, está sendo reduzida ao longo do tempo.

Ao analisar as vazões máximas, os resultados indicaram que os decréscimos registrados para as vazões defluentes, ao longo dos anos, são maiores do que os encontrados para as vazões afluentes, de maneira que, a jusante do reservatório, as vazões reduzem na ordem de 10,286 m³/s ao ano, enquanto as vazões afluentes apresentaram reduções anuais de 7,125 m³/s.



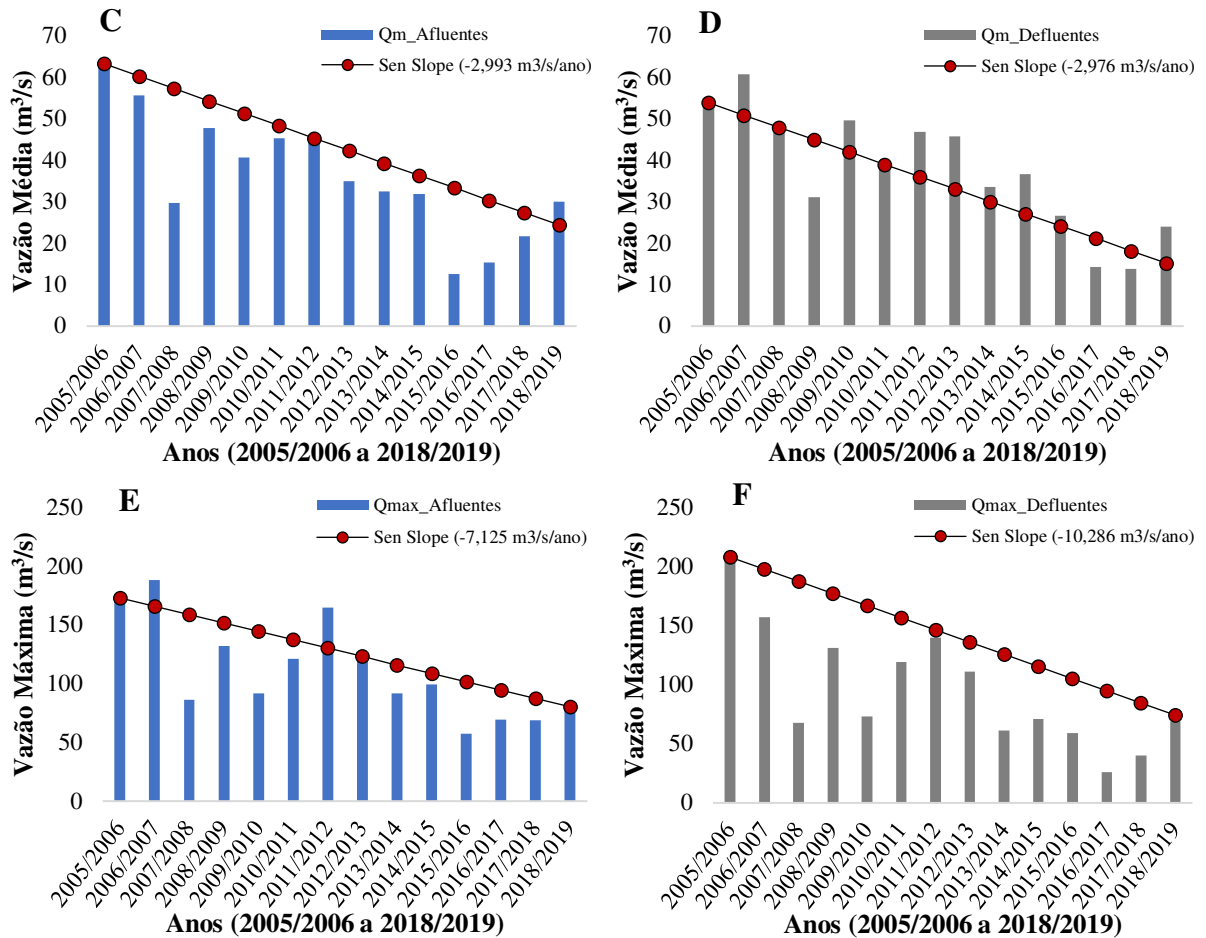


Figura 7 – Vazões Q_7 , Q_m e $Q_{m\acute{a}x}$ afluentes (A, C e E) e defluentes (B, D e F) do reservatório UHE Queimado e tendência das reduções expressas pelo estimador de Sen Slope, no período de 2005/2006 a 2018/2019.

A diferença entre esses resultados é esperada em função da regularização que o reservatório proporciona, retendo as vazões de maiores magnitudes no período chuvoso para então liberá-las no período seco. Dessa forma, as vazões máximas que passam pelo barramento estão submetidas ao controle operacional da usina, de tal maneira que parte fica retida e parte é disponibilizada a jusante.

Na Tabela 5 estão apresentados os resultados das análises de tendências para as estações convencionais da ANA localizadas a jusante do reservatório UHE Queimado.

Tabela 5 – Resultados dos testes de tendência de Mann-Kendall (MK)/Mann-Kendall Modificado (MKM) e Sen Slope para as vazões Q_7 , Q_m e $Q_{m\acute{a}x}$ das estações Q1, Q2 e Q3 localizadas a jusante do reservatório UHE Queimado

Séries Históricas	Variáveis	Mann-Kendall / Mann-Kendall Modificado (5%)				Sen Slope (5%) ($m^3/s/ano$)		
		p-valor	Z-valor	Tau de Kendall	Tendência*	Slope	Intervalo de confiança Superior	Inferior
Q1	Q_7	0,028	2,197	-0,451	↓	-1,111	-1,928	-0,112
	Q_m	0,000	3,504	-0,714	↓	-3,072	-4,384	-1,983
	$Q_{m\acute{a}x}$	0,004	2,847	-0,582	↓	-11,129	-17,496	-4,286
Q2	Q_7	0,021	2,299	-0,472	↓	-1,490	-3,201	-0,267
	Q_m	0,000	3,832	-0,780	↓	-4,270	-5,958	-2,668
	$Q_{m\acute{a}x}$	0,000	4,911	-0,648	↓	-17,825	-27,670	-11,098
Q3	Q_7	0,012	2,518	-0,516	↓	-2,917	-4,266	-1,028
	Q_m	0,000	3,504	-0,714	↓	-8,689	-10,514	-5,410
	$Q_{m\acute{a}x}$	0,324	0,985	-0,209	-	-	-	-

* (↓) tendência de redução; (↑) tendência de elevação e (-) sem tendência.

As séries de vazões mínimas, médias e máximas para as estações Q1 e Q2 também apresentaram tendências decrescentes, o que demonstra a influência da operação do reservatório nestas seções de monitoramento. As vazões mínimas em Q1 e Q2 apresentaram reduções de -1,111 e -1,490 $m^3/s/ano$, respectivamente, próximas às verificadas para as vazões afluentes e defluentes ao reservatório da UHE Queimado. Para as vazões médias e máximas das estações Q1 e Q2, as magnitudes das reduções anuais encontradas aumentaram conforme o aumento da área de drenagem das estações, sendo, portanto, maiores os decréscimos encontrados para a estação Q2, com vazões médias reduzindo 4,270 $m^3/s/ano$ e vazões máximas na taxa de 17,825 m^3/s ao ano.

Em Q3, estação mais distante do reservatório, as vazões mínimas apresentaram reduções de 2,917 $m^3/s/ano$ e as médias reduções de 8,689 $m^3/s/ano$. Essa estação apresenta um incremento na área de drenagem de cerca de 60% em relação a posição do reservatório. Isso, provavelmente, pode explicar a maior redução das vazões mínimas e média nessa estação em relação as estações Q1 e Q2, uma vez que se terão maiores variabilidades quanto ao tipo de solo, geomorfologia e, ainda, quanto à demanda de água pela irrigação. Contudo, para as vazões máximas não foi identificada tendência, caracterizando-se assim, como estacionárias ao longo do período analisado.

Diante das reduções das vazões (mínimas, média e máximas) observadas no presente estudo, aplicou-se também análises de tendência estatística para as séries de precipitação das estações utilizadas no estudo (Tabela 6) para avaliar se as reduções apresentadas estariam sendo influenciadas pelo comportamento climático da região.

A variabilidade da chuva dentro do período considerado, de maneira geral, se mostrou estacionária para o período anual, indicando que a chuva não apresentou mudanças significativas ao longo dos anos.

Tabela 6 – Resultados para os testes de tendência de Mann-Kendall (MK)/Mann-Kendall Modificado (MKM) para as séries de dados pluviométricos das estações selecionadas na bacia do rio Preto

Estações (Código)	Precipitação Total Anual			
	p-valor	Z-valor	Tau de Kendall	Tendência
01546005**	0,721	0,358	-0,111	-
01547003	0,228	1,204	-0,253	-
01547004**	0,381	0,876	-0,187	-
01547009**	0,858	0,179	-0,067	-
01547013**	0,428	0,793	-0,179	-
01547014**	0,246	1,159	-0,256	-
01547017**	0,721	0,358	-0,111	-
01547018**	1,000	0,000	0,011	-
01547021	0,324	0,985	-0,209	-
01547022	0,742	0,328	-0,077	-
01645019**	0,047	1,989	-0,455	↓
01646000	0,742	0,328	-0,077	-
01646001	0,079	1,752	-0,362	-
01646003	0,661	0,438	-0,099	-
01646004	0,127	1,525	-0,333	-
01647008	0,194	1,556	-0,382	-
01746001**	0,101	1,642	-0,341	-
01746002**	0,033	2,135	-0,461	↓
01746017**	0,033	2,135	-0,461	↓

* (↓) tendência de redução; (↑) tendência de elevação e (-) sem tendência.

** Estações localizadas no entorno da bacia do rio Preto

Apenas três estações indicaram mudanças significativas na precipitação total anual, cujos códigos são representados por 01645019, 01746002 e 01746017. Todas essas estações estão localizadas próximas à foz, fora dos limites da bacia, o que indica que as reduções das vazões nas posições de montante da bacia não apresentam influências consideráveis do comportamento da chuva e, portanto, não é possível afirmar que o comportamento das chuvas nessas posições é responsável pela redução das vazões afluentes e defluentes ao reservatório da UHE Queimado e nas estações Q1 e Q2.

Os resultados demonstram que a redução das vazões não pode ser associada a variação do regime pluviométrico anual, destacando uma possível preponderância das atividades

antrópicas sobre as condições climáticas, já que a bacia do rio Preto é uma fronteira de expansão agrícola com uso acentuado da irrigação. Essas evidências corroboram com aqueles obtidos por Xue et al. (2017), os quais observaram que as atividades humanas exerceram papel dominante na redução das vazões do rio Tarim, quando comparado com as variáveis climáticas, bem como com os resultados de Gao et al. (2011), que também identificaram que as contribuições das atividades humanas foram significativamente mais fortes para as reduções na vazão que a precipitação na bacia do rio Amarelo, na China.

Machado e Netto (2010) já mencionavam que na bacia do rio Preto, inserida no território do Distrito Federal, havia uma predominância de atividades agrícolas com uso expressivo de pivôs centrais para irrigação, que já contribuía para a redução da disponibilidade hídrica, especialmente no período de estiagem. Em seu estudo, os autores identificaram, que até o ano de 2007 haviam 207 pivôs centrais, representando 17.620 hectares de terras irrigadas ao longo de toda a extensão da bacia, com maiores concentrações na região noroeste.

A ANA, em parceria com a Embrapa, realizou o mapeamento do número de equipamentos de irrigação por pivô central no Brasil entre 1985 e 2019. Tomando como base o período considerado por este estudo (2005 a 2019), na bacia do rio Preto, o número de pivôs apresentou um crescimento muito significativo. Em 2010, a quantidade de pivôs centrais registrados representava 370 unidades, com área irrigada de 28.539,9 hectares, aumentando, em 2014, para 556 unidades, o que representou um acréscimo de 33,6% na área irrigada (42.979,8 hectares). Em 2017, o mapeamento passou a identificar a presença de 682 pivôs centrais distribuídos em toda a bacia, cuja implantação aumentou a área irrigada para 49.087,3 hectares. Já em 2019, o total de pivôs registrados passou a ser de 786 unidades, 53% a mais em relação a 2010, de maneira que cerca 488 destes, estavam concentrados na região noroeste da bacia (ANA, 2021).

Sendo assim, apesar do notório crescimento da atividade de irrigação por pivô-central na bacia do rio Preto, verifica-se na Figura 8 que a concentração desses equipamentos ocorre de forma mais acentuada nas regiões de cabeceira, a montante do reservatório UHE Queimado.

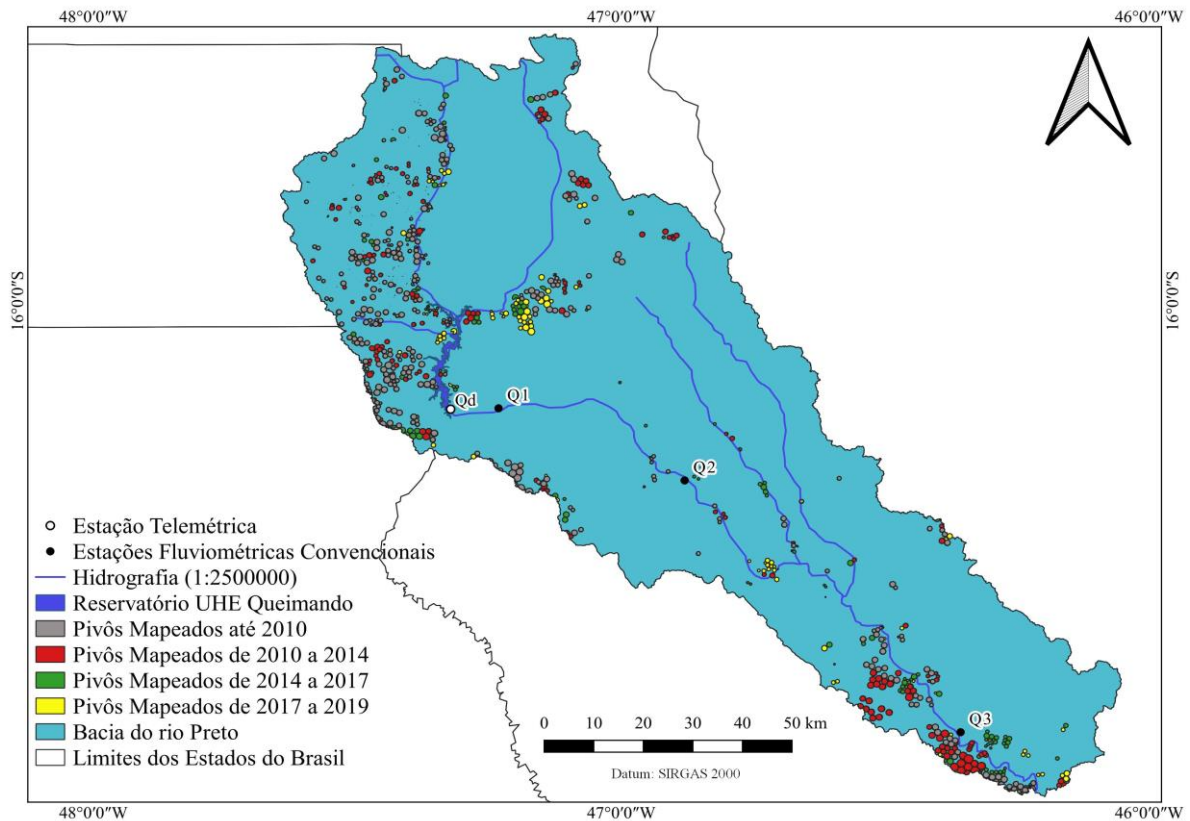


Figura 8 – Distribuição dos pivôs centrais na bacia do rio Preto para os anos de 2010, 2014, 2017 e 2019.

O incremento acentuado no número de pivôs centrais na bacia do rio Preto pode representar uma parcela significativa das reduções das vazões ao longo dos anos, especialmente para aquelas que afluem o reservatório UHE Queimado, impactando assim todas as demais atividades influenciadas por ele.

Associado ao crescimento do número de pivôs na bacia, conflitos pelo uso da água vem tomando maiores proporções ao longo dos anos, principalmente no período de estiagem, tornando necessária a mediação por parte dos órgãos gestores. Segundo Maniçoba (2019) a área do Comitê de Bacia Hidrográfica do Preto do Distrito Federal (CBH Preto – DF) é irrigada para produção de grãos e apresenta conflitos de água que têm sido tratados em reuniões de alocação negociada.

No ano de 2016, a ANA, através da Resolução nº 934, autorizou a redução da descarga mínima do reservatório da UHE Queimado, levando em consideração a importância de se preservar o estoque de água disponível e, ainda, mencionando a situação hidrometeorológica desfavorável para o trecho a montante do reservatório. A redução da vazão defluente mínima passou de 17 m³/s para 10 m³/s, sendo permitida apenas entre os meses de agosto a outubro daquele ano.

Apesar da resolução mencionar a crise hidrometeorológica naquele momento, é importante lembrar que a área irrigada de 2007 até 2017 quase que triplicou e, ainda, permanece em constante expansão, onde dificuldades de gestão e gerenciamento já se tornaram uma realidade da bacia.

É interessante mencionar, ainda, que pequenos reservatórios desempenham um papel importante no apoio à economia local e são usados principalmente para o fornecimento de água para irrigação e abastecimento do gado (Althoff et al., 2019). Na bacia do rio Preto, segundo levantamento feito pelo estudo de Rodrigues et al. (2012), foram identificados a presença de 147 desses pequenos barramentos, com área superficial entre 1 e 50 ha. Segundo os autores, a porção do Distrito Federal e Goiás, que se enquadram na bacia do rio Preto, considerada nesse estudo como a mais representativas para a área a montante do reservatório UHE Queimado, possuíam capacidade de armazenar, respectivamente, 14% (um a cada 30 km² - 44 reservatórios) e 29% (um a cada 70 km² - 32 reservatórios) da água da bacia nessas estruturas. Apesar disso, foram encontrados registros de outorgas, entre 2011 e 2019, para apenas 30 dessas unidades considerando a análise da mesma região, conforme informações disponibilizadas na plataforma da Agencia Reguladora de Águas, Energia e Saneamento do Distrito Federal (ADASA).

Essas informações permitem chamar a atenção para o fato de que a presença desses pequenos barramentos, com a finalidade de aumentar a disponibilidade hídrica em determinados trechos da hidrográfica, também podem estar contribuindo para redução das vazões afluentes ao reservatório UHE Queimado, especialmente no que diz respeito às vazões máximas.

2.4. CONCLUSÕES

- As vazões afluentes e defluentes ao reservatório UHE Queimado apresentaram fortes correlações, indicando o alto nível de importância das vazões afluentes em relação a operacionalização do reservatório, sendo os melhores resultados obtidos quando se considerou, na sua grande maioria, intervalos trimestrais de dados.

- A variação das vazões médias da estação Q1 e Q2 é muito fortemente dependente das vazões médias defluentes do reservatório, enquanto a variação das vazões média da estação Q3, localizada mais distantes do reservatório, é menos influenciada pela operação do reservatório e

mais dependentes do comportamento hidrológico da bacia e do avanço das atividades usuárias (ex. irrigação) dos recursos hídricos.

- Todas as vazões analisadas (Q_7 , Q_m e $Q_{m\acute{a}x}$), afluentes e defluentes ao reservatório UHE Queimado, apresentaram tendência decrescente no período de 2005 a 2019.

- As estações fluviométricas convencionais localizadas a jusante do reservatório UHE Queimado (Q1, Q2 e Q3) apresentaram tendência decrescente para todas as vazões analisadas (Q_7 , Q_m e $Q_{m\acute{a}x}$), à exceção daquela localizada mais distante (Q3) na qual não se evidenciou tendência para $Q_{m\acute{a}x}$.

- A precipitação total anual não apresentou tendência em nenhuma das estações localizadas na bacia do rio Preto, de maneira que a ela não se pode atribuir justificativas para as reduções encontradas nas vazões afluentes e defluentes ao reservatório UHE Queimado assim como nas estações convencionais localizadas a jusante.

- As reduções observadas nas vazões afluentes e defluentes ao reservatório UHE Queimado, assim como nas estações localizadas a jusante, podem estar relacionadas ao aumento de quase três vezes na área irrigada por pivô-central na bacia do rio Preto no período entre 2005 e 2019, especialmente a montante do reservatório UHE Queimado, dando um forte indicativo para o comprometimento das vazões ao longo da hidrografia.

2.5. REFERÊNCIAS

- Abeysingha, N. S., Singh, M., Sehgal, V. K., Khanna, M., Pathak, H. 2016. Analysis of trends in streamflow and its linkages with rainfall and anthropogenic factors in Gomti River basin of North India. *Theor. Appl. Climatol*, 123. pp.785-799.
- Althoff, D., Rodrigues, L. N., Silva, D. D. da., Bazame, H. C. 2019. Improving methods for estimating small reservoir evaporation in the Brazilian Savanna. *Agricultural Water Management*. doi: 10.1016/j.agwat.2019.01.028.
- Agência Nacional de Águas e Saneamento Básico – ANA. 2021. Atlas Irrigação: Uso da água na agricultura irrigada. 2ª ed. Brasília. 130p.
- Back, A. J. 2001. Aplicação de análise estatística para identificação de tendências climáticas. *Revista Pesquisa Agropecuária Brasileira*, v.36, n.5. p.717-726.

- Bozzoni, P.L., Mello Junior, A.V. 2020. Previsões de precipitação de modelos atmosféricos como subsídio à operação de sistemas de reservatório. *Revista Brasileira de Meteorologia*, v. 35, n. 1, p. 99-109.
- Busker, T., Roo, A. de., Gelati, E., Schwatke, C., Adamovic, M., Bisselink, B., Pekel, J-F., Cottam, A. 2018. A global lake and reservoir volume analysis using a surface water dataset and satellite altimetry. *Hydrology and Earth System Sciences*
- Chang, L.-C., Chang, F.-J. 2001. Intelligent control for modelling of real-time reservoir operation. *Hydrological Processes*, 15(9), 1621–1634. doi:10.1002/hyp.226.
- Chauluka, F., Singh, S., Kumar, R. 2021. Raifall and streamflow trends of Thuchila River, Southern Malawi. *Materials Today: Proceedings* 31. p. 846-855. doi: <https://doi.org/10.1016/j.matpr.2020.06.228>.
- Dubey, S. K., Sharma, D., Babel, M. S., Mundetia, N. 2020. Application of hydrological model for assessment of water security using multi-model ensemble of CORDEX-South Asia experiments in a semi-arid river basin of India. *Ecological Engineering*, 143, 105641. doi:10.1016/j.ecoleng.2019.105641
- Ferrari, A. L. 2012. Variabilidade e tendência da temperatura e pluviosidade nos municípios de Pirassununga, Rio Claro, São Carlos e São Simão (SP): Estudo sobre mudança climática de curto prazo em escala local. Tese. Universidade de São Paulo-USP. São Carlos/SP, 156 p.
- Gao, L., Mu, X. M., Wang, F., Li, R. 2011. Changes in streamflow and sediment discharge and the response to human activities in the middle reaches of the Yellow River. *Hydrology and Earth System Sciences*, 15, p.1-10.
- Hamed, K. H., Rao, R. 1998. “A modified Mann-Kendall trend test for auto-correlated data”. *Journal of Hydrology*, 204, pp.182-196.
- Helsel, D. R., Hirsch, R. M. 2002. “Statistical methods in water resources”. U.S. Geological Survey. 522 p.
- Helsel, D. R., Hirsch, R. M., Ryberg, K. R., Archfield, S. A., Gilroy, E. J., 2020, *Statistical methods in water resources: U.S. Geological Survey Techniques and Methods*, book 4, chapter A3, 458 p., <https://doi.org/10.3133/tm4a3>. [Supersedes USGS Techniques of Water-Resources Investigations, book 4, chapter A3, version 1.1.]
- Hirsch, R. M., Slack, R. J., Smith, R. 1982. Techniques of trend analysis for monthly water quality data. *Water Resources Research*. v.18, n.1, p.107-121.
- Kendall, M. G. 1975. *Rank Correlation Methods*. Charles Griffin, London.
- Maniçoba, A. M. 2019. Um modelo de gestão participativa: o uso compartilhado de recursos hídricos na Unidade Hidrográfica do Ribeirão Extrema – DF. Dissertação de Mestrado – Pós-Graduação em Saúde Pública, Fundação Oswaldo Cruz. Brasília-DF
- Mann, H. B. 1945. Nonparametric tests against trend. *Econometrica* 13, 245-259.

- Machado, N. D. S., Netto, O. M, C. 2010. Análise econômica aplicada à decisão sobre alocação de água entre os usos irrigação e produção de energia elétrica: o caso da bacia do rio Preto (DF/GO/MG). *Revista Brasileira de Recursos Hídricos*, v.15, n.4. p.105-119.
- Neeti, N., Eastman, J. R. 2011. A Contextual Mann-Kendall Approach for the Assessment of Trend Significance in Image Time Series. *Transactions in GIS*, 15(5): 599-611.
- Nourani, V., Rouzegari, N., Molajou, A., Baghanam, A. H. 2020. An integrated simulation-optimization framework to optimize the reservoir operation adapted to climate changes cenários. *Journal of Hydrology*, 587. doi: <https://doi.org/10.1016/j.jhydrol.2020.125018>.
- Organização Meteorológica Mundial – OMM. 1975. Drought and agriculture. WMO Tech. Note 138, Publ. WMO-392, Geneva, Switzerland, 127p.
- Organização Meteorológica Mundial – OMM. 1988. Analyzing long timeseries of hydrological data with respect to climate variability. Wcap-3, WMO/TD 224.
- Operador Nacional do Sistema Elétrico – ONS. 2020. Plano de Operação Energética 2020/2024. Disponível em: < [REVISTA_PEN_2020_versao20201112.pdf](#) (ons.org.br)>, acesso em 17 de maio de 2021.
- Operador Nacional do Sistema Elétrico – ONS. 2020. Manual de procedimentos de operação: Módulo 5 – Submódulo 5.12. Disponível em: < http://www.ons.org.br/%2FMPO%2FDocumento%20Normativo%2F3.%20Instru%C3%A7%C3%B5es%20de%20Opera%C3%A7%C3%A3o%20-%20SM%205.12%2F3.6.%20Opera%C3%A7%C3%A3o%20de%20Reservat%C3%B3rios%2FIO-OR.SF.SFR_Rev.23.pdf> Acesso em: 18 de maio de 2021.
- Passaia, O. A., Paiva, R. C. D. 2019. Análise exploratória do comportamento de reservatórios do sistema interligado nacional. XXIII Simpósio Brasileiro de Recursos Hídricos. Foz do Iguaçu.
- Pereira, G., Cardozo, F. da S., Negreiros, A. B. de., Zanin, G. D., Costa, J. C. da., Lima, T. E. R., Rufino, P. R., Ramos, R. de C. 2018. Análise da variabilidade da precipitação para o estado de minas gerais (1981-2017). *Revista Brasileira de Climatologia*. doi: <http://dx.doi.org/10.5380/abclima.v1i0.61028>
- Rippl, W. 1883. The capacity of storage-reservoirs for water-supply. *Proc. Inst. Civil Engrs*, 71, p. 270-278.
- Rodrigues, L. N., Sano, E. E., Steenhuis, T. S., Passo, D. P. 2012. Estimation os small reservoir storage capacities with remote sensing in the Brazilian Savannah region. *Water Resour Manage*. p. 873-882. doi: 10.1007/s11269-011-9941-8.
- Salviano, M. F., Groppo, J. D., Pellegrino, G. Q. 2016. Análise de Tendências em Dados de Precipitação e Temperatura no Brasil. *Revista Brasileira de Meteorologia*, 31(1), 64–73. doi:10.1590/0102-778620150003

- Schaeffli, B. 2015. Projecting hydropower production under future climates: a guide for decision-makers and modelers to interpret and design climate change impact assessments. *Wiley Interdisciplinary Reviews: Water*, 2(4), 271–289. doi:10.1002/wat2.1083
- Silva, F. B., Almeida, L. T. de., Vieira, E. O. de., Silva, D. D. da, Maciel, I. P., Parma Júnior, F. 2020. Pluviometric and fluviometric trends in association with future projections in areas of conflict for water use. *Journal of Environmental Management*. doi: 10.1016/j.jenvman.2020.110991
- Silva, R. M. da., Santos, C. A. G., Moreira, M., Corte-Real, J., Silva, V. C. L., Medeiros, I. C. 2015. Rainfall and river flow trends using Mann–Kendall and Sen’s slope estimator statistical tests in the Cobres River basin. *Natural Hazards*, 77(2), 1205–1221. doi:10.1007/s11069-015-1644-7
- Sen, P. K. 1968. Estimates of the regression coefficient based on Kendall’s Tau. *J. Am. Stat. Association*, v. 63, p. 1379-1389.
- Shadkam, S., Ludwig, F., van Oel, P., Kirmit, Ç., Kabat, P. 2016. Impacts of climate change and water resources development on the declining inflow into Iran’s Urmia Lake. *Journal of Great Lakes Research*, 42(5), 942–952. doi:10.1016/j.jglr.2016.07.033
- Spalding-Fecher, R., Chapman, A., Yamba, F., Walimwipi, H., Kling, H., Tembo, B., Nyambe, I., Cuamba, B. 2014. The vulnerability of hydropower production in the Zambezi River Basin to the impacts of climate change and irrigation development. *Mitig. Adapt. Strateg. Glob. Chang*, 19, 1–22.
- Setti, A. A., Lima, J. E. F. W., Chaves, A. G. de M., Pereira, I. de C. 2000. *Introdução ao gerenciamento de recursos hídricos*. Agência Nacional de Energia Elétrica, Superintendência de Estudos e Informações Hidrológicas. 2ª Ed. Brasília. 207p.
- Tucci, C. E. M., Braga, B. 2003. *Clima e Recursos Hídricos no Brasil*. Porto Alegre: ABRH.
- Wu, J., Liu, Z., Yao, H., Chen, X., Chen, X., Zheng, Y., He, Y. 2018. Impacts of reservoir operations on multi-scale correlations between hydrological drought and meteorological drought. *Journal of Hydrology*, 563, 726–736. doi:10.1016/j.jhydrol.2018.06.053
- Xue, L., Yang, F., Yang, C., Chen, X., Zhang, L., Chi, Y., Yang, G. 2017. Identification of potential impacts of climate change and anthropogenic activities on streamflow alterations in the Tarim River Basin, China. *Scientific Reports*.
- Yue, S., Pilon, P., Cavadias, G. 2002. “Power of the Mann-Kendall and Spearman’s rho tests for detecting monotonic trends in hydrological series”. *Journal of Hydrology* 259, pp. 254-271

CAPÍTULO 3:

Proposta metodológica para avaliar a estimativa de vazões pelos métodos do balanço hídrico e da curva-chave baseado no método de regressão logística

RESUMO: Para a estimativa dos valores de vazão defluente aos reservatórios de regularização é utilizado o método do balanço hídrico, de maneira que os dados são coletados, por força de lei, em intervalos horários, possibilitando a observação quanto ao comportamento diário do escoamento no curso d'água. Já no método da curva-chave, normalmente adotado pelas estações fluviométricas convencionais, os valores são estimados a partir da média de duas leituras diárias, fortemente influenciados pelas características da área da sessão de medição do escoamento. Além disso, o método da curva-chave requer o uso de técnicas de extrapolação a fim de possibilitar a estimativa de valores fora dos limites de validação. Sendo assim, o objetivo deste trabalho foi analisar a variação das vazões defluentes horárias do reservatório e ainda propor um método para avaliar se a estimativa das vazões diárias defluentes do reservatório UHE Queimado, obtidas pelo método do balanço hídrico, são semelhantes aos valores de vazão obtidos em uma estação fluviométrica (estimadas pelo método da curva-chave) localizada a jusante do referido reservatório. Os resultados mostraram que os valores médios obtidos entre as observações feitas a partir da média de 2 ou 24 leituras foram muito próximos, considerando assim que os dados estimados para duas leituras diárias apresentam boa representatividade quanto a variabilidade das vazões na região. Os métodos do balanço hídrico e da curva-chave apresentaram desempenho estatístico semelhantes, para a maioria dos dados considerados. A regressão logística foi capaz de identificar dados atípicos nas séries históricas de vazões analisadas, especialmente no período chuvoso, estando estes associados principalmente aos métodos adotados para extrapolação da curva-chave. Assim, o estudo traz vantagens por permitir que o processo de consistência das estimativas de vazões pelo método da curva-chave se torne menos trabalhoso e mais seguro, quando os critérios para sua aplicação puderem ser atendidos.

Palavras-chave: extrapolação, modelagem, vazões defluentes, estações fluviométricas

ABSTRACT: The water balance method is used to estimate the values of effluent flow to the regularization reservoirs, so that the data are collected, by force of law, at hourly intervals, allowing the observation of the daily behavior of the flow in the watercourse. In the rating curve method, normally adopted by conventional fluviometric stations, the values are estimated from

the average of two daily readings, strongly influenced by the characteristics of the area of the flow measurement session. Furthermore, the rating curve method requires the use of extrapolation techniques in order to allow the estimation of values outside the validation limits. Therefore, the objective of this work was to analyze the variation of the reservoir's hourly effluent flows and also to propose a method to evaluate whether the estimate of the daily effluent flows of the Queimado UHE reservoir, obtained by the water balance method, are similar to the flow values obtained at a fluviometric station. The results showed that the average values obtained between the observations made from the average of 2 or 24 readings were very close, considering that the estimated data for two daily readings present good representation regarding the variability of the flows in the region. The water balance and rating curve methods presented similar statistical performance for most of the data considered. The logistic regression was able to identify atypical data in the historical series of analyzed flows, especially in the rainy season, which are mainly associated with the methods adopted for extrapolation of the rating curve. Thus, the study has advantages because it allows the process of consistency of flow estimates using the rating curve method to become less laborious and safer, when the criteria for its application can be met.

Keywords: extrapolation, modeling, effluent flows, fluviometric stations

3.1. INTRODUÇÃO

As usinas hidrelétricas podem apresentar diversas finalidades construtivas pois, conforme afirmam Wannasin et al. (2021), além do foco na geração de energia, os reservatórios de acumulação de água das usinas podem atenuar eventos de inundação e seca ao regularizar o fluxo, com picos de inundação menores e vazões mínimas mais elevadas. Contudo, segundo os mesmos autores mencionados, quando a gestão é ineficaz, os extremos hidrológicos podem ser intensificados, podendo, em alguns casos, ter até maior influência que as mudanças climáticas.

Cheng et al. (2019) comentam que o controle da vazão em unidades hidrelétricas é extremamente importante por diversos motivos, dentro os quais pode-se destacar a avaliação do desempenho quanto à operação, ao monitoramento dos equipamentos, à garantia das condições necessárias para preservação do ambiente aquático saudável a jusante do barramento e, ainda, para cumprir parâmetros de liberação específicos de água aos quais são submetidos e que podem variar sazonalmente e/ou devido a condições específicas (inundações ou secas).

É fundamental que o monitoramento da operação nessas estruturas seja feito de maneira adequada, em intervalos regulares, para a compreensão do comportamento e, conseqüentemente, para o apoio às tomadas de decisões. A Resolução Conjunta ANEEL/ANA nº 03, de 10 de agosto de 2010, determina que, em situações em que há a presença de usinas hidrelétricas com reservatórios de regularização, as estações hidrométricas com monitoramento pluviométrico, linimétrico e fluviométrico sejam automatizadas e telemetrizadas, de maneira que os dados coletados por essas estações possam ser registrados em intervalo máximo de uma hora.

Nesses casos, para o cálculo da quantidade de água que entra e sai de um reservatório, em certo período de tempo, normalmente é usado o método do balanço hídrico, sendo que para a determinação das vazões defluentes, é necessária a obtenção de duas variáveis: a vazão turbinada, retirada do reservatório pelo conduto forçado que conduz a água para as turbinas visando a geração de energia, e a vazão vertida, liberada pelos vertedores do reservatório e que não passa pelas turbinas e, portanto, não gera energia (Molina, 2016).

A exigência da coleta de dados em intervalos horários em usinas hidrelétricas permite que 24 leituras distribuídas ao longo do dia sejam realizadas, o que aumenta a precisão das estimativas médias diárias, uma vez que minimiza erros causados pela variabilidade do escoamento ao longo do dia e, ainda, ajuda a compreender a dinâmica das demandas por energia elétrica na região. Sendo assim, é possível avaliar os horários de maiores e menores turbinagens de água, o que contribui para que leituras de vazão diária média apresentem maior representatividade quanto ao regime de variação de vazões.

Por outro lado, em trechos onde não há a presença de usinas hidrelétricas, o método mais utilizado para estimativas das vazões nos cursos d'água é o da curva-chave, que é baseado na medição de dados de cota ou nível de água do rio. A construção da curva-chave é limitada pelo número de medições realizadas na seção considerada, de modo que, é feita uma leitura do nível da água, através de uma régua linimétrica, e a medição da respectiva vazão, levando em consideração a velocidade média e a área molhada da seção transversal. Essa área, contudo, pode sofrer modificações ao longo do tempo, já que as características do leito do rio se alteram com o passar dos anos, o que significa que a curva-chave pode perder sua representatividade, tornando necessárias campanhas frequentes de medição de vazão em campo (Manfreda, 2018).

A relação entre cota e vazão, medidas pontualmente ao longo do tempo para fins de obtenção da curva-chave, pode ser expressa por meio de uma equação, o que permite que leituras contínuas de cota sejam convertidas em séries de vazões. Nas estações pertencentes à Rede Hidrometeorológica Nacional (RHN), sob responsabilidade da Agência Nacional de

Águas e Saneamento Básico (ANA), a leitura de cotas é feita diariamente em dois períodos do dia, às 7 e 17 horas, sendo o valor médio das duas leituras utilizado como representativo para fins da estimativa da vazão média diária.

Em situações em que se tem vazões máximas, é comum, entretanto, a aplicação de metodologias para extrapolação da curva-chave, já que normalmente, por problemas operacionais, não se têm medições de vazões associadas aos períodos mais críticos de cheias, o que segundo Grison e Kobiyama (2009) se justifica por tornar perigoso para a equipe de hidrometria realizar medições nestas situações onde a velocidade de escoamento é alta.

Sendo assim, torna-se importante avaliar a estimativa das vazões considerando a comparação entre os métodos do balanço hídrico e da curva-chave, uma vez que ambos são reconhecidamente respeitados e aplicados mundialmente. Dessa forma, o estudo partiu do princípio de que o método do balanço hídrico apresentaria menos limitações que o método da curva-chave em função, principalmente, da natureza das extrapolações que se faz necessária para a obtenção das vazões máximas no método da curva-chave. Conforme explicam Peña-Arancibia et al. (2015) há a presença de heterocedasticidade (variâncias distintas) entre a vazão medida e os valores estimados pelas curvas-chaves, especialmente nos maiores valores de vazão, para várias estações, enfatizando a alta incerteza potencialmente causada pela extrapolação das curvas.

Diante do exposto, este estudo teve por objetivo: i) analisar o comportamento das vazões defluentes horárias ao reservatório da Usina Hidrelétrica de Queimado (UHE Queimado); e ii) propor uma metodologia para avaliar as vazões diárias defluentes ao reservatório UHE Queimado, estimadas pelo método do balanço hídrico, comparativamente as vazões diárias da estação fluviométrica convencional, estimadas pelo método da curva-chave, localizada a jusante do referido reservatório, utilizando para isso, o método de regressão logística.

3.2. MATERIAL E MÉTODOS

3.2.1. Área de estudo

O estudo foi realizado utilizando informações relacionadas ao reservatório da Usina Hidrelétrica de Queimado (UHE Queimado), localizado na bacia do rio Preto (10.325 km²), sub-bacia do rio Paracatu, inserida, por sua vez, na bacia hidrográfica do rio São Francisco. A

bacia do rio Preto abrange parte do Distrito Federal (13%) e dos estados de Goiás (22%) e Minas Gerais (65%).

A UHE Queimado (Figura 1), de propriedade do Consórcio entre a Companhia Energética de Minas Gerais (CEMIG) e a Companhia Energética de Brasília (CEB), está localizada em Palmital de Minas, distrito pertencente ao município mineiro de Cabeceira Grande. A Usina entrou em operação em 2004, apresenta área de contribuição de aproximadamente 3.657 km² e o seu reservatório possui um volume útil de 389,46 hm³, devendo sua operação respeitar os usos múltiplos.

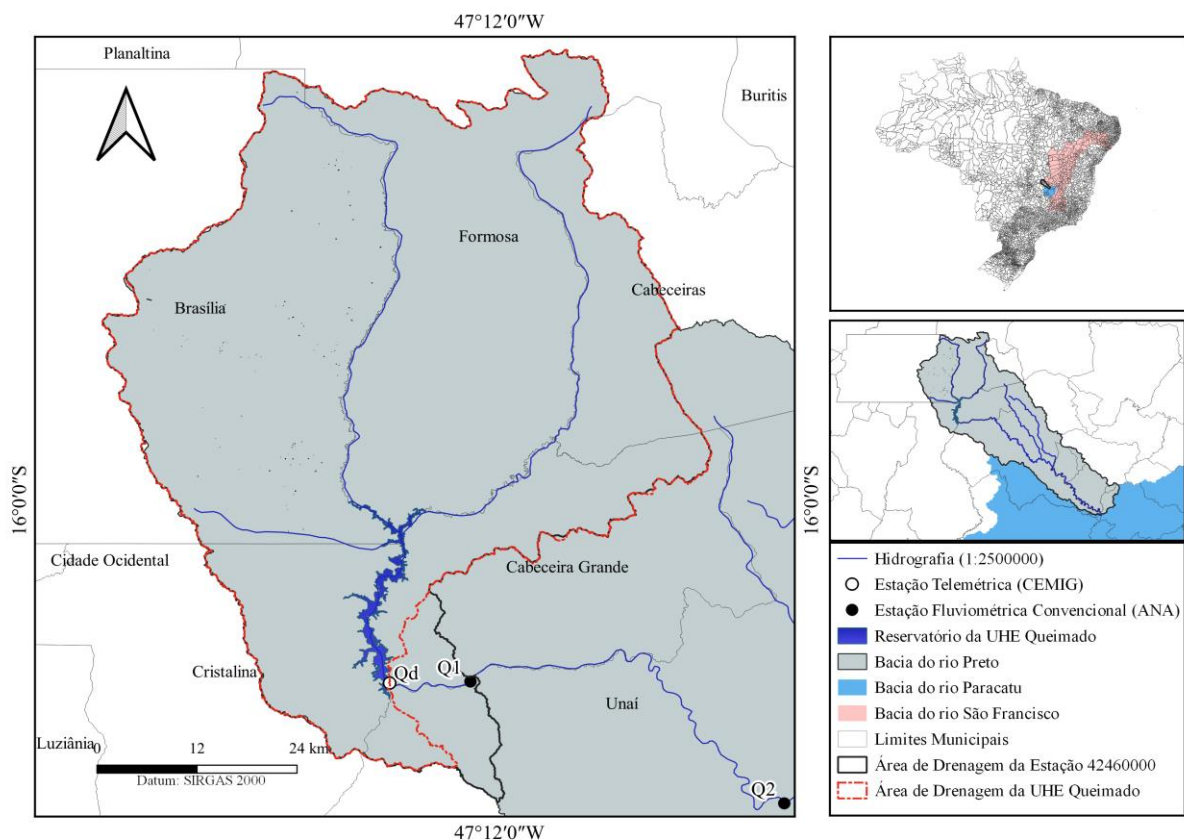


Figura 1 - Localização da área de estudo.

Imediatamente a jusante do reservatório encontra-se localizada a estação fluviométrica convencional da ANA (Agência Nacional de Águas e Saneamento Básico), denominada Fazenda Limeira (4246000), com área de drenagem equivalente a 3.890 km², representada na Figura 1 por Q1. A distância da estação Q1 da seção onde ocorre a defluência do reservatório, Qd, é de aproximadamente 10 km.

A região apresenta duas estações do ano bem definidas, com início do período chuvoso em novembro e término em abril, e estação seca, com baixa incidência de chuvas, entre maio e outubro.

3.2.2. Dados hidrológicos

Os dados das vazões horárias defluentes ao reservatório da UHE Queimado foram obtidos junto à CEMIG, de modo que para cada dia encontram-se registradas 24 leituras horárias, com início às 0:00 e término às 23:00 horas, no período de 2004 a 2019. Em função das inconsistências encontradas nas séries de vazões horárias para o período de 2004 a 2005, provavelmente justificadas pelas instabilidades decorrentes do início das operações em 2004, optou-se por excluir essas informações e trabalhar com o período de 2006 a 2019.

Dentro do período trabalhado, também foram excluídos os dias que apresentaram falhas em pelo menos um horário do dia, já que a ausência de uma dessas informações poderia comprometer as análises, especialmente aquelas cujo objetivo era justamente avaliar a variabilidade das vazões ao longo do dia.

Os dados históricos e a curva-chave da estação fluviométrica Q1, localizada imediatamente a jusante do reservatório UHE Queimado, foram obtidos na plataforma Hidroweb (<https://www.snirh.gov.br/hidroweb/serieshistoricas>), que inclui dados da RHN (Rede Hidrometeorológica Nacional) de responsabilidade da ANA. Entretanto, esta estação só apresenta dados coerentes até o ano de 2014, o que conduziu ao estabelecimento do período de 2006 a 2014 para fins da análise comparativa com os dados de vazão defluente ao reservatório (Qd).

A fim de apoiar as análises entre as seções de interesse foram obtidos, ainda, na plataforma do Hidroweb, os dados de precipitação da estação pluviométrica Fazenda Limeira (01647008), localizada próxima à seção da estação fluviométrica Q1, considerando o período de 2006 a 2014.

3.2.3. Análise do comportamento horário das vazões defluentes

A variabilidade das vazões defluentes ao reservatório (Qd) ao longo dos dias é muito grande, e por isso foi avaliado o comportamento da vazão ao longo do período de 24 horas utilizando o valor médio mensal de cada um dos horários considerados, para o período de 2006 a 2019, buscando identificar os momentos em que se tem maiores amplitudes das turbinagens de água em função da produção de energia elétrica. Foram avaliadas, ainda, possíveis diferenças entre as médias das 24 horas e dos horários de 7 e 17 horas (utilizados na RHN), buscando verificar se há diferenças consideráveis no comportamento das vazões ao longo do dia.

3.2.4. Comparação dos métodos para estimativa de vazão

Para esta análise foi aplicado um critério para a seleção dos dados, considerando a identificação e remoção de valores incompatíveis entre as vazões defluentes (Q_d) e as da estação Q1. Sendo assim, foram removidos os dados que apresentaram relação $Q1/Q_d \leq 0,90$, uma vez que se partiu da premissa básica de que estações localizadas a jusante tendem a apresentar valores de vazões diárias superiores aos das estações de montante, o que é reforçado no presente estudo em função de não terem sido identificadas captações expressivas de água entre a UHE Queimado e a estação Q1. Portanto, 10% foi o limite de permissividade incorporado em função do nível de precisão dos equipamentos e de possíveis captações não identificadas ao longo do trecho do curso d'água entre as duas seções consideradas. Além disso, também foram eliminados os dados cujas relações $Q1/Q_d$ foram classificadas como outliers a partir da aplicação do boxplot.

Na sequência, os dados de vazão defluente (Q_d) e da estação fluviométrica convencional da ANA (Q1), em m^3/s , foram convertidos para vazão específica, em $L/s.km^2$ (q_d e q_1 , respectivamente), a fim de viabilizar a comparação dos valores obtidos com os diferentes métodos de estimativa de vazão para as diferentes áreas de drenagem, a saber: curva-chave (adotado pela ANA para estimativa de Q1) e balanço hídrico (adotado pela CEMIG para estimativa de Q_d).

Para que as metodologias fossem avaliadas de maneira mais uniforme foi considerado o mesmo intervalo de tempo para obtenção dos dados, tanto na vazão defluente quanto na estação convencional da ANA. Sendo assim, para esta análise foi levada em consideração as vazões específicas diárias q_{d7-17} , obtidas a partir da média dos valores observados às 7 e 17 horas, bem como as vazões específicas diárias q_1 , que são estimadas utilizando o mesmo intervalo de tempo.

Inicialmente uma análise da distribuição dos dados de vazão defluente e da estação fluviométrica convencional foi feita, via gráfico quantil-quantil (qq-plot). Essa ferramenta foi proposta por Wilk e Gnanadesikan (1968) e conforme explicam Nobre et al. (2017), é um gráfico que é utilizado para observar se dois conjuntos de dados pertencem a mesma distribuição de probabilidades, sendo que os pontos são formados pelos quantis amostrais, de modo que se no resultado os pontos se alinharem, formando uma reta, as duas amostras poderão ser consideradas as mesmas.

Na sequência, as séries originais de vazão (q_{d7-17} e q_1) foram testadas a fim de verificar se poderiam ou não ser classificadas como estatisticamente semelhantes. Para isso foi aplicado

o teste F, na forma modificada de Graybill (1976). O teste F de Graybill é recomendado quando se objetiva saber se as estimativas obtidas por uma equação (Y_j) são estatisticamente iguais aos valores observados (Y_1) ou se as estimativas de uma determinada variável, obtidas por dois métodos de avaliação ($Y_1 = Y_j$), são estatisticamente iguais, o que justifica a sua aplicação nesse trabalho.

Se as estimativas de Y_1 e Y_j forem exatamente iguais, após o ajuste de uma regressão linear, $Y_j = \beta_0 + \beta_1 Y_1 + \varepsilon_i$, os coeficientes da equação seriam, simultaneamente, $\beta_0 = 0$ e $\beta_1 = 1$, definindo uma reta com 45° passando na origem.

Sendo assim, a série de dados de vazão defluente, obtida a partir do método do balanço hídrico (qd₇₋₁₇), foi aqui considerada como o método padrão, tendo as vazões estimadas na estação na estação da ANA (q1), pelo método da curva-chave, como método de estimativa alternativo.

O teste F de Graybill foi aplicado, conforme apresentado na equação 1, no qual foi testada a hipótese $H_0: \beta_0 = 0$ e $\beta_1 = 1$.

$$F(H_0) = \frac{(\beta - \theta)'(Y_1'Y_1)(\beta - \theta)}{2QMR} \quad \text{Equação 1}$$

em que:

$$\beta = \begin{bmatrix} \hat{\beta}_0 \\ \hat{\beta}_1 \end{bmatrix} \quad \text{Equação 2}$$

$$\theta = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \quad \text{Equação 3}$$

$$(Y_1'Y_1) = \begin{bmatrix} n & \sum Y_1 \\ \sum Y_1 & \sum Y_1^2 \end{bmatrix} \quad \text{Equação 4}$$

O QMR é o quadrado médio dos resíduos da regressão $Y_j = \beta_0 + \beta_1 Y_1$. Quando $F(H_0) \geq F_\alpha(2, n - 2)$, a hipótese foi rejeitada. Quando $F(H_0) < F_\alpha(2, n - 2)$, a hipótese não foi rejeitada, admitindo a identidade entre os dois métodos avaliados, isto é $Y_1 = Y_j$ a um nível α de significância. Neste estudo foi considerado o nível de 1% de significância (0,01). A aplicação deste teste foi feita utilizando os recursos do pacote **forestmangr** (Braga et al., 2021).

Constatada as diferenças estatísticas entre as séries, estas foram novamente submetidas ao um novo teste F, agora considerando apenas as observações em que o módulo da diferença entre q1 e qd₇₋₁₇ (*diferença absoluta* = $|q1 - qd_{7-17}|$) eram menores ou iguais ao desvio

padrão de q_1 , a fim de comprovar se as mesmas passariam ou não a ser estatisticamente semelhantes.

Essa análise foi importante para que se pudesse prosseguir na aplicação dos modelos de regressão logística. Segundo Fernandes et al. (2020), o modelo de regressão logística adota variáveis dependentes categóricas binárias para identificação do evento de interesse, sendo uma metodologia que permite estimar a probabilidade associada à sua ocorrência. Os mesmos autores explicam que um evento de interesse pode ser classificado em “1” e a sua ausência como “0”, de tal modo que a regressão logística pode ser interpretada como um caso particular de modelos lineares generalizados em que a variável dependente é dicotômica.

Neste estudo, estimativas que assumiram o valor 1 foram enquadrados em uma condição em que q_{d7-17} e q_1 apresentaram discrepância de comportamento entre si, devendo receber alguma “atenção”, e aquelas que assumiram o valor 0 foram considerados como dados que apresentaram comportamento “semelhante” entre si.

Para a implementação do modelo de regressão logística é necessário a adoção de alguns critérios preliminares que viabilizem a sua aplicação e, sendo assim, o primeiro estágio tratou da imposição de um limite de decisão, cuja função foi permitir que o algoritmo fosse capaz de separar as duas classes. Em seguida procedeu-se à estimativa dos parâmetros da regressão logística e o ajuste do modelo através da seleção das variáveis independentes, utilizando duas formas diferentes de amostragem: aleatória e estratificada.

O valor do limite de decisão necessário para o ajuste do modelo de regressão logística foi baseado, inicialmente, no valor do desvio padrão de q_1 , utilizado anteriormente para filtragem das séries de vazão quando se aplicou o teste F.

Para a estimativa dos parâmetros da regressão logística é utilizado a função de máxima verossimilhança, que segundo King e Zeng (2001), é conhecida por sofrer algum viés para pequenas amostras. O grau do viés é fortemente dependente do número de casos menos frequente nas duas categorias. Quando o número de casos de interesse é inferior a 5% do total de amostras o modelo pode enquadrá-los como fenômeno raro, sendo necessário aplicar correções específicas para lidar com a situação buscando, assim, diminuir o viés (King e Zeng, 2001; Fernandes et al. 2020). Por esse motivo é que se optou por aplicar o método de máxima verossimilhança penalizada proposto por Firth (1993), cujo objetivo é produzir estimativas finas para os parâmetros do modelo, introduzindo um pequeno viés na função escore.

Ao se proceder às análises foi verificado que o limite de decisão adotado inicialmente para que o ajuste do modelo de regressão pudesse ser realizado foi insuficiente para que a função de máxima verossimilhança (que estima os parâmetros dos modelos) conseguisse

convergir. Em função disso, um novo critério foi adotado para aumentar o número de casos relacionados às vazões discrepantes, passando a ser considerado, então, dois terços (2/3) do desvio padrão de q_1 . Assim, houve um pequeno aumento das vazões discrepantes (de 19 para 46), que embora aparentemente não apresente grande expressividade, foi suficiente para que o modelo pudesse convergir.

Posteriormente foram verificados os requisitos técnicos para a modelagem da regressão logística, onde foram determinados o número de variáveis e observações a serem testadas. Como a regressão logística é muito sensível a multicolinearidade nas variáveis independentes, aquelas covariáveis que apresentaram coeficiente de correlação de Pearson maior ou igual a 0,95 foram eliminadas, usando ferramentas do pacote **caret**.

A Tabela 1 apresenta a quantidade de variáveis independentes testadas, destacando aquelas efetivamente utilizadas e aquelas descartadas na modelagem após o pré-processamento.

Tabela 1 – Variáveis independentes testadas e selecionadas/descartadas para aplicação do modelo de regressão logística com variável dependente dicotômica binária (0 ou 1)

<i>Variáveis Independentes</i>	<i>Representação</i>	<i>Situação</i>
Vazão específica média diária obtida a partir da estação fluviométrica 42460000 da ANA, considerando duas leituras.	q_1	Descartada
Vazão específica média diária defluente ao reservatório UHE Queimado obtida considerando 24 leituras.	q_d	Descartada
Vazão específica média diária defluente ao reservatório UHE Queimado obtida considerando 2 leituras (7 e 17 horas)	$q_{d_{7,17}}$	Utilizada
Dia do ano juliano no qual a observação foi registrada	d	Utilizada
Mês do ano juliano na qual a observação foi registrada	m	Descartada
Diferença entre q_1 e $q_{d_{7,17}}$	Dif	Utilizada

Afim de garantir que a seleção dos dados amostrais para o treinamento dos modelos fosse feita com maior representatividade, levando em conta especialmente os períodos mais chuvosos do ano, onde provavelmente ocorrem as extrapolações da curva-chave, os conjuntos de dados foram divididos nos procedimentos de aprendizado de máquina para as estratégias de treinamento, onde os hiperparâmetros do modelo de regressão logística penalizada são ajustados; e teste, onde o melhor modelo ajustado é usado para classificar um conjunto de dados não conhecidos por ele, de maneira a conter todos os meses do ano para que não houvesse riscos de ajuste de um modelo de regressão tendencioso.

Sendo assim, foram adotadas duas estratégias de divisão dos dados: uma realizando a amostragem estratificada, visando conter a mesma distribuição dos dados, com valores de vazão

entre as menores e as maiores magnitudes registradas nas séries, nos conjuntos de treinamento e teste; e a segunda estratégia realizando a amostragem aleatória, ambas com 70% da quantidade de dados para o treinamento e 30% para o teste, divisão também adotada por Gharib e Davis (2021). A divisão do conjunto de dados nas duas abordagens foi realizada usando recursos do pacote **rsample** (Silge et al., 2021).

As métricas utilizadas para avaliar os modelos foram a acurácia, que representa a proporção de casos que foram corretamente previstos; e o coeficiente Kappa, que fornece um valor numérico para comparar o modelo treinado com o modelo aleatório (Czodrowki, 2014).

O coeficiente Kappa avalia o grau de concordância entre a classificação e o valor real de uma mesma amostra e seu valor pode variar de -1 a 1. De acordo com Al Kafy et al. (2021), um coeficiente Kappa maior que 0,61 pode ser considerado um bom ajuste.

Foi avaliado, ainda, a importância de cada uma das variáveis utilizadas na modelagem usando a função **vip**, do pacote **vip**. Nesse caso, a função analisou o desempenho do modelo (Acurácia e Kappa), quando as variáveis foram, através de permutação, omitidas.

Posteriormente, o modelo de melhor ajuste para a regressão logística foi adotado e utilizado para classificação das vazões a partir de um conjunto de dados nunca antes trabalhado pelo modelo (conjunto de teste) e, assim, pôde-se produzir a matriz de contingência, onde o layout de uma tabela é construído para visualização do desempenho do algoritmo, permitindo que os valores de vazão, cuja classificação é atribuída com o valor 0 (“semelhantes”), fossem filtrados.

Feita a classificação no conjunto de teste foi aplicado novamente o teste F de Graybill para verificar se as observações classificadas pelo modelo com valor “0”, presentes nas séries de vazão, poderiam ser considerados como estatisticamente semelhantes.

Todo o processo de análise, visualização e modelagem foi realizado com recursos da linguagem de programação R (R Core Team, 2021), usando o software RStudio. A modelagem, treinamento, validação e teste dos modelos foi realizada usando recursos dos pacotes **caret** (Kuhn, 2020) e **glmnet** (Friedman, Hastie e Tibshirami, 2010), que possui uma implementação de Regressão Logística Penalizada.

Por fim, para os dados cuja classificação foi atribuída com o valor 1 (“atenção”) foram procedidas análises em associação com a curva-chave e os dados de chuva obtidos na estação pluviométrica (01647008), próxima ao local onde se encontra posicionada a estação pluviométrica Q1, objetivando melhorar a compreensão acerca destes comportamentos.

3.3. RESULTADOS E DISCUSSÃO

3.3.1. Análise do comportamento horário das vazões defluentes ao reservatório UHE Queimado

Os resultados obtidos a partir da obtenção das médias horárias mensais para a vazões defluentes ao reservatório do UHE Queimado, para o período chuvoso (novembro a abril) da série histórica relativa ao período de 2006 a 2019, são apresentados na Figura 2.

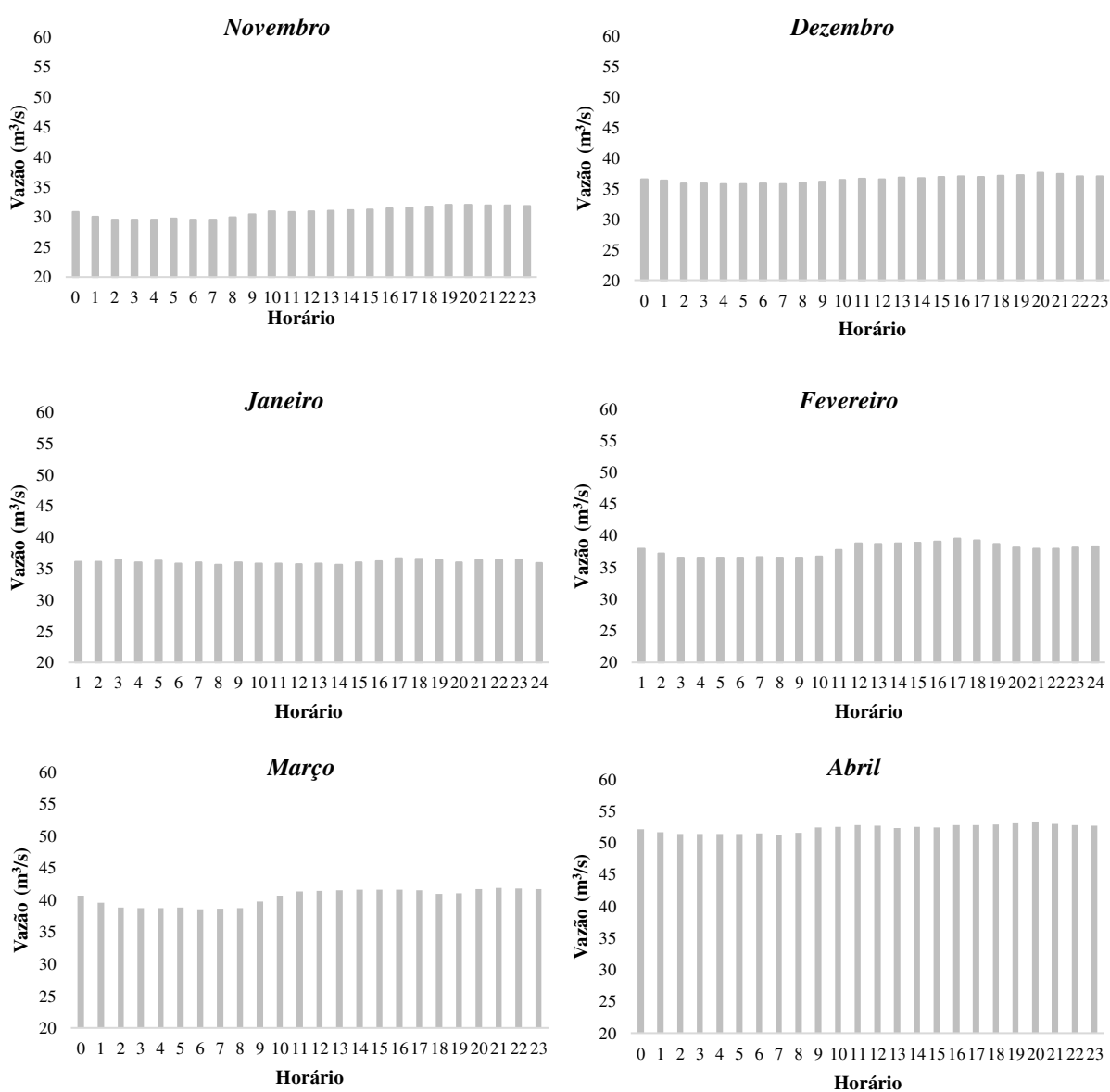
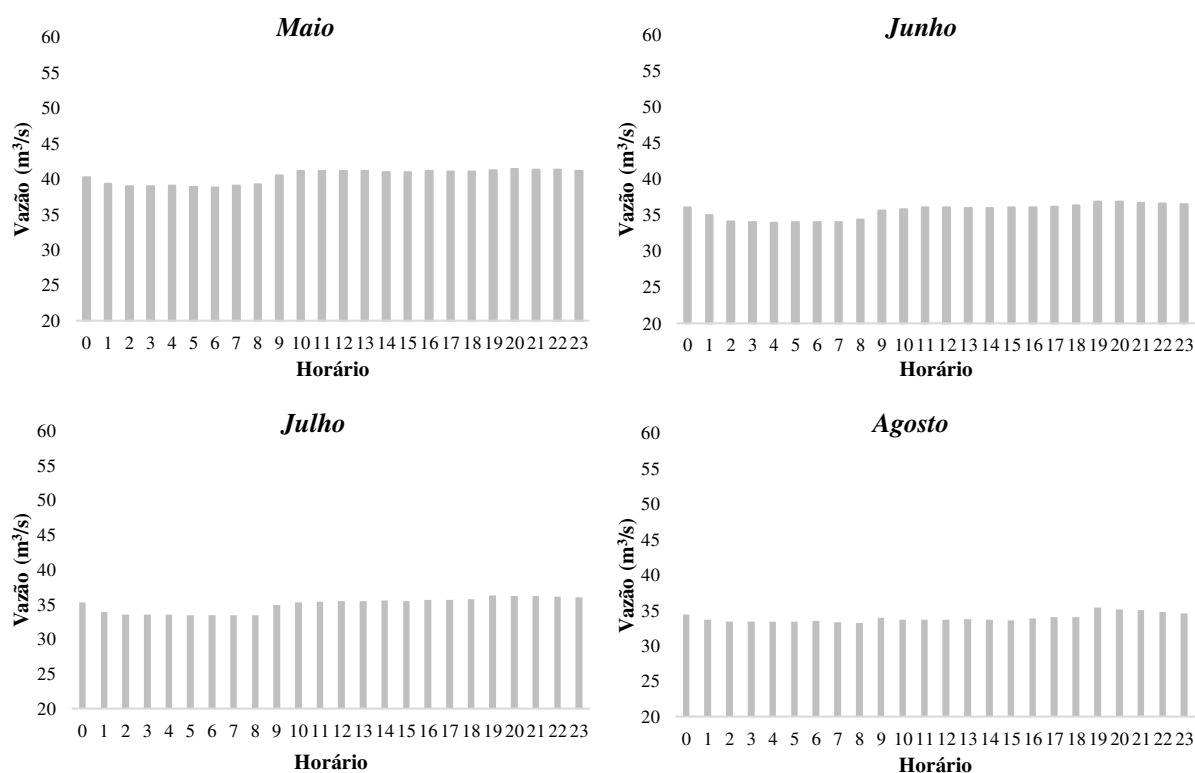


Figura 2 – Variação da vazão média horária defluente do reservatório UHE Queimado, observada ao longo do dia, considerando o período chuvoso (novembro a abril) da bacia do rio Preto.

É possível verificar na Figura 2 que, de maneira geral, as vazões defluentes não apresenta variações horárias bruscas para o período chuvoso (novembro a abril), o que significa que a produção de energia elétrica praticamente se mantém constante ao longo de todo o dia neste período. Entretanto, pode-se notar um ligeiro aumento na magnitude média das vazões, em alguns meses (novembro, fevereiro e março), nos períodos vespertino e noturno em relação ao matutino, mas ainda assim com diferenças médias da ordem de 1 a, no máximo, 2 m³/s, o que é irrelevante (2,5 a 5%) considerando a magnitude das vazões regularizadas a jusante (aproximadamente 40 m³/s). A Figura 3 apresenta a variação das vazões horárias a jusante do reservatório UHE Queimado no período de estiagem (maio a outubro), podendo-se verificar um comportamento relativamente distinto em comparação ao da vazão no período chuvoso (Figura 4).



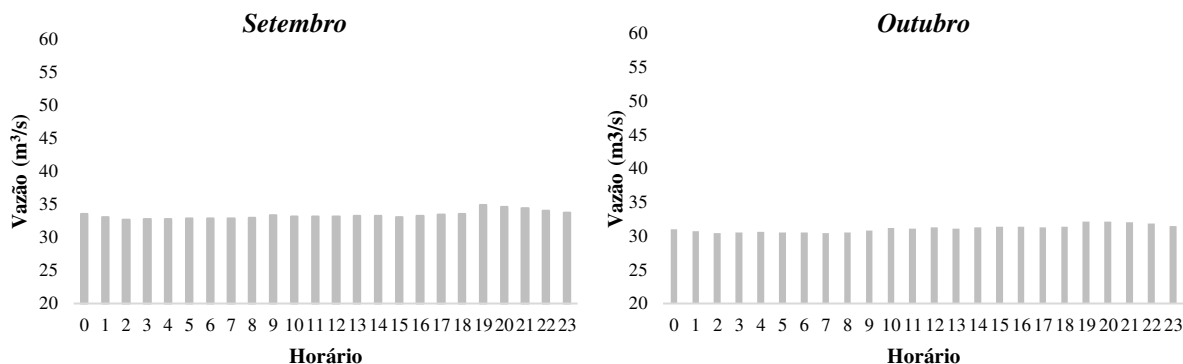


Figura 3 – Comportamento da vazão média defluente ao reservatório UHE Queimado, observada para as 24 horas do dia, considerando o período de estiagem (maio a outubro) na região.

Quando se analisa mais criteriosamente o comportamento das vazões horárias no período seco (Figura 3) percebe-se leves variações entre os diferentes turnos do dia, a exemplo dos meses de maio (maiores valores entre 10 e 0 horas), junho, julho, agosto, setembro e outubro (maiores valores a partir das 19 horas).

Os menores valores de vazão ocorrem quase que predominantemente no período matutino, porém apesar das variações identificadas no período seco, a maior diferença média entre a maior e a menor vazão horária não ultrapassa o valor de $4 \text{ m}^3/\text{s}$. Sendo assim, apesar das maiores diferenças nos valores de vazão horárias terem sido encontradas, aproximadamente entre 7 e 19 horas, uma avaliação comparativa do comportamento da vazão defluente média de 24 horas e dos valores de vazão defluentes calculados a partir da média de 7 e 17 horas, como é feito atualmente pela ANA, permitiram constatar que quase não existem diferenças nos valores para os dois intervalos de tempo, o que permite dizer que os resultados apresentam boa representatividade quanto ao comportamento das vazões quando se utiliza a média dos valores obtidos às 7 e 17 horas.

A título de exemplo, a Figura 4 comprova que, excetuando alguns pontos isolados, não existem diferenças consideráveis no comportamento das vazões defluentes ao reservatório quando se considera as médias de 24 horas de leitura ou a média das 7 e 17 horas, justamente em função da mudança pouco considerável do comportamento das vazões ao longo do dia na UHE Queimado.

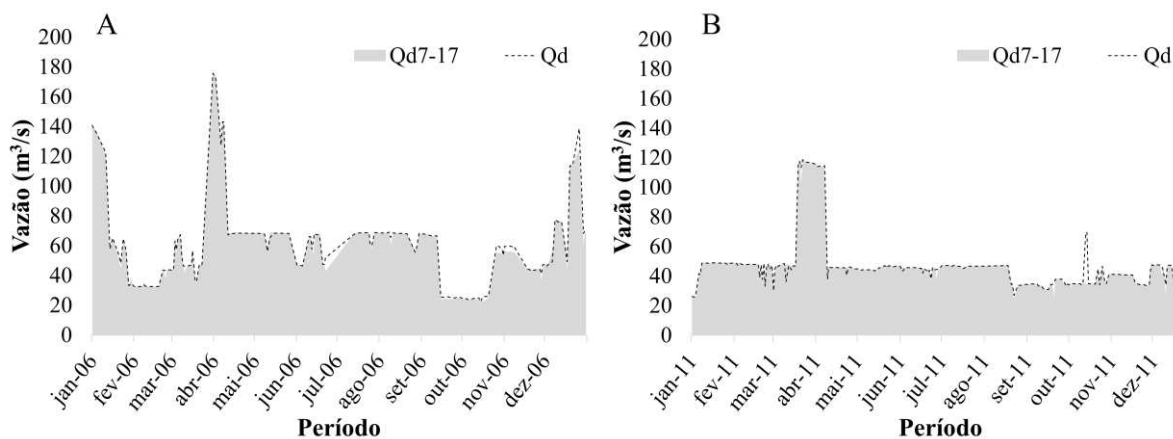


Figura 4 – Comparação do comportamento das vazões defluentes ao reservatório UHE Queimado quando considerada as estimativas diárias a partir das médias de 24 leituras (Qd) e obtidas com base na média de duas leituras às 7 e 17 horas (Qd₇₋₁₇) para os anos de 2006 (A) e 2011 (B).

Os resultados obtidos indicam que não ocorrerão problemas expressivos em relação à estimativa das vazões nas estações da ANA localizadas a jusante do reservatório da UHE Queimado em função dos horários de observação (7 e 17 horas). Porém, é importante avaliar se diferenças podem ser encontradas em função da metodologia adotada para quantificação das vazões na RHN, uma vez que o método utilizado para estimativa das vazões é diferente daquela adotada para a obtenção das vazões defluentes aos reservatórios de acumulação de água.

3.3.2. Comparação dos métodos para estimativa de vazão

A Figura 5 apresenta o gráfico de normalidade obtido pelo qq-plot onde foi possível analisar os dados de vazão estimados pelo método do balanço hídrico e os da curva-chave, verificando que há presença de dispersões entre os valores, principalmente quando estes ultrapassam cerca de 30 L/s.km².

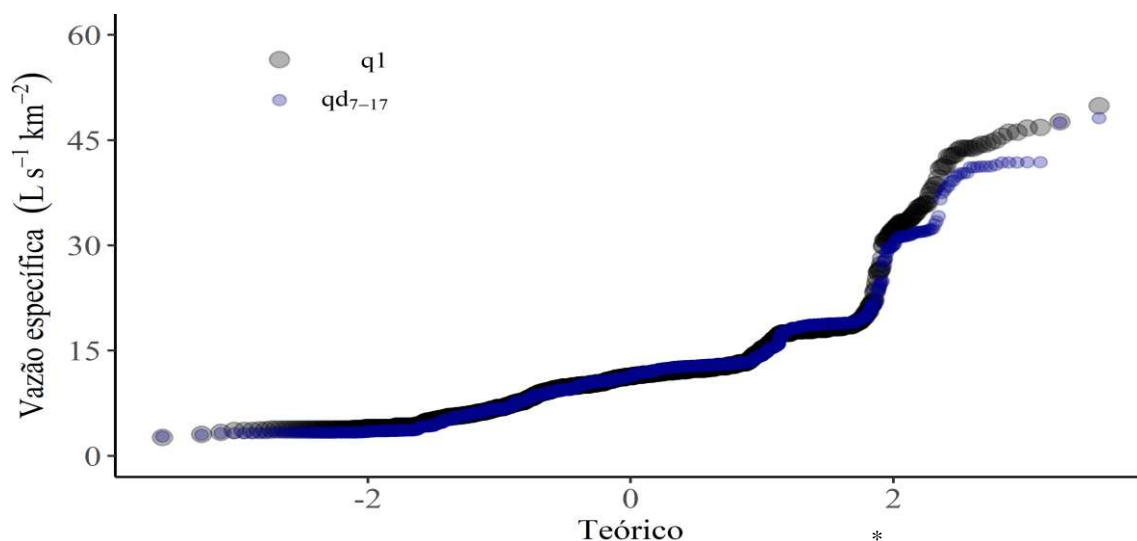


Figura 5 – Gráfico de normalidade de distribuição das vazões qd_{7-17} e q_1 para o período de 2006 a 2014.

Aplicando o teste F nas séries de dados constatou-se que as séries originais de qd_{7-17} e q_1 , para o período de análise considerado, foram caracterizadas como estatisticamente diferentes, de modo que, aplicando o teste F novamente, considerando apenas as observações em que o módulo da diferença entre q_1 e qd_{7-17} eram menores ou iguais ao desvio padrão de q_1 , pôde-se comprovar que as mesmas passaram a ser estatisticamente semelhantes, como mostra a Figura 6.

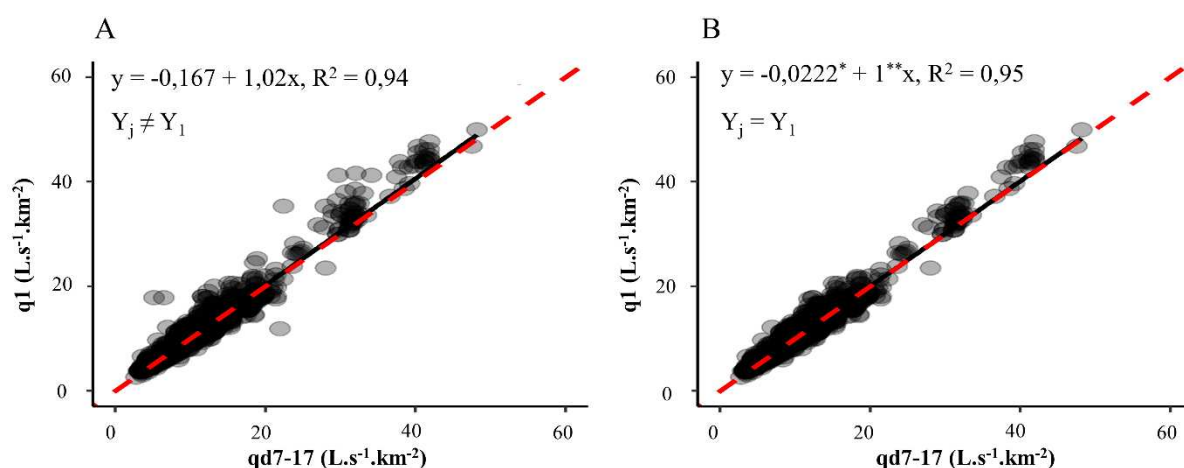


Figura 6 - Gráficos de dispersão entre qd_{7-17} e q_1 (x, y) resultantes do teste F de Graybill ao nível de significância de 1%, em que: (A) representa as séries originais estatisticamente diferentes e (B) representa as séries após ser aplicado o critério de filtragem das amostras pelo desvio padrão, caracterizadas como estatisticamente iguais ($*\beta_0 = 0$, $**\beta_1 = 1$).

Na Figura 6A há 2766 observações e na Figura 6B o número de observações é de 2747. Isso significa que 19 observações foram identificadas como discrepantes entre as séries,

causando alterações quanto ao grau de concordância entre elas, já que em (A) as séries são estatisticamente diferentes e em (B) passaram a ser estatisticamente iguais.

Após as verificações apresentadas, os modelos ajustados de regressão logística via amostragens estratificada e aleatória alcançaram, conforme apresentado na Tabela 2, bons resultados tanto para a fase de treinamento quanto para a de validação.

Tabela 2 - Índices estatísticos obtidos no treinamento e no teste do modelo de regressão logística utilizando amostragens estratificada e aleatória

Etapas	Amostragem	Índices estatísticos	
		Acurácia	Kappa
Treinamento	Estratificada	0,994	0,817
	Aleatória	0,995	0,789
Teste	Estratificada	0,992	0,621
	Aleatória	0,996	0,901

Verifica-se na Tabela 2 que os valores de acurácia do modelo, considerando as duas amostragens, alcançaram percentuais próximos a 100%, o que é um excelente indicativo. Em relação aos valores do Kappa, percebe-se que ocorreram reduções do treinamento para a fase de validação na amostragem estratificada. A acentuada redução neste tipo de amostragem indica a presença de overfitting e viés. Por outro lado, na modelagem utilizando amostragem aleatória o índice Kappa subiu, indicando capacidade de generalização do modelo nessa abordagem. Destaca-se, ainda, que os valores de Kappa encontrados para os dois modelos são maiores que 0,61, sendo classificados como bons (Al Kafy et al., 2021).

Procurando verificar se a modelagem aplicada seria capaz de identificar as observações classificadas como de “atenção” e assegurar a igualdade estatística entre as vazões específicas q_{d7-17} e q_1 , se aplicou novamente o teste F de Graybill aos dados classificados como “0” na regressão logística, tendo sido constatada a eficácia da metodologia utilizando as duas amostragens, como mostra a Figura 7.

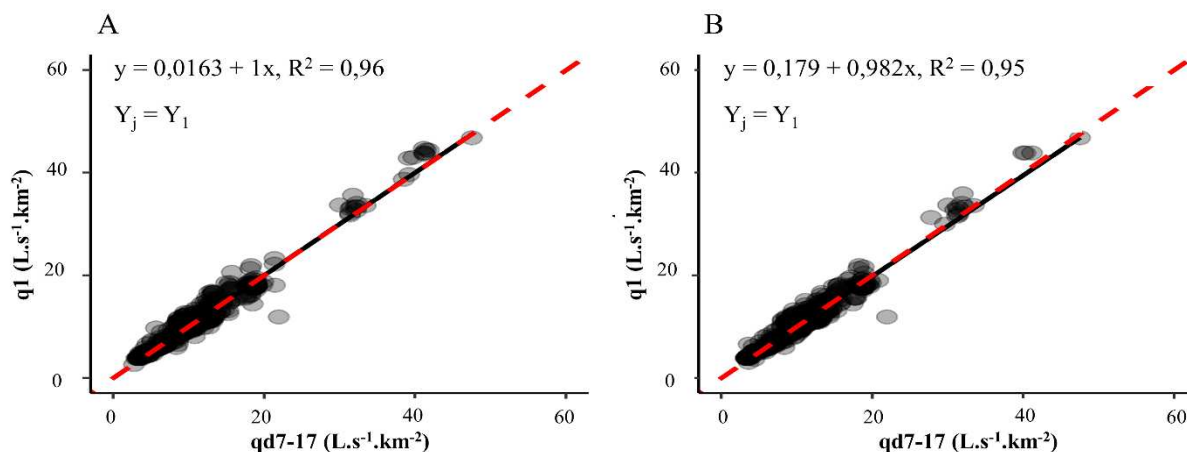


Figura 7 – Gráficos de igualdade estatística entre os métodos da curva-chave (q_1) e do balanço hídrico (q_{d7-17}) a partir da identificação e remoção das observações classificadas como de “atenção”, ao nível de significância de 1%, considerando as amostragens estratificada (A) e aleatória (B).

Além da igualdade estatística verificada para as séries selecionadas, pode-se observar que os valores dos coeficientes de determinação foram iguais a 0,96 e 0,95 nas amostragens estratificada e aleatória, respectivamente, indicando a existência de alta correlação entre os valores de vazão obtidos pelas duas metodologias.

Na fase de treinamento, através do teste para avaliação da importância de cada uma das variáveis utilizadas no modelo de regressão logística, tanto para a amostragem estratificada quanto para a aleatória, indicou que a variável relacionada ao dia do ano (d) foi a que apresentou menor importância, de modo que a sua remoção provocaria a menor perda de desempenho na modelagem em relação às demais variáveis testadas, o que não significa que sua presença não seja capaz de melhorar o desempenho do modelo.

Já a variável relacionada à diferença (Dif) entre q_1 e $q_{d7,17}$ foi a de maior importância para a modelagem, sendo que, aquelas diferenças que provocaram maior influência no processo de decisão classificatória do modelo (códigos 0 ou 1) ocorreram para valores de vazão associados às maiores magnitudes, isto é, no período chuvoso.

Ao avaliar isoladamente apenas aqueles dados que foram considerados como “atenção” pela regressão logística, foram quantificados para o conjunto teste 14 casos quando se considerou a amostragem aleatória e 5 casos quando se utilizou a amostragem estratificada.

A Figura 8 deixa evidente os casos encontrados pelo modelo de regressão logística com classificação “1”, ou seja, “atenção”, de modo que a maior parte das incompatibilidades

encontradas para os valores de vazão ocorreram predominantemente no período chuvoso, isto é, entre novembro e abril.

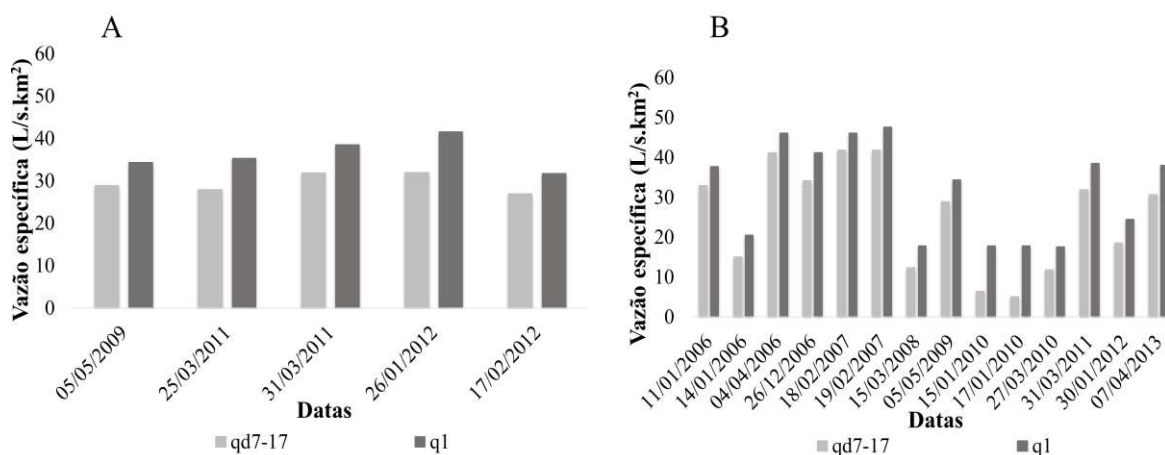


Figura 8 – Observações classificadas pelo modelo de regressão logística como de “atenção” considerando as amostragens estratificada (A) e aleatória (B).

Os maiores valores de qd_{7-17} e $q1$ encontrados para o período de 2006 a 2014 foram estimados em 48,2 e 49,9 L/s.km², respectivamente, representando vazões correspondentes a 176,2 e 194,2 m³/s. Assim sendo, é possível evidenciar, que quando a regressão logística foi realizada utilizando a amostragem aleatória o modelo identificou situações de “atenção” em $q1$ para vazões pequenas, ou seja, incompatibilidade entre as metodologias da curva-chave e do balanço hídrico mesmo em situações em que a magnitude dos valores é considerada baixa, isto é, conforme observado na Figura 5, menor que 30 L/s.km² e próximas da média de 11,7 L/s.km². Assim, estes valores, muito provavelmente, estão posicionados abaixo da faixa superior de extrapolação da curva-chave.

Isso não significa que tenha havido erro na modelagem, já que podem existir outros motivos pelos quais as diferenças entre as estimativas podem ter ocorrido (exemplo de erros de observação ou alteração na seção do escoamento), dado que a variável de maior importância para o modelo foi a diferença entre qd_{7-17} e $q1$. O modelo chama a atenção para vazões que não apresentam comportamento semelhante entre si e cabe ao analista verificar o motivo da ocorrência de tal situação, podendo ser ou não um dado incorretamente estimado.

A modelagem feita utilizando amostragem aleatória foi mais sensível às diferenças existentes, a partir de um limite pré-estabelecido entre as estimativas nos dois trechos analisados, do que quando se utilizou a amostragem estratificada. Para esta última os dados de “atenção” foram encontrados apenas para vazões específicas acima de 30 L/s.km². Assim, a capacidade de predição e o comportamento de um modelo depende significativamente do conjunto amostral de dados selecionados para a realização do seu treinamento.

Buscando refinar ainda mais as análises, foi avaliado se as discrepâncias encontradas entre as vazões q_{d7-17} e q_1 poderiam ou não estar associadas à ocorrência de precipitações (Figura 9).

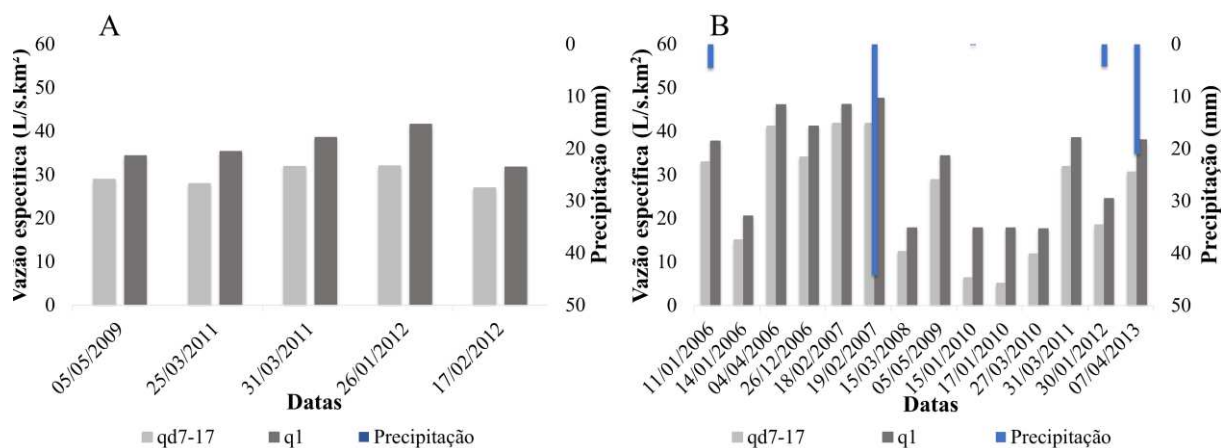


Figura 9 – Comparação do comportamento das vazões q_{d7-17} e q_1 com os dados de precipitação para as amostragens estratificada (A) e aleatória (B).

Verifica-se que, de maneira geral, a precipitação não pode ser considerada, como o principal fator relacionado à discrepância entre os valores de vazão diárias entre os trechos (Figura 9), uma vez que na maioria dos casos identificados não houve a incidência de chuvas na região que pudesse influenciar nessas alterações.

Contudo, na modelagem em que se utilizou a amostragem aleatória (Figura 9B), em dois dos 14 dias classificados com valores de vazão incompatíveis pode ter ocorrido alguma influência da precipitação em decorrência da magnitude dos valores registrados (dias 19/02/2007 e 07/04/2013), mas ainda assim não se pode afirmar com certeza, já que não se dispõe dos valores de precipitação na área de drenagem a montante do local onde os dados de q_{d7-17} foram coletados para que se pudesse fazer uma comparação mais adequada.

Por fim, ao analisar os dados de vazão efetivamente medidos em campo para fins de calibração da curva-chave na seção onde se encontra localizada a estação fluviométrica convencional da ANA durante o período analisado neste trabalho (Figura 10), verifica-se que os valores de vazão medidos não ultrapassaram $90 m^3/s$, ou seja, algo em torno de $24 L/s.km^2$, enquadrando os valores acima de $30 L/s.km^2$ dentro de uma zona de extrapolação e, portanto, com maiores probabilidades de erros de estimativa.

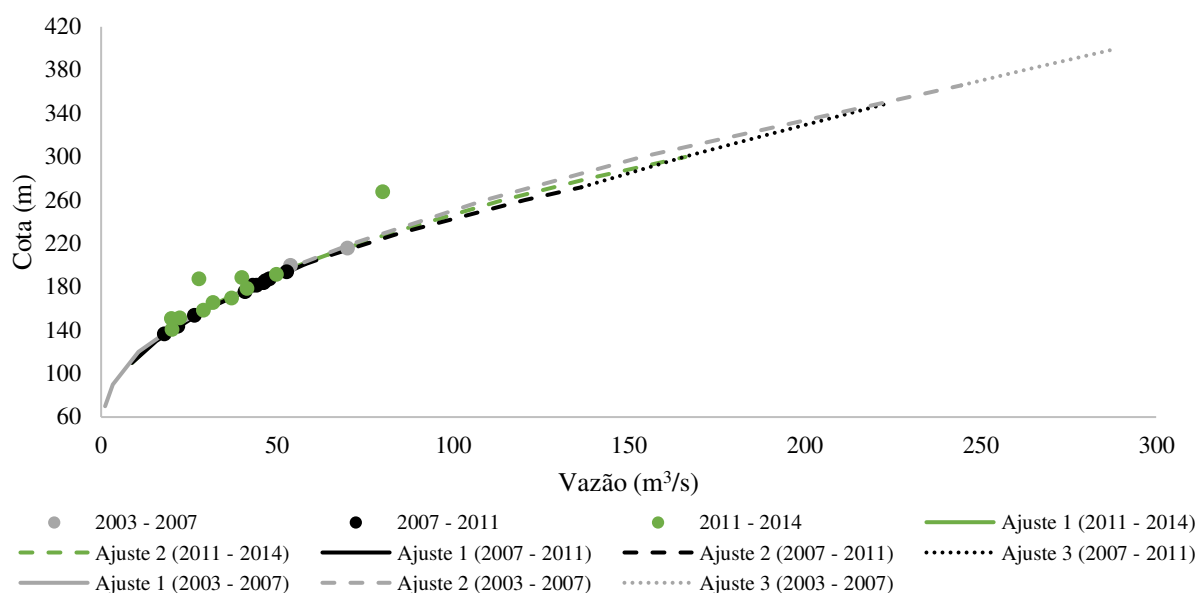


Figura 10 – Medições de cota e vazão realizadas na seção transversal da estação fluviométrica da ANA para produção das curvas-chaves válidas entre os períodos de 2003 a 2007, 2007 a 2011 e 2011 a 2014, utilizadas para estimativa de vazão.

Diante das análises realizadas, constatou-se que a metodologia que utiliza o método da curva-chave apresentou, para a maioria dos dados de vazão coletados, estimativas muito semelhantes aos valores de vazão estimados pelo método do balanço hídrico, o que reforça a eficácia do método para a obtenção dessa medida de escoamento.

Apesar disso, foram evidenciadas, ainda que em poucos casos, estimativas de vazões discrepantes entre os métodos da curva-chave e do balanço hídrico, comprovadas pelas diferenças estatísticas entre as séries e, dessa forma, passíveis de correção (Figura 8).

Sendo assim, a metodologia proposta por este trabalho apresentou potencialidade para a identificação de dados que podem apresentar inconsistências relacionadas tanto a problemas decorrentes da extrapolação da curva-chave, assumidos como predominantes nesse estudo, quanto por quaisquer outros fatores que devem ser cautelosamente checados.

O grande diferencial da metodologia proposta é tornar o processo de consistência das estimativas de vazões pelo método da curva-chave menos trabalhoso e mais seguro, desde que exista a viabilidade de se comparar os dados a um método padrão, que apresente menor variabilidade dos parâmetros hidráulicos, como o balanço hídrico. Assim, o processo de correções das informações hidrológicas inconsistentes ou incoerentes poderá ocorrer de forma menos trabalhosa e subjetiva.

3.4. CONCLUSÕES

- A variação das vazões horárias defluentes ao reservatório UHE Queimado apresenta baixa variabilidade diária ao longo dos meses do ano, sendo que no período seco, ocorrem a maior amplitude dos dados entre o período matutino e noturno.
- Em função da reduzida variabilidade horária encontrada para a vazão diária defluente do reservatório UHE Queimado, ficou evidente que os dados estimados a partir da média entre essas duas leituras diárias (estações convencionais da RHN) apresentam boa representatividade do comportamento do escoamento nessa região.
- Os valores de vazão diárias estimados pelo método do balanço hídrico (vazão defluente ao reservatório) e pelo método da curva-chave (vazão na estação fluviométrica da ANA, localizada a jusante) apresentam estimativas estatisticamente semelhantes para as vazões mínimas e médias, evidenciando uma boa compatibilidade entre os métodos.
- As maiores diferenças encontradas na estimativa de vazões pelos métodos do balanço hídrico e da curva-chave ocorreram, em sua maioria, no período chuvoso, onde há a ocorrência de vazões máximas. As diferenças encontradas nesse estudo foram associadas à aplicação de métodos para extrapolação da curva-chave, já que para a sua calibração não foram verificados valores medidos de vazão na seção do curso d'água superiores a 24 L/s.km².
- A aplicação do método de regressão logística foi capaz de identificar dados atípicos nas séries históricas de vazões analisadas, considerando a comparação entre os métodos do balanço hídrico e da curva-chave, podendo ser usado como metodologia de consistência dos dados de vazões para situações em que existir a possibilidade de se comparar os dados com um método padrão.

3.5. REFERÊNCIAS

Braga, S. R., Oliveira, M. L. R. de., Gorgens, E. B. 2021. forestmangr: Forest Mensuration and Management (R package version 0.9.4). <https://cran.r-project.org/package=forestmangr>

- Cheng, Z., Lee, K., Kim, D., Muste, M., Vidmar, P., Hulme, J. 2019. Experimental evidence on the performance of rating curves for continuous discharge estimation in complex flow situations. *Journal of Hydrology*, p. 959-971, doi: <https://doi.org/10.1016/j.jhydrol.2018.11.021>
- Czodrowki, P. 2014. Count on kappa. *Journal of Comput-Aided Molecular Design*, v. 28, p. 1049-1055, doi: <https://doi.org/10.1007/s10822-014-9759-6>
- Fernandes, A. A. T., Figueiredo Filho, D. B., Rocha, E. C. da., Nascimento, W. da S. 2020. Leia este artigo se você quiser aprender regressão logística. *Revista de Sociologia e Política*, v. 28, n.74, doi: 10.1590/1678-987320287406
- Firth, D. 1993. Bias reduction of maximum likelihood estimates. *Biometrika*, v.80, n.1, p. 87-38.
- Friedman, J. H., Hastie, T., Tibshirani, R. 2010. Regularization Paths for Generalized Linear Models via Coordinate Descent. *Journal of Statistical Software*, 33(1), 1–22. <https://doi.org/10.18637/JSS.V033.I01>
- Gharib, A., Davies, E. G. R. 2021. A workflow to address pitfalls and challenges in applying machine learning models to hydrology. *Advances in Water Resources*, v. 152, doi: <https://doi.org/10.1016/j.jenvman.2021.113359>
- Graybill, F. A. 1976. *Theory and application of the linear model*. Belmont: Duxbury Press, 704p.
- Kafy, A.-A., Faisal, A.-A.-, Shuvo, R. M., Naim, M. N. H., Sikdar, M. S., Chowdhury, R. R., Islam, M. A.; Sarker, M. H. S.; Khan, M. H. H., Kona, M. A. 2021. Remote sensing approach to simulate the land use/land cover and seasonal land surface temperature change using machine learning algorithms in a fastest-growing megacity of Bangladesh. *Remote Sensing Applications: Society and Environment*, v. 21, doi:10.1016/j.rsase.2020.100463
- King, G., Zeng, L. 2001. Logistic regression in rare events data. *Political analysis*, v.9, p. 137-163, doi: 10.1093/oxfordjournals.pan.a004868
- Kuhn, M. 2020. caret: Classification and Regression Training (6.0-86). <https://cran.r-project.org/package=caret>
- Lessa, A. C. R.; Santos, M. A. dos.; Maddock, J. E. L., Bezerra, C. dos S. 2015. Emissions of greenhouse gases in terrestrial areas pre-existing to hydroelectric plant reservoirs in the Amazon: The case of Belo Monte hydroelectric plant. *Renewable and Sustainable Energy Reviews*, v. 51, 1728-1736, doi: <https://doi.org/10.1016/j.rser.2015.07.067>
- Manfreda, S. 2018. On the derivation of flow rating curves in data-scarce environments. *Journal of Hydrology*, v. 562, p. 151-154, doi: <https://doi.org/10.1016/j.jhydrol.2018.04.058>
- Molina, Z. B. de la C. 2016. Avaliação da acurácia do modelo de simulação a usinas individualizadas baseadas no pós-operativo. Dissertação de Mestrado – Pós-Graduação em Engenharia Elétrica, Universidade Estadual de Campinas.

- Nobre, J. P. G., Almeida, H. C. de, Ribeiro, L. E., Silva, F. D. dos S., Costa, R. L. 2017. Estimativa da produtividade do milho nas diferentes mesorregiões do Estado de Alagoas. *Revista Brasileira de Geografia Física*, v.10, n.14, p. 1100-1109.
- Peña-Arancibia, J. L., Zhang, Y., Pagendm, D. E., Viney, N. R., Lerat, J., Van Dijk, A. I. J. M., Vaze, J., Frost, A. J. 2015. Streamflow rating uncertainty: characterization and impacts on model calibration and performance. *Environ. Model. Softw*, v. 63, p. 32-44, doi: <https://doi.org/10.1016/j.envsoft.2014.09.011>
- R Core Team (2021). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>
- Silge, J., Chow, F., Kuhn, M., Wickham, H. (2021). *rsample: General Resampling Infrastructure* (R package version 0.1.0). <https://cran.r-project.org/package=rsample>
- Wannasin, C., Brauner, C. C., Uijlenhoet, R., Van Verseveld, W. J., e Weerts, A. H. 2021. Daily flow simulation in Thailand Part II: Unraveling effects of reservoir operation. *Journal of Hydrology: Regional Studies*, v.34. doi: <https://doi.org/10.1016/j.ejrh.2021.100792>
- Wilk, M. B., Gnanadesikan, R. (1968). Probability plotting methods for the analysis for the analysis of data. *Biometrika*, v. 55, n.1, p.1–17. doi:10.1093/biomet/55.1.1