**UNIVERSIDADE FEDERAL DE VIÇOSA**

# Uso de técnicas de Machine Learning para geração de séries históricas de vazões na bacia do rio Doce

Anderson de Oliveira Pinheiro
*Magister Scientiae*

**VIÇOSA - MINAS GERAIS**
**2026**

**ANDERSON DE OLIVEIRA PINHEIRO**

# Uso de técnicas de Machine Learning para geração de séries históricas de vazões na bacia do rio Doce

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

Orientador: Demetrius David da Silva

Coorientador: Michel Castro Moreira

**VIÇOSA - MINAS GERAIS**
**2026**

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

T
P654u
2026

Pinheiro, Anderson de Oliveira, 1998-
Uso de técnicas de machine learning para geração de séries históricas de vazão na bacia do Rio Doce / Anderson de Oliveira Pinheiro. – Viçosa, MG, 2026.
1 dissertação eletrônica (91 f.): il. (algumas color.).

Orientador: Demetrius David da Silva.
Dissertação (mestrado) - Universidade Federal de Viçosa, Departamento de Engenharia Agrícola, 2026.
Referências bibliográficas: f. 74-91.
DOI: https://doi.org/10.47328/ufvbbt.2026.575
Modo de acesso: World Wide Web.

1. Estações fluviométricas - Doce, Rio, Bacia (MG e ES). 2. Medidores de fluxo - Modelos matemáticos. 3. Análise multivariada. 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 Pires CRB-6/2523

**ANDERSON DE OLIVEIRA PINHEIRO**

**Uso de técnicas de Machine Learning para geração de séries históricas de vazões na bacia do rio Doce**

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

APROVADA: 3 de março de 2026.

Assentimento:

_____

Anderson de Oliveira Pinheiro
Autor

_____

Demetrius David da Silva
Orientador

Essa dissertação foi assinada digitalmente pelo autor em 01/10/2026 às 10:00:08 e pelo orientador em 01/10/2026 às 13:26:56. As assinaturas têm validade legal, conforme o disposto na Medida Provisória 2.200-2/2001 e na Resolução nº 37/2012 do CONARQ. Para conferir a autenticidade, acesse https://siadoc.ufv.br/validar-documento. No campo 'Código de registro', informe o código **G83P.CTS4.W5UJ** e clique no botão 'Validar documento'.

## AGRADECIMENTOS

A realização deste trabalho representa muito mais do que a conclusão de uma etapa acadêmica. Foi um período de muito aprendizado, desafios e amadurecimento, no qual aprendi não apenas a fazer ciência, mas também a lidar com as dificuldades que fazem parte desse caminho. Essa conquista, porém, não foi construída de forma individual. Muitas pessoas contribuíram, de diferentes maneiras.

Dessa forma, agradeço, primeiramente, a Deus, por me dar força, coragem e sabedoria para enfrentar os desafios e seguir em frente, mesmo nos momentos mais difíceis.

Aos meus pais, Elza e Valdevino, por sempre acreditarem na educação e me incentivarem a estudar desde pequeno. Agradeço por todo o esforço para que eu pudesse realizar minha graduação, pelo apoio para continuar na pós-graduação e, principalmente, por estarem ao meu lado.

Aos meus familiares e amigos, em especial Aprigio, Hellen, Geane, Lucas Andrade e Lucas Guilherme, e a todos aqueles que estiveram presentes durante essa caminhada, agradeço pela amizade, pelas conversas, pela paciência nos momentos de dificuldade e pelas experiências compartilhadas. A companhia de vocês tornou esse período mais leve e deixou boas lembranças para além da vida acadêmica.

Um agradecimento especial à minha avó Maria Madalena, por todo o carinho, pelas orações e por sempre torcer por mim.

Aos meus orientadores, professor Demétrius e professor Michel, agradeço pela confiança, dedicação, paciência e disponibilidade ao longo dessa formação. Obrigado pelas orientações, pelas discussões, pelos ensinamentos e pelas contribuições não apenas para este trabalho, mas também em todos os demais trabalhos.

À Universidade Federal de Viçosa (UFV), pela oportunidade de realizar minha pós-graduação e por proporcionar um ambiente de aprendizado, pesquisa e desenvolvimento profissional.

Ao Programa de Pós-Graduação em Engenharia Agrícola e ao Departamento de Engenharia Agrícola (DEA), pelo suporte e pelas oportunidades oferecidas ao longo do curso.

Este trabalho foi realizado com o apoio das seguintes agências de pesquisa brasileiras: Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Código de Financiamento 001, Fundação de Amparo à Pesquisa do Estado de Minas Gerais (FAPEMIG) e Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq).

A todos que, direta ou indiretamente, fizeram parte dessa trajetória, meu muito obrigado.

## RESUMO

PINHEIRO, Anderson de Oliveira, M.Sc., Universidade Federal de Viçosa, março de 2026. **Uso de técnicas de Machine Learning para geração de séries históricas de vazões na bacia do rio Doce**. Orientador: Demetrius David da Silva. Coorientador: Michel Castro Moreira.

O conhecimento do regime de vazões é fundamental para o planejamento e a gestão dos recursos hídricos em bacias hidrográficas, porém falhas nas séries fluviométricas podem limitar sua utilização em estudos hidrológicos. Com isso, este trabalho teve como objetivo avaliar a capacidade dos modelos Random Forest (RF), Extreme Gradient Boosting (XGB), Multilayer Perceptron (MLP) e Long Short-Term Memory (LSTM) de gerar séries diárias de vazão em sete sub-bacias do rio Doce com diferentes características fisiográficas e sob diferentes períodos consecutivos de falha. Foram utilizados dados diários de precipitação e vazão disponibilizados pela Agência Nacional de Águas e Saneamento Básico (ANA), considerando como variáveis de entrada a precipitação do dia corrente e dos cinco dias antecedentes, além da vazão do dia antecedente. Foram simulados blocos consecutivos de 10, 30, 60, 90 e 120 dias de falha. Os modelos foram avaliados a partir de duas estratégias de treinamento: sequencial, na qual o treinamento foi atualizado ao longo da série e os valores previamente estimados passaram a compor o processo de preenchimento das falhas; e tradicional, baseada em um único treinamento para posterior aplicação no período de teste. A LSTM foi aplicada apenas pela estratégia sequencial. O desempenho dos modelos foi avaliado com base em métricas estatísticas (RMSE, sMAPE, PBIAS, $R^2$ e KGE) e hidrológicas (curva de permanência, vazões de referência e vazões médias mensais), considerando a capacidade de reproduzir adequadamente os diferentes componentes do regime de vazões. Os modelos apresentaram melhor desempenho nos blocos de 10 dias, com redução da capacidade de reconstrução à medida que a duração das falhas aumentou, embora as diferenças entre os períodos tenham se tornado menos evidentes a partir de 60 dias. Os modelos baseados em árvores treinados pela metodologia tradicional apresentaram os melhores resultados, com destaque para o XGB2, seguido pelo RF2, enquanto a LSTM apresentou o melhor desempenho entre os modelos treinados pela metodologia sequencial. De modo geral, observou-se maior dificuldade na representação das condições extremas do regime hidrológico, principalmente dos picos de vazão e das condições mais críticas de baixa vazão. A importância das covariáveis variou entre os modelos, com maior contribuição da vazão do dia anterior e das precipitações ocorridas um e

dois dias antes. Apesar das diferenças observadas entre as sub-bacias, os resultados apresentaram uma tendência semelhante em relação ao desempenho dos modelos. A estratégia de treinamento e a extensão das falhas tiveram influência na qualidade da reconstrução das séries, sendo observado um melhor desempenho dos modelos baseados em árvores quando utilizados na abordagem tradicional.

Palavras-chave: Regime fluviométrico; Modelos não lineares; Análise comparativa

## ABSTRACT

PINHEIRO, Anderson de Oliveira, M.Sc., Universidade Federal de Viçosa, March, 2026. **Use of Machine Learning Techniques for the Generation of Streamflow Time Series in the Doce River Basin**. Adviser: Demetrius David da Silva. Co-adviser: Michel Castro Moreira.

Knowledge of streamflow regimes is essential for water resources planning and management in river basins; however, gaps in streamflow time series may limit their use in hydrological studies. In this context, this study aimed to evaluate the ability of Random Forest (RF), Extreme Gradient Boosting (XGB), Multilayer Perceptron (MLP), and Long Short-Term Memory (LSTM) models to reconstruct daily streamflow time series in seven sub-basins of the Doce River with different physiographic characteristics and under different durations of consecutive data gaps. Daily precipitation and streamflow data provided by the Brazilian National Water and Basic Sanitation Agency (ANA) were used. The input variables included precipitation on the current day and the five preceding days, as well as streamflow from the previous day. Consecutive data gaps of 10, 30, 60, 90, and 120 days were simulated. The models were evaluated using two training strategies: a sequential approach, in which model training was updated throughout the time series and previously estimated values were incorporated into the gap-filling process; and a traditional approach, based on a single training procedure followed by application to the testing period. The LSTM model was applied only using the sequential strategy. Model performance was assessed using statistical metrics (RMSE, sMAPE, PBIAS, $R^2$, and KGE) and hydrological indicators (flow duration curves, reference streamflows, and mean monthly streamflows), considering the models' ability to adequately reproduce different components of the streamflow regime. The models performed best for 10-day data gaps, with a decline in reconstruction performance as gap duration increased, although differences among gap durations became less pronounced from 60 days onward. Tree-based models trained using the traditional approach showed the best overall performance, particularly XGB2, followed by RF2, whereas LSTM achieved the best performance among the models trained using the sequential approach. Overall, the models showed greater difficulty in representing extreme hydrological conditions, particularly peak flows and the most critical low-flow conditions. Covariate importance varied among models, with the previous day's streamflow and precipitation from one and two days earlier making the greatest contributions. Despite differences among the sub-basins, the results showed a similar pattern in terms of model performance. Both the training

strategy and gap duration influenced the quality of time-series reconstruction, with tree-based models showing better performance when implemented using the traditional approach.

Keywords: Streamflow regime; Nonlinear models; Comparative analysis

## LISTA DE FIGURAS

**Figura 1:** Mapa de localização da bacia do Rio Doce ..... 25
**Figura 2:** Bacias hidrográficas delimitadas para a modelagem hidrológica na área de estudo. ..... 28
**Figura 3:** Variação dos valores de RMSE das 10 repetições feitas em cada bacia e quantidade de dias seguidos de falhas. ..... 51
**Figura 4:** Variação dos valores de PBIAS das 10 repetições feitas em cada bacia e quantidade de dias seguidos de falhas ..... 52
**Figura 5:** Variação dos valores de SMAPE das 10 repetições feitas em cada bacia e quantidade de dias seguidos de falhas ..... 53
**Figura 6:** Variação dos valores de KGE das 10 repetições feitas em cada bacia e quantidade de dias seguidos de falhas ..... 54
**Figura 7:** Variação dos valores de $R^2$ das 10 repetições feitas em cada bacia e quantidade de dias seguidos de falhas ..... 55
**Figura 8:** Curva de permanência para todos os modelos em cada bacia e quantidade de falhas ..... 57
**Figura 9:** Razão entre os valores de $Q_{10}$ estimados e observados nos diferentes cenários de ausência de dados ..... 59
**Figura 10:** Razão entre os valores de $Q_{50}$ estimados e observados nos diferentes cenários de ausência de dados ..... 60
**Figura 11:** Razão entre os valores de $Q_{90}$ estimados e observados nos diferentes cenários de ausência de dados ..... 61
**Figura 12:** Razão entre os valores de $Q_{95}$ estimados e observados nos diferentes cenários de ausência de dados ..... 62
**Figura 13:** Razão entre os valores de $Q_{99}$ estimados e observados nos diferentes cenários de ausência de dados ..... 63
**Figura 14:** Vazões médias mensais observadas e estimadas por cada modelo de machine learning nos diferentes cenários de ausência de dados (10 a 120 dias) ..... 65
**Figura 15:** Comparação entre os resíduos das vazões observadas e estimadas por cada modelo de machine leanrnig nos diferentes cenários de ausência de dados, pelo teste de Friedman e Wilcoxon ..... 67
**Figura 16:** Hidrogramas das vazões observadas e simuladas pelos modelos XGB2, RF2 e LSTM, em cada cenário de ausência de dados, para a bacia BH1 ..... 69
**Figura 17:** Hidrogramas das vazões observadas e simuladas pelos modelos XGB2, RF2 e LSTM, em cada cenário de ausência de dados, para a bacia BH2 ..... 69
**Figura 18:** Hidrogramas das vazões observadas e simuladas pelos modelos XGB2, RF2 e LSTM, em cada cenário de ausência de dados, para a bacia BH3 ..... 70
**Figura 19:** Hidrogramas das vazões observadas e simuladas pelos modelos XGB2, RF2 e LSTM, em cada cenário de ausência de dados, para a bacia BH4 ..... 70
**Figura 20:** Hidrogramas das vazões observadas e simuladas pelos modelos XGB2, RF2 e LSTM, em cada cenário de ausência de dados, para a bacia BH5 ..... 71
**Figura 21:** Hidrogramas das vazões observadas e simuladas pelos modelos XGB2, RF2 e LSTM, em cada cenário de ausência de dados, para a bacia BH6 ..... 71
**Figura 22:** Hidrogramas das vazões observadas e simuladas pelos modelos XGB2, RF2 e LSTM, em cada cenário de ausência de dados, para a bacia BH7 ..... 72

**SUMÁRIO**


1. INTRODUÇÃO ..... 11
2. REVISÃO DE LITERATURA ..... 14
2.1. Bacia do Rio Doce ..... 14
2.2. Importância das vazões médias diárias na gestão hídrica ..... 15
2.3. Informações fluviométricas ..... 16
2.4. Modelagem hidrológica ..... 17
2.4.1. Random Forest (RF) ..... 19
2.4.2. Extreme Gradient Boosting – XGBoost (XGB) ..... 20
2.4.3. Multilayer Perceptron (MLP) ..... 21
2.4.4. Long Short-Term Memory (LSTM) ..... 21
2.5. Covariáveis importantes na modelagem Chuva-Vazão ..... 23
3. MATERIAIS E MÉTODOS ..... 24
3.1. Área de estudo ..... 24
3.2. Obtenção da base de dados ..... 25
3.3. Análise de Tendências nas séries vazões médias diárias ..... 28
3.3.1. Função de Autocorrelação (ACF) ..... 29
3.3.2. Mann-Kendall (MK) e Mann-Kendall Modificado (MKM) ..... 30
3.3.3. Correlação de Spearman ..... 32
3.3.4. Teste de Pettitt ..... 33
3.3.5. Período de treino e teste ..... 34
3.4. Pré-processamento dos dados de chuva vazão e definição das covariáveis ..... 35
3.4.1. Vazão ..... 35
3.4.2. Precipitação ..... 36
3.4.4. Normalização dos dados ..... 36
3.5. Modelos de Machine Learnig (ML) ..... 37
3.5.1. Random Forest (RF) ..... 40
3.5.2. Extreme Gradient Boosting (XGBoost – XGB) ..... 40
3.5.3. Multilayer Perceptron (MLP) ..... 41
3.5.4. Long Short-Term Memory ..... 41
3.6. Avaliação do Desempenho dos Modelos ..... 42
3.6.1. Análise das métricas de desempenho entre as repetições ..... 42
4. RESULTADOS E DISCUSSÃO ..... 49

4.1. Avaliação da influência das repetições feitas em cada bacia e quantidade de dias seguidos de falhas nas métricas de desempenho ..... 49
4.2. Importância das Covariáveis nos Modelos ..... 55
4.3. Avaliação das Curvas de permanência nos diferentes cenários de ausência de dados 56
4.3.1. Análise dos valores de $Q_{10}$, $Q_{50}$, $Q_{90}$, $Q_{95}$ e $Q_{99}$ ..... 57
4.4. Regime hidrológico mensal ..... 63
4.5. Comparação entre os modelos de Machine Learning ..... 65
4.6. Hidrogramas das vazões simuladas para os modelos XGB2, RF2 e LSTM ..... 67
5. CONCLUSÕES ..... 73
REFERÊNCIAS ..... 74

## 1. INTRODUÇÃO

A disponibilidade hídrica superficial em uma bacia hidrográfica é relacionada ao regime de vazões dos sistemas fluviais, responsáveis pelo atendimento de diversos usos consuntivos e não consuntivos, como abastecimento humano, irrigação, geração de energia, diluição de efluentes e manutenção dos ecossistemas aquáticos (Falkenmark; Lannerstad, 2005; Valdés-Pineda et al., 2014; Mwaguni et al., 2016). O conhecimento quantitativo da disponibilidade hídrica nos rios torna-se um elemento central para o planejamento e a gestão integrada dos recursos hídricos. Na bacia hidrográfica do Rio Doce, assim como em outras bacias de grande extensão, essa necessidade é evidente, considerando sua extensão territorial, diversidade de usos da terra e importância socioeconômica para os estados de Minas Gerais e Espírito Santo (CBH-Doce, 2026).

Entretanto, a obtenção de informações de vazão representativas de toda a bacia é limitada pela distribuição espacial das estações fluviométricas, uma vez que cada estação representa as condições hidrológicas de sua área de drenagem a montante. Além disso, falhas de equipamentos, dificuldades de manutenção e descontinuidades das séries podem resultar em registros incompletos, comprometendo análises hidrológicas e a gestão dos recursos hídricos (Fasipe; Izinyon, 2021; Wilby et al., 2017).

Diante dessas limitações, a estimativa de vazões a partir de variáveis hidrometeorológicas surge como uma alternativa para a obtenção de séries de vazão. Dentre essas variáveis, a precipitação destaca-se por sua relação direta com a geração do escoamento e com a variabilidade temporal das vazões (Xie et al., 2023). Na bacia do Rio Doce, as vazões apresentam sazonalidade associada ao regime de precipitação, com menores valores durante o inverno, caracterizado pelo período de estiagem, e maiores vazões no verão, quando ocorre maior concentração das chuvas (Cavalcanti, 2016).

Todavia, a conversão da precipitação em vazão representa um desafio, especialmente em bacias de grande extensão e elevada heterogeneidade. Uma mesma quantidade de precipitação pode produzir respostas hidrológicas distintas entre sub-bacias em razão de diferenças em suas características fisiográficas, como área de drenagem, relevo, declividade, características dos solos e cobertura e uso da terra, além das condições climáticas regionais (H. Liu et al., 2025; Y. Zhang et al., 2025). Essas características controlam processos como infiltração, armazenamento, evapotranspiração e propagação do escoamento, resultando em diferentes respostas de vazão para eventos de precipitação semelhantes (Sánchez et al., 2024). Dessa forma, a capacidade de um modelo em representar adequadamente a relação chuva–vazão pode variar entre bacias com diferentes características fisiográficas.

Nesse sentido, compreender e modelar os padrões não lineares que governam a relação chuva–vazão é fundamental para reconstruir séries de vazão e representar as diferentes respostas hidrológicas existentes em uma bacia hidrográfica (Herath et al., 2021). A modelagem chuva–vazão pode ser utilizada na avaliação de impactos ambientais, na análise da disponibilidade hídrica e no suporte à tomada de decisão na gestão dos recursos hídricos, principalmente em regiões com baixa disponibilidade de informações fluviométricas (Chagas et al., 2020).

Tradicionalmente, a modelagem chuva–vazão baseia-se em modelos físicos e conceituais. Os modelos físicos representam os processos hidrológicos por meio de princípios físicos, como a conservação da massa e da energia, enquanto os modelos conceituais utilizam representações matemáticas simplificadas desses processos e, geralmente, demandam menor quantidade de informações (Fathi et al., 2025).

Embora fundamentados em bases teóricas consistentes, esses modelos podem apresentar limitações na aplicação prática, relacionadas às simplificações adotadas na representação dos processos e às incertezas associadas à parametrização do sistema hidrológico (Dakhlaoui et al., 2017; Xiong et al., 2019). Além disso, a hipótese de parâmetros constantes no tempo pode limitar a representação das mudanças no comportamento da bacia, uma vez que a calibração com um único conjunto de parâmetros tende a representar predominantemente suas condições médias (Fowler et al., 2018; Pathiraja et al., 2018).

Os métodos de machine learning têm emergido como alternativas promissoras para a modelagem hidrológica, sobretudo na relação chuva–vazão (Li et al., 2021; Mendonça et al., 2021; Nourani et al., 2021). Esses métodos destacam-se pela capacidade de extrair padrões diretamente dos dados observados e representar relações complexas e não lineares entre as variáveis de entrada e a vazão (Okkan et al., 2021). Algoritmos como Random Forest (RF), Extreme Gradient Boosting (XGBoost), Multilayer Perceptron (MLP) e Long Short-Term Memory (LSTM) têm sido empregados na modelagem hidrológica, apresentando capacidade de representar relações complexas e não lineares entre variáveis hidrometeorológicas (Gholami & Sahour, 2022; Saadi et al., 2019; Sedighi et al., 2016; Shen et al., 2022; S. Wang et al., 2022).

Apesar desses avanços, o desempenho dos modelos de machine learning não depende apenas do algoritmo utilizado, podendo variar em função das características da bacia em que são aplicados. Diferenças de área de drenagem, relevo, declividade, cobertura e uso da terra e condições climáticas podem modificar a relação entre precipitação e vazão e, consequentemente, influenciar a capacidade dos modelos de reconstruir as séries hidrológicas. Assim, mais do que identificar um único modelo de melhor desempenho, torna-se importante avaliar sua consistência em bacias com características fisiográficas distintas.

Além das características fisiográficas das bacias, a extensão das falhas também pode influenciar a capacidade dos modelos de machine learning de reconstrução das séries de vazão. Falhas de curta duração preservam maior quantidade de informações próximas ao período ausente, enquanto blocos consecutivos mais longos aumentam a dificuldade de representar as variações hidrológicas ocorridas durante o intervalo sem observações (Lopes Martins et al., 2023; Solanki; Mishra, 2026). Sendo assim, avaliar diferentes extensões de falhas permite verificar até que ponto os modelos mantêm seu desempenho à medida que aumenta o período contínuo sem dados.

A bacia do Rio Doce apresenta condições adequadas para essa avaliação devido à sua elevada heterogeneidade espacial. Ao longo de sua extensão, suas sub-bacias apresentam diferenças de área de drenagem, relevo, altitude, declividade, uso e cobertura da terra e regime pluviométrico, resultando em respostas hidrológicas distintas (Marcuzzo et al., 2011). Essa variabilidade permite avaliar se modelos treinados sob diferentes condições fisiográficas mantêm capacidade semelhante de representar a dinâmica das vazões ou se seu desempenho é condicionado pelas características de cada sub-bacia.

Além dessa heterogeneidade, a bacia do Rio Doce apresenta relevância ambiental, social e econômica, abrangendo importantes centros urbanos, áreas agrícolas, remanescentes de vegetação nativa e atividades industriais e minerárias (Miranda et al., 2021). Essa multiplicidade de usos intensifica a pressão sobre os recursos hídricos e reforça a necessidade de ferramentas capazes de reconstruir séries de vazão de forma consistente em diferentes condições hidrológicas, contribuindo para análises de disponibilidade hídrica e para a gestão dos recursos hídricos.

Portanto, o presente trabalho objetivou avaliar a capacidade dos modelos Random Forest, Extreme Gradient Boosting, Multilayer Perceptron e Long Short-Term Memory em reconstruir séries de vazão a partir da relação chuva–vazão em sub-bacias do Rio Doce com diferentes características fisiográficas, considerando diferentes quantidades de dias consecutivos de falhas.

# 2. REVISÃO DE LITERATURA

## 2.1. Bacia do Rio Doce

A bacia hidrográfica do rio Doce localiza-se na região Sudeste do Brasil, abrangendo aproximadamente 86.700 km², dos quais cerca de 86% encontram-se no estado de Minas Gerais e 14% no estado do Espírito Santo. A bacia compreende aproximadamente 228 municípios e drena uma extensa área inserida predominantemente no domínio do bioma Mata Atlântica, um dos hotspots mundiais de biodiversidade, caracterizado por elevada riqueza biológica e intensa fragmentação decorrente da ocupação antrópica (Almeida et al., 2018; CBH-Doce, 2026).

Segundo a classificação climática de Köppen, predominam na bacia os tipos Cwa, caracterizado por clima subtropical de altitude com inverno seco e verão quente e chuvoso, e Aw, correspondente ao clima tropical com estação seca no inverno (Martins et al., 2018). O regime pluviométrico apresenta forte sazonalidade, concentrando aproximadamente 80% da precipitação anual entre os meses de outubro e março, enquanto o período de abril a setembro corresponde à estação seca (Cupolillo et al., 2022). A precipitação média anual é da ordem de 1.200 mm, porém apresenta elevada variabilidade espacial, influenciada principalmente pela topografia, pela distância em relação ao litoral e pelos sistemas atmosféricos atuantes sobre a região (Jesus et al., 2020). As temperaturas médias anuais variam entre aproximadamente 18 °C e 26 °C, sendo menores nas regiões serranas e maiores nas áreas de menor altitude (Martins et al., 2018).

A distribuição espacial das chuvas é controlada pela atuação conjunta de diferentes sistemas atmosféricos. Durante a estação chuvosa, a Zona de Convergência do Atlântico Sul (ZCAS) constitui o principal mecanismo responsável pelos eventos persistentes de precipitação sobre a bacia, frequentemente associada à ocorrência de enchentes e inundações. Além da ZCAS, frentes frias provenientes do sul do continente, sistemas convectivos de mesoescala, linhas de instabilidade e o transporte de umidade oriundo da Amazônia pelo Jato de Baixos Níveis da América do Sul (JBN) desempenham papel fundamental na formação das chuvas (Lima & Cupolillo, 2018; Montini et al., 2019). Em escala interanual, a variabilidade climática é influenciada por fenômenos oceânico-atmosféricos como o El Niño–Oscilação Sul (ENOS) e o Dipolo do Atlântico Tropical, que modificam a distribuição espacial e temporal das precipitações sobre a região Sudeste (Bombardi; Carvalho, 2011; Reboita et al., 2021).

O relevo da bacia apresenta elevada heterogeneidade, com amplitude altimétrica superior a 1.700 m, variando desde áreas montanhosas localizadas nas cabeceiras, associadas principalmente à Serra da Mantiqueira, Serra do Espinhaço Meridional e Quadrilátero Ferrífero, até o nível do mar na foz do rio Doce, localizada no litoral do Espírito Santo. Essa variação

altimétrica favorece a ocorrência de gradientes climáticos, especialmente de temperatura e precipitação, produzindo diferenças significativas na resposta hidrológica entre as sub-bacias. As regiões de maior altitude apresentam relevo fortemente dissecado, elevadas declividades e maiores índices pluviométricos decorrentes do efeito orográfico, enquanto as porções média e baixa da bacia caracterizam-se por relevo ondulado a suavemente ondulado e extensas planícies fluviais (Sondermann et al., 2022).

Quanto ao uso e cobertura da terra, a bacia do rio Doce apresenta elevado grau de antropização. As pastagens constituem a classe predominante, seguidas por áreas agrícolas, remanescentes de Mata Atlântica, florestas plantadas, principalmente de eucalipto, áreas urbanizadas, corpos hídricos e áreas de mineração. Estima-se que aproximadamente 76% da bacia esteja ocupada por atividades agropecuárias, refletindo um intenso processo histórico de conversão da vegetação nativa em áreas produtivas (CBH-Doce, 2026).

A economia regional é diversificada e fortemente dependente dos recursos hídricos. Destacam-se a agropecuária, especialmente os cultivos de café, cana-de-açúcar e eucalipto, a pecuária de corte e leiteira, a mineração, a produção de celulose, a geração de energia hidrelétrica e um dos maiores complexos siderúrgicos da América Latina. Aproximadamente 3,5 milhões de habitantes encontram-se distribuídos entre áreas urbanas e rurais, exercendo elevada pressão sobre os recursos naturais e ampliando os desafios relacionados à gestão integrada das águas (CBH-Doce, 2026).

Diante desse contexto fisiográfico, climático e socioeconômico, a bacia do rio Doce apresenta elevada complexidade hidrológica. A interação entre relevo acidentado, grande amplitude altimétrica, forte sazonalidade climática, variabilidade espacial das precipitações, diversidade de usos da terra e intensa ocupação humana resulta em respostas hidrológicas distintas entre suas sub-bacias, tornam a região ideal para entender como diferentes modelos de machine learning se saem ao estimar as vazões.

## 2.2. Importância das vazões médias diárias na gestão hídrica

A água constitui um recurso natural indispensável à manutenção da vida, ao equilíbrio dos ecossistemas e ao desenvolvimento das atividades humanas, incluindo o abastecimento público, a irrigação, a geração de energia e os processos industriais (Ejiohuo et al., 2025; Jung et al., 2023). No contexto das bacias hidrográficas, os rios assumem papel estruturante na dinâmica hidrológica, pois concentram o escoamento superficial proveniente das precipitações e as contribuições do

fluxo subterrâneo, funcionando como principais eixos de drenagem e integração dos processos físicos que ocorrem no território (Depetris, 2021; Smakhtin, 2001).

A disponibilidade hídrica de uma bacia está diretamente condicionada ao comportamento temporal das vazões. Com isso, o conhecimento do regime hidrológico, especialmente sob a perspectiva quantitativa e sazonal, é fundamental para avaliar a capacidade do sistema em atender às múltiplas demandas consuntivas e não consuntivas (Collins et al., 2024). A definição de vazões de referência assume papel estratégico, uma vez que esses indicadores sintetizam características do regime de escoamento e fornecem base técnica para o planejamento, a alocação e a regulação do uso da água (Silva et al., 2017).

Assim, a gestão de recursos hídricos fundamenta-se na quantificação robusta e contínua da disponibilidade hídrica. Conforme destacado por Moreira et al. (2022), a análise de séries históricas de vazões médias diárias é essencial para a estimativa de vazões de referência, subsidiando processos decisórios relacionados à outorga, ao enquadramento dos corpos d'água e à definição de estratégias de alocação. Dessa maneira, o monitoramento e a análise sistemática das vazões tornam-se instrumentos centrais para a promoção da segurança hídrica e da sustentabilidade na gestão de bacias hidrográficas.

## 2.3. Informações fluviométricas

A aquisição de dados de vazão constitui um dos pilares do monitoramento hidrológico, sendo fundamental para a gestão de recursos hídricos, previsão de cheias, planejamento de obras hidráulicas, estudos de variabilidade climática, além de serem a base da modelagem hidrológia (Dobriyal et al., 2017). Essa coleta é realizada por meio de redes de monitoramento compostas por estações fluviométricas convencionais e estações fluviométricas automáticas (telemétricas), responsáveis pelo registro sistemático do nível d'água e pela estimativa da descarga dos cursos fluviais (Mazi et al., 2023; Roza dos Santos et al., 2017).

Nas estações convencionais, a medição do nível d'água (cota) é realizada manualmente por meio de réguas linimétricas instaladas na seção do rio. Técnicos visitam periodicamente o local para efetuar as leituras e realizar medições de vazão com equipamentos específicos, como o molinete hidrométrico ou o medidor acústico Doppler (ADCP). Com base nessas medições de descarga, é construída ou atualizada a curva-chave, que estabelece a relação entre nível e vazão, permitindo a estimativa contínua das vazões a partir das cotas observadas (Gravelle, 2015). Apesar de sua relevância histórica e da confiabilidade quando bem mantidas, essas estações dependem de visitas regulares e estão sujeitas a limitações como falhas de leitura, intervalos longos entre

observações e dificuldades de acesso durante eventos extremos, o que pode comprometer a continuidade e a representatividade das séries (Birgand et al., 2013).

Por sua vez, as estações automáticas utilizam sensores eletrônicos para medir continuamente o nível da água, com registro em intervalos curtos e transmissão dos dados, em tempo quase real, para bancos de dados centrais via telefonia ou satélite. A vazão também é estimada a partir da relação nível-vazão previamente estabelecida (Sauer; Turnipseed, 2010). Embora a telemetria amplie expressivamente a frequência e a agilidade na obtenção das informações, essas estações não estão isentas de problemas. Podem ocorrer falhas de sensores, ruídos eletrônicos, interrupções na comunicação, inconsistências de registro e alterações físicas na seção de medição, como assoreamento ou mudanças no perfil transversal, que afetam a confiabilidade da curva-chave (Gacu et al., 2025). Além disso, a menor frequência de inspeções presenciais pode retardar a identificação de falhas instrumentais ou ambientais (Maggiotto et al., 2007). Essas incertezas impactam diretamente análises hidrológicas, como estimativas de frequência de cheias, estudos de tendência e modelagem de séries temporais (Fortesa et al., 2019).

Além disso, a cobertura espacial das redes de monitoramento ainda é insuficiente em diversas regiões do mundo, especialmente em áreas remotas ou países em desenvolvimento (Gacu et al., 2025). Muitos cursos d'água permanecem sem monitoramento contínuo, reduzindo a disponibilidade de dados para calibração e validação de modelos hidrológicos (Fischer et al., 2025). Essa limitação dificulta a compreensão dos regimes hidrológicos, compromete estimativas de eventos extremos e influencia diretamente a definição de vazões de projeto e a avaliação de riscos de inundação, com implicações relevantes para o planejamento de obras e para a formulação de políticas de gestão da água (Velásquez; Krajewski, 2024).

## 2.4. Modelagem hidrológica

A modelagem hidrológica é uma das principais ferramentas para compreender e representar os processos que controlam a transformação da precipitação em vazão, permitindo analisar a disponibilidade hídrica, prever cheias e subsidiar o planejamento de recursos hídricos. Os modelos buscam expressar matematicamente os mecanismos de interceptação, infiltração, armazenamento, evapotranspiração e escoamentos superficial e subterrâneo, a partir de um conjunto de variáveis de entrada que tipicamente incluem precipitação, temperatura do ar, radiação, umidade do solo, características fisiográficas da bacia e parâmetros relacionados ao uso e cobertura da terra (Jaiswal et al., 2020). A literatura tradicionalmente organiza os modelos em abordagens físicas (ou distribuídas com base em processos) e conceituais.

Os modelos físicos procuram representar explicitamente as leis de conservação de massa, energia e momento, descrevendo os fluxos de água por meio de equações diferenciais que dependem de propriedades do solo, topografia e condições de contorno (Brunner; Simmons, 2012; Grayson et al., 1992), cuja formulação integra escoamento superficial, fluxo em zona não saturada e aquíferos, fornecendo uma estrutura abrangente para simular a dinâmica hidrológica em escala de bacia. Embora esses modelos ofereçam uma interpretação fisicamente consistente dos processos, sua aplicação requer grande quantidade de dados detalhados e parametrizações difíceis de obter em muitas regiões do mundo (Xiong et al., 2019). A incerteza associada à heterogeneidade espacial, aliada à diferentes combinações de parâmetros que produzem resultados semelhantes, compromete a identificação de soluções únicas e limita a confiabilidade das simulações (Fowler et al., 2018).

Em contraste, os modelos conceituais simplificam a bacia em reservatórios interligados que representam armazenamentos médios e fluxos agregados, mantendo um equilíbrio entre realismo físico e parcimônia de dados (Vaze et al., 2011). Apresentam estruturas que exigem menos informações espaciais, utilizando principalmente séries de precipitação, temperatura e estimativas de evapotranspiração potencial (Lee et al., 2005). Contudo, essa representação agregada dos processos impõe limitações importantes. Embora seus parâmetros possuam interpretação física, os valores obtidos por calibração nem sempre correspondem diretamente às características físicas da bacia. Além disso, a transferência dos modelos para bacias não monitoradas permanece incerta, e seu desempenho pode ser reduzido em cenários de alterações no regime climático ou no uso e cobertura da terra (Seibert, 1999).

Além disso, tanto em modelos físicos quanto em modelos conceituais a calibração constitui uma etapa crítica no uso de modelos hidrológicos, visto que os parâmetros devem ser ajustados para que as simulações reproduzam adequadamente as séries observadas de vazão (por exemplo, precipitação–vazão). A escolha da função objetivo a ser otimizada, como por exemplo, coeficientes que quantificam a similaridade entre hidrogramas simulados e observados, influencia fortemente os parâmetros finais obtidos e, por conseguinte, o desempenho do modelo em diferentes condições hidrológicas (tais como vazões de pico e baixas vazões) (Althoff; Rodrigues, 2021).

Diante das dificuldades associadas à modelagem hidrológica, considerando que os processos de transformação de chuva em vazão são inerentemente não lineares, persistentes e não estacionários, conforme destacado por Ouali et al. (2017), torna-se necessário o emprego de técnicas capazes de representar adequadamente essa complexidade. Métodos de aprendizado de máquina têm se mostrado alternativas promissoras na reconstrução de séries históricas, pois

apresentam elevada eficiência na captura de padrões não lineares diretamente a partir dos dados, sem a imposição prévia de equações que descrevam explicitamente os mecanismos físicos da bacia (Solomatine; Ostfeld, 2008). Abordagens como redes neurais tradicionais e recorrentes, árvores de decisão e métodos *ensemble* têm se destacado pela capacidade de representar relações complexas entre variáveis hidrometeorológicas, alcançando bom desempenho preditivo em diferentes aplicações hidrológicas.

### 2.4.1. Random Forest (RF)

Entre os algoritmos de aprendizado de máquina normalmente empregados em estudos hidrológicos e ambientais, como demonstrado por Jeung et al. (2019), X. Liu et al. (2022), Nazeri Tahroudi et al. (2023), Ouali et al. (2017), Schoppa et al. (2020), destaca-se o Random Forest (RF). Proposto por Breiman, (2001), o RF é um método de ensemble aplicado tanto a problemas de classificação quanto de regressão. Modelos baseados em árvores de decisão são frequentemente utilizados nas ciências ambientais devido à sua elevada capacidade preditiva, simplicidade estrutural e à possibilidade de interpretação mais transparente dos resultados (Vilaseca et al., 2023).

O funcionamento do Random Forest baseia-se na construção de um grande número de árvores de decisão independentes, geradas a partir de reamostragens aleatórias do conjunto de dados original por meio da técnica de bootstrap, conforme descrito por (J. Hu; Szymczak, 2023). Em cada nó da árvore, apenas um subconjunto aleatório de variáveis preditoras é avaliado para a realização das divisões, o que aumenta a diversidade entre as árvores e reduz a correlação entre elas (Breiman, 2001).

Nesse processo, múltiplos modelos são ajustados a diferentes subconjuntos dos dados e posteriormente combinados por meio de agregação, geralmente pela média das estimativas individuais. Em problemas de regressão, cada árvore fornece uma estimativa baseada nas observações presentes no nó terminal correspondente, e a predição final do modelo resulta da média das previsões de todas as árvores da floresta, conforme discutido por (Biau; Scornet, 2016).

Diante dessas características, o Random Forest apresenta elevada capacidade de capturar os padrões de variabilidade das vazões, especialmente em sistemas hidrológicos marcados por não linearidades e interações complexas entre variáveis meteorológicas e hidrológicas.

### 2.4.2. Extreme Gradient Boosting – XGBoost (XGB)

Outra técnica de aprendizado de máquina que tem apresentado resultados expressivos na modelagem chuva-vazão é o Extreme Gradient Boosting (XGB), proposto por (Chen; Guestrin, 2016). Esse algoritmo consiste em uma implementação otimizada do método de *gradient boosting*, desenvolvida para aumentar a eficiência computacional e a capacidade preditiva dos modelos. Sua estrutura baseia-se na construção sequencial de árvores de decisão, em que cada nova árvore é ajustada para minimizar a função de perda por meio da correção dos erros residuais produzidos pelas árvores previamente construídas, promovendo o refinamento progressivo das estimativas. Além disso, o algoritmo incorpora um termo de regularização à função objetivo, bem como técnicas de *shrinkage* e amostragem de variáveis (*column subsampling*), que contribuem para reduzir o sobreajuste e melhorar a capacidade de generalização do modelo (Chen; Guestrin, 2016).

O XGBoost combina a construção sequencial de árvores de decisão com mecanismos de regularização que controlam a complexidade do modelo e favorecem sua capacidade de generalização. Além disso, incorpora estratégias para tratamento de dados esparsos, amostragem aleatória, redução da complexidade computacional e paralelização do processo de treinamento, tornando o algoritmo eficiente para aplicações envolvendo grandes conjuntos de dados (Bentéjac et al., 2021).

Nos últimos anos, diversos estudos têm demonstrado a eficiência do XGBoost na modelagem hidrológica. Trabalhos desenvolvidos por W. Liu et al. (2025), Parisouj et al. (2024) e Sumith (2025) evidenciaram elevado desempenho do algoritmo na previsão de vazões, destacando sua capacidade de representar as relações não lineares entre precipitação e escoamento e de produzir estimativas consistentes em diferentes condições hidrológicas. Esses estudos mostram que o ajuste sequencial das árvores favorece a redução progressiva dos erros de previsão, resultando em maior precisão na simulação das séries de vazão.

Entretanto, assim como outros modelos baseados em aprendizado de máquina, o desempenho do XGBoost depende diretamente da qualidade e da representatividade dos dados utilizados no treinamento, bem como da definição adequada de seus hiperparâmetros. Em situações nas quais o conjunto de treinamento não representa adequadamente as condições hidrológicas de aplicação, ou quando os parâmetros do modelo não são convenientemente ajustados, sua capacidade de generalização pode ser reduzida, comprometendo a precisão das estimativas (W. Liu et al., 2025; Parisouj et al., 2024).

Apesar dessas limitações, o XGBoost tem sido empregado em estudos de modelagem hidrológica, principalmente em aplicações relacionadas à previsão de vazões, preenchimento de

falhas em séries hidrológicas e modelagem chuva-vazão como nos trabalhos de (Magyari-Sáska et al., 2025; Marcos Junior et al., 2024; Ni et al., 2020).

### 2.4.3. Multilayer Perceptron (MLP)

As Redes Neurais Artificiais (RNAs) são modelos de aprendizado de máquina inspirados na estrutura e no funcionamento das redes neurais biológicas, capazes de aprender relações complexas e não lineares entre variáveis a partir de dados observados. Entre as diferentes arquiteturas existentes, a Multilayer Perceptron (MLP) destaca-se como uma das mais utilizadas na modelagem hidrológica e na previsão de vazões, apresentando resultados satisfatórios em diferentes condições climáticas e fisiográficas (Lukas et al., 2024; Sahoo et al., 2022; Shin; Jung, 2025; Valles, 2024).

A MLP é formada por uma camada de entrada, uma ou mais camadas ocultas e uma camada de saída. Durante o treinamento, os neurônios artificiais ajustam continuamente seus pesos sinápticos por meio da comparação entre os valores estimados e os valores observados, permitindo que a rede aprenda os padrões presentes nos dados. Dessa forma, o modelo é capaz de representar relações complexas entre as variáveis sem que seja necessário definir previamente uma formulação matemática para o processo hidrológico (Darbandi; Pourhosseini, 2018; Dias et al., 2016).

Essa capacidade tem favorecido o emprego da MLP em aplicações como modelagem chuva-vazão, previsão de vazões e preenchimento de falhas em séries hidrológicas. Entretanto, seu desempenho depende da qualidade e da representatividade dos dados utilizados no treinamento, bem como da definição adequada da arquitetura da rede. Quando esses aspectos não são devidamente considerados, o modelo pode apresentar sobreajuste (*overfitting*) ou perda da capacidade de generalização, reduzindo a precisão das estimativas em novos conjuntos de dados (Piotrowski; Napiorkowski, 2013; Senthil Kumar et al., 2005).

### 2.4.4. Long Short-Term Memory (LSTM)

Entre as técnicas de aprendizado profundo empregadas na modelagem hidrológica, destacam-se as redes neurais recorrentes do tipo Long Short-Term Memory (LSTM), amplamente utilizadas para representar a elevada não linearidade e a dependência temporal características dos processos hidrológicos. Essa arquitetura, proposta por Hochreiter e Schmidhuber (1997), foi desenvolvida para superar as limitações das redes neurais recorrentes convencionais, como a dificuldade em aprender relações de longo prazo presentes em séries temporais.

A principal característica das redes LSTM é a capacidade de armazenar e utilizar informações de estados anteriores da sequência durante o processo de aprendizagem. Diferentemente das redes neurais tradicionais, que tratam cada observação de forma isolada, a LSTM considera a dependência entre eventos consecutivos, permitindo que informações relevantes sejam preservadas ao longo do tempo. Essa propriedade é importante na modelagem hidrológica, uma vez que a resposta de uma bacia hidrográfica não depende apenas da precipitação ocorrida em um determinado instante, mas também das condições antecedentes de umidade do solo, do armazenamento de água e da dinâmica do escoamento ao longo do tempo (Kratzert et al., 2018).

O funcionamento da arquitetura baseia-se em uma célula de memória capaz de controlar o fluxo de informações por meio de três mecanismos principais: a porta de esquecimento (*forget gate*), responsável por remover informações consideradas pouco relevantes; a porta de entrada (*input gate*), que seleciona quais novos dados serão incorporados ao estado da memória; e a porta de saída (*output gate*), que determina quais informações armazenadas serão utilizadas para produzir a resposta da rede em cada passo temporal (Hochreiter; Schmidhuber, 1997).

Essa capacidade de preservar informações ao longo da sequência temporal permite que a LSTM represente a memória hidrológica da bacia, incorporando indiretamente os efeitos de processos como armazenamento de água no solo, contribuição do escoamento subterrâneo e resposta retardada da precipitação sobre a vazão, sem a necessidade de especificar explicitamente esses mecanismos físicos (Kratzert et al., 2018).

Nos últimos anos, as redes LSTM passaram a ocupar posição de destaque na hidrologia devido à sua capacidade de aprender relações complexas entre as variáveis hidrometeorológicas diretamente a partir dos dados observados. Diversos estudos demonstram que essa arquitetura apresenta elevado desempenho na simulação de vazões, frequentemente superando ou igualando modelos hidrológicos conceituais em diferentes condições climáticas e fisiográficas (Cho; Kim, 2022; Fathi et al., 2025; Girihagama et al., 2022; Hunt et al., 2022; Marasini; Pokhrel, 2024; Sabzipour et al., 2023). Além da elevada capacidade preditiva, a LSTM apresenta maior facilidade para representar relações altamente não lineares e dependências temporais complexas, características inerentes aos processos hidrológicos.

Entretanto, assim como outros modelos baseados em aprendizado de máquina, as redes LSTM também apresentam limitações. Seu desempenho depende diretamente da qualidade, da quantidade e da representatividade dos dados utilizados durante o treinamento, podendo apresentar redução da capacidade preditiva quando aplicadas em condições hidrológicas que não estejam

adequadamente representadas no conjunto de treinamento ou que diferem significativamente daquelas observadas durante a calibração (Lu et al., 2021).

## 2.5. Covariáveis importantes na modelagem Chuva-Vazão

Em aplicações de modelagem chuva–vazão, especialmente em regiões com limitação de dados hidrológicos, a precipitação é adotada como principal variável de entrada por representar a forçante dominante do sistema hidrológico e o principal aporte de água responsável pela geração de escoamento superficial e pela resposta de vazão nos rios. Estudos como os de Hu et al. (2022) e Xie et al. (2023) destacam que a precipitação constitui a informação fundamental para a simulação do processo chuva–vazão, uma vez que controla o início, a intensidade e a variabilidade temporal do escoamento.

Entretanto, a resposta hidrológica não depende exclusivamente da chuva incidente. Variáveis ambientais como temperatura, umidade do solo e cobertura do terreno exercem influência expressiva na partição da água entre infiltração, armazenamento e escoamento superficial. A temperatura do ar e do solo modula a evapotranspiração e o balanço hídrico, afetando a quantidade de água efetivamente disponível para gerar vazão, conforme discutido por Materia et al. (2022). A dinâmica da umidade do solo, por sua vez, controla diretamente a capacidade de infiltração e o armazenamento, regulando a intensidade da resposta da bacia a eventos de chuva (Penna et al., 2011; Scaife; Band, 2017; Ye et al., 2023). Além disso, alterações na cobertura e no uso da terra, como urbanização ou mudanças na vegetação, modificam propriedades físicas da superfície, como rugosidade e infiltração, influenciando a conectividade hidrológica e a geração de escoamento (Tuladhar et al., 2019).

Apesar da reconhecida importância dessas variáveis, a obtenção de dados detalhados sobre temperatura e umidade do solo nem sempre é viável, sobretudo em regiões com monitoramento limitado (Carvalho et al., 2015; Vilanova et al., 2012). Além disso, a representação estacionária das características fisiográficas da bacia pode comprometer a capacidade dos modelos hidrológicos em reproduzir adequadamente a variabilidade espacial e temporal dos processos, especialmente quando submetidos a diferentes padrões de precipitação (Loritz et al., 2021). A utilização exclusiva da precipitação torna-se uma alternativa metodológica mais acessível. Para representar, ainda que de forma indireta, as condições antecedentes de umidade do solo, pode-se recorrer à inclusão da chuva antecedente nos modelos hidrológicos.

A chuva antecedente refere-se aos volumes precipitados em dias anteriores ao evento analisado, enquanto a precipitação antecedente acumulada corresponde à soma desses volumes ao

longo de uma janela temporal definida (por exemplo, 1, 3, 5, 7 ou 10 dias). Essa abordagem incorpora o efeito de memória hidrológica da bacia, reconhecendo que a resposta do escoamento não depende apenas da chuva do dia corrente, mas também do histórico recente de precipitações (Lakew et al., 2026). A precipitação antecedente é considerada um dos principais fatores que controlam a umidade do solo e, consequentemente, a geração de escoamento superficial (Ali et al., 2010; Bennett et al., 2018).

Quando o solo se encontra previamente úmido ou próximo à saturação, sua capacidade de infiltração é reduzida, favorecendo maior proporção de escoamento superficial e respostas de vazão mais rápidas e intensas. Em contraste, sob condições de solo seco, parcela significativa da precipitação inicial é utilizada para recompor o armazenamento hídrico, retardando e atenuando o aumento das vazões (Schoener; Stone, 2019; Song; Wang, 2019).

Além da precipitação, as vazões antecedentes também têm sido utilizadas como covariáveis em modelos de aprendizado de máquina voltados à previsão de vazões. Enquanto a precipitação representa a principal forçante do processo hidrológico, as vazões observadas nos dias anteriores refletem a resposta integrada da bacia aos diferentes processos que ocorrem ao longo do tempo, como infiltração, armazenamento de água no solo e nos aquíferos, contribuição do escoamento de base e propagação do escoamento pela rede de drenagem. Dessa forma, essas informações complementam as variáveis meteorológicas ao fornecer uma representação do estado hidrológico da bacia no momento da previsão (Ashfaq et al., 2026; Tongal;  Booij, 2018).

A incorporação de uma ou mais vazões defasadas permite que os modelos considerem a persistência temporal do escoamento, característica frequentemente denominada memória hidrológica da bacia. Como a vazão de um determinado dia é fortemente influenciada pelas condições hidrológicas dos dias anteriores, a utilização dessas informações contribui para aumentar a capacidade preditiva dos modelos, especialmente em previsões de curto prazo. Diversos estudos mostram que as vazões antecedentes figuram entre as variáveis mais relevantes para a estimativa da vazão futura, muitas vezes apresentando maior contribuição para o desempenho dos modelos do que outras covariáveis hidrometeorológicas (Hao; Yan, 2025; Pulgarín-Morales et al., 2026).

## 3. MATERIAIS E MÉTODOS

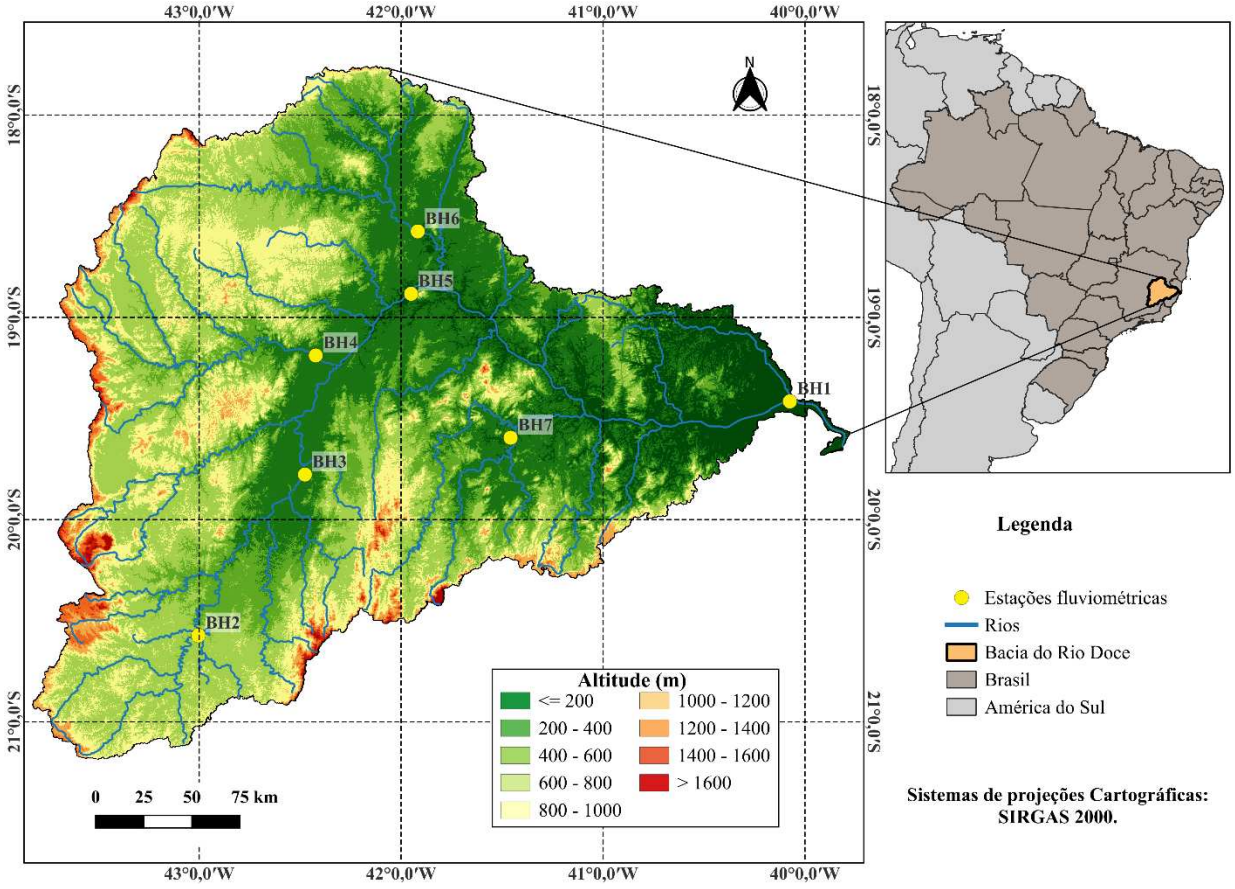
### 3.1. Área de estudo

A modelagem hidrológica foi desenvolvida na bacia hidrográfica do Rio Doce (Figura 1), que possui área aproximada de 86.715 km$^2$ e está localizada na região Sudeste do Brasil. Cerca de

86% de sua área encontra-se inserida no estado de Minas Gerais, enquanto a porção remanescente se estende pelo estado do Espírito Santo (CBH-Doce, 2026).

Do ponto de vista fisiográfico, a bacia apresenta elevada variabilidade altimétrica, com cotas que variam desde o nível do mar, na região da foz, localizada no estado do Espírito Santo, até altitudes superiores a 1.700 m nas áreas de cabeceira, próximas às nascentes (Cupolillo et al., 2022).

**Figura 1:** Mapa de localização da bacia do Rio Doce




## 3.2. Obtenção da base de dados

Para a modelagem chuva–vazão na bacia hidrográfica do rio Doce, foram utilizados dados diários de precipitação e vazão média. Os dados de vazão foram obtidos em sete estações fluviométricas (Tabela 1), disponíveis no sistema HidroWeb, da Agência Nacional de Águas e Saneamento Básico (ANA). Foram consideradas as séries disponíveis entre 2001 e 2023, sendo

selecionadas, para cada estação, as informações correspondentes aos 10 anos completos mais próximos de 2023.

Após a seleção das estações, foi realizada a análise da estacionariedade das séries de vazão, com o objetivo de definir o período de estudo a partir da identificação de mudanças de tendência, conforme apresentado no item 3.3. No período definido, todas as estações apresentaram séries de vazão contínuas, sem falhas. A estação BH1 foi a única que não apresentou 10 anos completos de dados após a mudança de tendência identificada, apresentando 8 anos de dados de vazão completos. Ainda assim, optou-se por sua inclusão no estudo por estar localizada mais próxima ao exutório da bacia e, consequentemente, representar uma área de drenagem mais próxima da área total da bacia.

Com base nas estações fluviométricas selecionadas (Tabela 1), foram delimitadas as bacias hidrográficas a montante de cada estação (Figura 2). Esse procedimento foi realizado no ambiente do software QGIS, por meio da integração das ferramentas do módulo GRASS GIS.

Para a delimitação das bacias hidrográficas, utilizou-se o Modelo Digital de Elevação (MDE) Copernicus, com resolução espacial de 30 m, obtido por meio do complemento OpenTopography DEM Downloader no software QGIS. Inicialmente, realizou-se o preenchimento de depressões espúrias (*fill sinks*) para garantir a continuidade do escoamento superficial. Em seguida, foram determinadas a direção (*flow direction*) e a acumulação de fluxo (*flow accumulation*), sendo esta última utilizada para identificar as células de maior concentração do escoamento e definir a rede de drenagem, como indicado por (Nicolete & Piroli, 2024). A partir da rede de drenagem e das estações fluviométricas, utilizadas como pontos de exutório, foram determinadas as respectivas áreas de contribuição e assim delimitadas as bacias hidrográficas, posteriormente convertidas para o formato vetorial, conforme apresentado na Figura 2.

Tabela 1. Descrição das estações fluviométricas

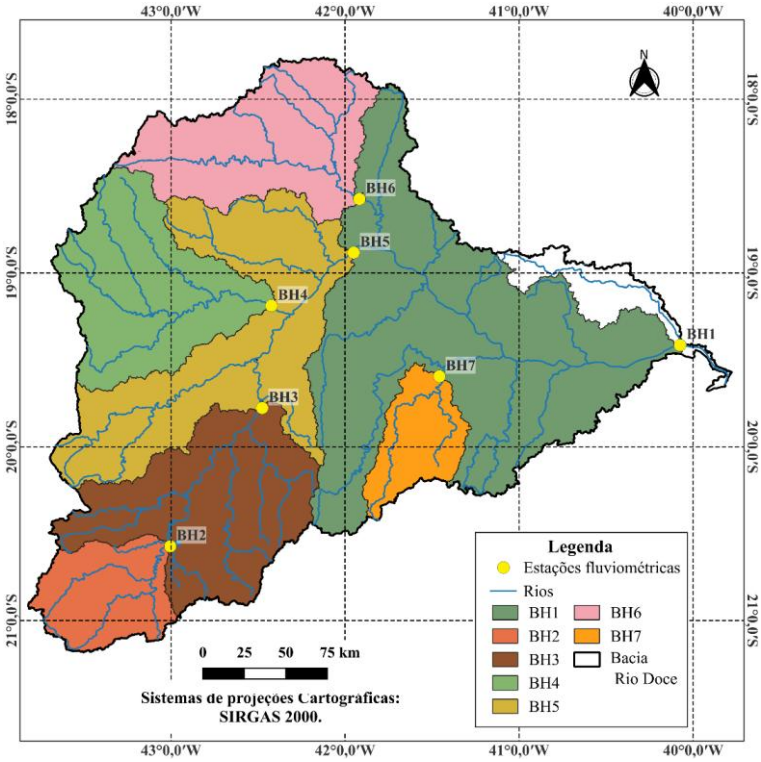
| **Código da estação** | **Código ANA** | **Nome da Estação** | **Latitude** | **Longitude** | **Área (km²)** |
|---|---|---|---|---|---|
| BH1 | 56998200 | LINHARES CAIS DO PORTO | 19°24'55.1"S | 40°04'32.2"W | 82.100 |
| BH2 | 56094000 | PCH BRECHA MONTANTE 1 | 20°34'21.0"S | 43°00'15.1"W | 4.680 |

| | | | | | |
|---|---|---|---|---|---|
| BH3 | 56539000 | CACHOEIRA DOS ÓCULOS MONTANTE | 19°46'36.8"S | 42°28'35.0"W | 15.900 |
| BH4 | 56825000 | NAQUE VELHO | 19°11'17.2"S | 42°25'22.1"W | 10.200 |
| BH5 | 56850000 | GOVERNADOR VALADARES | 18°52'59.2"S | 41°57'02.9"W | 40.500 |
| BH6 | 56891900 | VILA MATIAS MONTANTE | 18°34'28.9"S | 41°55'04.1"W | 9.770 |
| BH7 | 56989400 | ASSARAI MONTANTE | 19°35'40.9"S | 41°27'29.2"W | 3.190 |

Os dados de precipitação foram obtidos a partir de estações pluviométricas pertencentes à Rede Hidrometeorológica Nacional, disponibilizadas no sistema Hidroweb da Agência Nacional de Águas e Saneamento Básico. Foram selecionadas as estações localizadas na bacia do rio Doce que apresentaram séries diárias completas de precipitação para o mesmo período dos dados de vazão correspondente a cada estação.

A precipitação média diária de cada bacia foi determinada pelo método dos polígonos de Thiessen, utilizado para representar a contribuição espacial das estações pluviométricas na estimativa da precipitação média em bacias hidrográficas, como feito por (Pereira et al., 2024). Os polígonos foram gerados a partir da distribuição espacial das estações e intersectados com os limites de cada uma das sete bacias, permitindo determinar a área de influência de cada estação. Em seguida, a precipitação média diária em cada bacia foi calculada por média ponderada, utilizando-se a proporção da área de influência de cada estação pluviométrica como fator de ponderação.

**Figura 2:** Bacias hidrográficas delimitadas para a modelagem hidrológica na área de estudo.




3.3. Análise de Tendências nas séries vazões médias diárias

A análise de estacionalidade foi aplicada às séries de vazões médias diárias para verificar a ocorrência de mudanças ao longo do tempo. A partir dessa análise, foi definido o período mais recente com comportamento hidrológico homogêneo, utilizado posteriormente nas etapas de treinamento e teste dos modelos de aprendizado de máquina.

Inicialmente, aplicou-se a Função de Autocorrelação (ACF), como demostrado no item 3.3.1, para verificar a presença de autocorrelação, analisando-se os coeficientes para diferentes defasagens temporais e sua significância estatística ao nível de 5% nas séries de vazão (Witt, 2007). Em seguida, utilizou-se o teste de aleatoriedade de Run (Equação 3) com o intuito de avaliar se as séries apresentavam comportamento aleatório ou estrutura temporal definida (Thom, 1966).

Após a verificação da dependência serial, procedeu-se à análise de tendência das séries. Para aquelas que não apresentaram autocorrelação significativa, aplicou-se o teste tradicional de Mann-Kendall (item 3.3.2.), adequado sob a hipótese de independência entre observações (Kendall, 1948; Mann, 1945).

Para as séries em que foi identificada autocorrelação significativa, utilizou-se a versão modificada do teste de Mann-Kendall, apropriada para dados temporalmente correlacionados, de

modo a evitar a superestimação da significância estatística decorrente da dependência serial (Hamed; Ramachandra Rao, 1998).

Também foi aplicado o coeficiente de correlação de Spearman para avaliar a presença de tendência monotônica nas séries de vazão (Ali Abd Al-Hameed, 2022).

Uma vez detectada tendência nos dados diários, aplicou-se o teste de Pettitt com o objetivo de identificar possíveis pontos de mudança abrupta na série temporal (Pettitt, 1979). A partir das quebras identificadas pelo teste de Pettitt, delimitou o período de estudo. Destaca-se que em todas as estações fluviométricas adotadas no estudo identificaram-se mudanças de tendencias a partir de 2013.

### 3.3.1. Função de Autocorrelação (ACF)

O número esperado de oscilações e sua variância foram determinados conforme descrito nas Equações (1) e (2), enquanto a estatística Z foi obtida de acordo com a Equação (3). Quando o valor calculado de Z ultrapassou o valor crítico da distribuição normal para o nível de significância adotado, a hipótese nula foi rejeitada.

A hipótese nula ($H_0$) considera que a sequência das vazões diárias apresenta comportamento aleatório ao longo do tempo, ou seja, que as variações observadas não seguem tendência, padrão sistemático ou dependência temporal, sendo resultado apenas de flutuações ao acaso. Assim, a rejeição de $H_0$ indica evidências de que a série das vazões, não é puramente aleatória, sugerindo a presença de tendência, persistência ou outra forma de estrutura nas séries de vazões, como descrito por (Jacomazzi et al., 2022).

$$E(n) = \frac{2N_1N_2}{N_1 + N_2} + 1 \tag{1}$$

$$Var(n) = \frac{2N_1N_2(2N_1N_2 - N_1 - N_2)}{(N_1 + N_2)^2(N_1 + N_2 - 1)} \tag{2}$$

$$\mathrm{Z} = \frac{n + E(n)}{\sqrt{Var(n)}} \tag{3}$$

onde:

$E(n)$ = distribuição amostral, adimensional;

$Var(n)$ = variância da distrivuição, adimensional;

$Z$ = estatística do teste Run, adimensional;

$n$ = número total de oscilações, adimensional;

$N_1$ = número de dados na série temporal abaixo da mediana, adimensional; e

$N_2$ = número de dados na série temporal acima da mediana, admensional.

### 3.3.2. Mann-Kendall (MK) e Mann-Kendall Modificado (MKM)

Séries temporais com extensos períodos de observação nem sempre atendem às premissas estatísticas exigidas por métodos paramétricos, especialmente quanto à forma da distribuição dos dados. A aplicação de abordagens não paramétricas mostra-se uma alternativa adequada, destacando-se o teste de Mann-Kendall. Conforme descrito por Cabral Júnior; Lucena (2020), a estatística do teste é expressa pela seguinte equação:

$$S = \sum_{i=2}^{n} \sum_{j=1}^{i=1} sinal\,(x_j - x_i) \tag{4}$$

$$sinal = \begin{cases} +1, se\ (x_j - x_i) > 0 \\ 0, se\ (x_j - x_i) = 0 \\ -1, se\ (x_j - x_i) < 0 \end{cases} \tag{5}$$

$$VAR = \frac{1}{18}\left[n(n-1)(2n+5) - \sum_{p=1}^{q} t_p(t_p - 1)(2t_p + 5)\right] \tag{6}$$

onde:

S = estatística do teste MK, adimensional;

$(x_j - x_i)$ = valores da série de vazão, $m^3\ s^{-1}$;

sinal = Função sinal $xj - xi$, adimensional;

$Var(n)$ = variância da estatística S sob a hipótese nula, adimensional;

n = tamanho da série temporal, adimensional;

$t_p$ = número de observações no p-ésimo grupo empatado

Quando o tamanho da amostra é igual ou superior a 10 observações, a estatística $S$ passa a ser adequadamente representada por uma aproximação à distribuição normal. Considerando a hipótese nula de inexistência de tendência na série, assume-se que o valor médio esperado de $S$ é nulo. A estimativa da variância dessa estatística incorpora a correção para a presença de valores repetidos (empates) nos dados, sendo determinada conforme a seguinte expressão:

$$Z_{MK} = \begin{cases} \frac{(\mathrm{S}-1)}{\sqrt{Var(S)}}, se\ S > 0 \\ 0\ , se\ S = 0 \\ \frac{(\mathrm{S}+1)}{\sqrt{Var(S)}}, se\ S < 0 \end{cases} \quad (7)$$

onde:

S = estatística do teste MK, adimensional;

$Z_{MK}$ = estatística Z padronizada, adimensional; e

$Var(S)$ = variância, $m^3\ s^{-1}$.

Quando foi identificada a presença de autocorrelação na série, aplicou-se o teste de Mann-Kendall Modificado, conforme proposto por (Hamed & Ramachandra Rao, 1998), no qual a variância da estatística é ajustada para considerar os efeitos da dependência serial nos dados.

$$Var_{MKM}(\mathrm{S}) = \frac{n(n-1)(2n+5)}{18}\ \mathrm{x}\ \frac{n}{ns} \quad (8)$$

$$\frac{n}{ns} = 1 + \frac{2}{n(n-1)(n-2)}\, x \sum_{i=1}^{n-1} (n-1)(n-i-1)(n-i-2)\rho_k \quad (9)$$

$$Z_{MKM} = \begin{cases} \dfrac{(\mathrm{S}-1)}{\sqrt{Var_{MKM}(\mathrm{S})}}, se\ S > 0 \\ 0\ , se\ S = 0 \\ \dfrac{(\mathrm{S}+1)}{\sqrt{Var_{MKM}(\mathrm{S})}}, se\ S < 0 \end{cases} \quad (10)$$

onde:

| | | |
|---|---|---|
| $\frac{n}{ns}$ | | fator de correção, admensional; |
| $n$ | | número de observações; |
| $\rho_k$ | | coeficiente de autocorrelação da série para a defasagem k, admensional; |
| S | = | estatística do teste MK, adimensional; |
| $Z_{MKM}$ | = | estatística Z padronizada, adimensional; e |
| $Var_{MKM}(\mathrm{S})$ | = | variância modificada, $m^3s^{-1}$. |

### 3.3.3. Correlação de Spearman

Utilizou-se o coeficiente de correlação de Spearman (ρ), utilizado na identificação de relações monotônicas em séries temporais (Sneyers, 1991). Diferentemente dos métodos paramétricos, o coeficiente de Spearman baseia-se na ordenação (postos) dos dados, não exigindo pressupostos quanto à forma da distribuição ou à linearidade da relação entre as variáveis.

Na análise de tendências temporais, o teste consiste em correlacionar os valores observados da série com a ordem cronológica (tempo). Valores positivos e estatisticamente significativos de ρ indicam tendência crescente, enquanto valores negativos sugerem tendência decrescente. Por se tratar de uma medida não paramétrica, o coeficiente de Spearman é menos sensível à presença de outliers e assimetrias, características frequentemente observadas em séries hidrológicas (Kahya & Kalaycı, 2004).

O coeficiente de correlação de Spearman foi determinado a partir da seguinte expressão:

$$\rho = 1 - \frac{6\sum_{i=1}^{n}[Xi - Yi]^2}{n(n^2-1)} \quad (11)$$

onde:

| | | |
|---|---|---|
| ρ | = | coeficiente de correlação de Spearman, adimensional; |
| $n$ | = | número de observações; e |
| Xi e Yi | = | postos (ordens) dos i-ésimos valores das variáveis X e Y, respectivamente. |

### 3.3.4. Teste de Pettitt

Após a detecção de tendências por meio dos testes de Mann-Kendall (MK) ou Mann-Kendall Modificado (MKM) e da inspeção visual das séries temporais para identificar padrões de acréscimo ou decréscimo, procedeu-se à aplicação do teste de Pettitt para identificar possíveis pontos de ruptura. O teste de Pettitt é um método não paramétrico de detecção de mudanças abruptas na mediana ou na tendência central de uma série temporal, baseado em estatísticas de postos e comparações entre todos os pares de observações antes e depois de cada possível ponto de ruptura (Rybski; Neumann, 2011).

Esse procedimento testa a hipótese nula de que não existe mudança na série contra a alternativa de que existe um ponto no tempo no qual ocorre um deslocamento estatisticamente significativo na distribuição dos dados. A localização do ponto de mudança é obtida a partir do máximo absoluto da estatística de Pettitt e sua significância é avaliada por meio de uma aproximação assintótica da probabilidade associada (R. M. Moreira et al., 2024).

O uso do teste de Pettitt em conjunto com testes de tendência é amplamente documentado na literatura hidrológica e climática como uma abordagem robusta para caracterizar não apenas a existência de tendências monotônicas, mas também a presença de mudanças estruturais ao longo do tempo, dividindo séries longas em subperíodos antes e depois da ruptura para análises subsequentes (Ďurigová et al., 2020).

Conforme descrito por Moreira et al (2024), o teste de Pettitt foi aplicado utilizando a seguinte equação:

$$U_{t,n} = \sum_{i=1}^{t} \sum_{j=1}^{n} sinal(X_t - X_j) \qquad (12)$$

onde:

$U_{t,n}$ = estatística do teste de Pettit, adimensional;

$n$ = número de observações;

sinal = sinal da operação ($Xt - Xj$), adimensional;

t e j = posição dos valores na série, adimensional; e

$X_t - X_j$ = valores da série, $m^3/s$ para vazão.

### 3.3.5. Período de treino e teste

Após a identificação do ponto de ruptura na série histórica de vazões, foram utilizados, para as etapas de treinamento e teste, os dados correspondentes ao período posterior à ruptura, buscando trabalhar com um intervalo temporal de comportamento mais homogêneo (Tabela 2). A partir desse período, os dados foram divididos em conjuntos de treinamento e teste, sendo 30% destinados ao treinamento dos modelos e 70% ao teste, conforme adotado por (Dastour; Hassan, 2023; Küçüktopcu et al., 2023).

Para a avaliação do preenchimento de falhas, os dois primeiros anos da série foram mantidos sem a inserção de falhas e já compuseram os 70% destinados ao treinamento, garantindo um período inicial contínuo e a disponibilidade das informações antecedentes necessárias para as estimativas dos modelos. A partir desse período inicial, as falhas foram distribuídas ao longo da série temporal utilizada na etapa de teste. Foram avaliados cinco cenários de ausência de dados, constituídos por blocos consecutivos de 10, 30, 60, 90 e 120 dias.

Para cada cenário de ausência de dados foram realizadas 10 repetições da simulação das falhas, variando-se a posição dos blocos de dados de falhas ao longo do conjunto de teste. Esse procedimento permitiu avaliar o desempenho dos modelos diante de diferentes períodos de ausência de dados, reduzindo a influência da localização das falhas sobre os resultados.

O preenchimento foi realizado de forma sequencial ao longo de cada bloco de falhas. Como a vazão do dia anterior foi utilizada como covariável de entrada dos modelos, após a estimativa do primeiro dia ausente o valor de vazão simulado passou a ser utilizado como informação antecedente para a estimativa do dia seguinte. Esse procedimento foi repetido sucessivamente até o preenchimento completo de cada bloco de 10, 30, 60, 90 ou 120 dias, reproduzindo uma situação em que os valores observados de vazão não estão disponíveis durante todo o período de falha.

**Tabela 2.** Extensão das séries temporais de cada estação fluviométrica, definido a partir do ponto de mudança identificado pelo teste de Pettitt

| Código da Estação | Série Histórica |
|---|---|
| BH1 | 01/01/2015 – 31/12/2023 |
| BH2 | 01/01/2012 – 31/12/2021 |
| BH3 | 01/01/2013 – 31/12/2023 |
| BH4 | 01/01/2013 – 31/12/2023 |
| BH5 | 01/01/2013 – 31/12/2023 |
| BH6 | 01/01/2013 – 31/12/2023 |
| BH7 | 01/01/2013 – 31/12/2023 |

## 3.4. Pré-processamento dos dados de chuva vazão e definição das covariáveis

### 3.4.1. Vazão

As séries de vazão observadas apresentam magnitudes distintas entre as sete bacias hidrográficas, refletindo as diferenças nas áreas de contribuição e, consequentemente, nos volumes de precipitação aportados a montante de cada estação fluviométrica. Assim, bacias com maiores áreas de drenagem tendem a apresentar vazões mais elevadas, enquanto bacias menores tendem a apresentar vazões menores (Erikson et al., 2024). Diante dessa heterogeneidade, tornou-se necessária a padronização das vazões, de modo a permitir a comparação adequada dos resultados entre bacias com diferentes áreas de drenagem.

Com isso, as vazões observadas foram convertidas em vazão específica diária por meio da sua divisão pela respectiva área de drenagem da respectiva bacia hidrográfica, de modo a expressar a resposta hidrológica em relação à área de captação. Com isso, o processo de modelagem passou a utilizar a precipitação média diária na área de drenagem de cada bacia e a vazão específica correspondente. As áreas de drenagem das bacias hidrográficas utilizadas no cálculo da vazão específica diária estão apresentadas na Tabela 1.

### 3.4.2. Precipitação

A precipitação utilizada na modelagem corresponde à precipitação média ponderada diária na área de drenagem a montante de cada estação fluviométrica, calculada conforme descrito no item 3.2.

### 3.4.3. Definição das covariáveis de entrada

Para representar a dinâmica chuva–vazão nos modelos RF, XGB e MLP, foram definidas variáveis explicativas a partir da precipitação diária e da vazão antecedente. Para a precipitação, foram considerados o valor correspondente ao dia da vazão a ser estimada e os valores registrados nos cinco dias anteriores. Para a vazão, foi utilizado o valor observado no dia imediatamente anterior. Assim, foram estabelecidas as seguintes covariáveis:

- $Q_1$: vazão do 1º dia antecedente;
- $P_0$: precipitação do dia corrente (associada à vazão observada no mesmo dia);
- $P_1$: precipitação do 1º dia antecedente;
- $P_2$: precipitação do 2º dia antecedente;
- $P_3$: precipitação do 3º dia antecedente;
- $P_4$: precipitação do 4º dia antecedente;
- $P_5$: precipitação do 5º dia antecedente.

Para o modelo LSTM, as entradas foram estruturadas de forma sequencial, utilizando as vazões de até 5 dias anteriores e os valores de precipitação dos dias antecedentes e do dia corrente.

Em todos os modelos, a variável de saída foi definida como a vazão específica média diária correspondente ao dia a ser estimado.

### 3.4.4. Normalização dos dados

Para realizar a modelagem por meio de modelos de Machine Learning foi necessária a normalização pelo método Min–Max, descrito na equação 13, de forma que todas as variáveis fossem reescalonadas para um intervalo comum, geralmente entre 0 e 1 (Muhammad Ali, 2022).

A padronização é importante para algoritmos baseados em distâncias, gradientes ou funções de ativação, como Redes Neurais e métodos de otimização iterativa, pois contribui para

uma convergência mais rápida e estável, além de reduzir problemas numéricos durante o ajuste dos parâmetros. Ao colocar todas as variáveis na mesma escala, garante-se que o aprendizado do modelo seja guiado pelos padrões reais dos dados, e não pela diferença de ordem de grandeza entre os atributos, resultando em modelos mais robustos e com melhor capacidade de generalização (Rahmad Ramadhan; Anne Mudya, 2024).

$$X_{norm} = \frac{X_{obs} - X_{min}}{X_{max} - X_{min}} \tag{13}$$

onde:

$X_{norm}$ = vazão normalizada, adimensional;
$X_{obs}$ = vazão observada (L $s^{-2}$ $m^{-2}$);
$X_{min}$ = vazão mínima observada (L $s^{-2}$ $m^{-2}$); e
$X_{max}$ = vazão máxima observada (L $s^{-2}$ $m^{-2}$).

## 3.5. Modelos de Machine Learnig (ML)

Os modelos utilizados neste estudo foram implementados e treinados na linguagem Python, por meio do ambiente Google Colaboratory (Google Colab). Foram avaliados os algoritmos Random Forest (RF), Extreme Gradient Boosting (XGBoost - XGB), Perceptron Multilayer (MLP) e Long Short-Term Memory (LSTM), considerando duas estratégias de treinamento e preenchimento das falhas. Para diferenciá-las, os modelos treinados pela metodologia sequencial receberam a identificação 1, enquanto aqueles treinados pela metodologia tradicional receberam a identificação 2. A LSTM foi mantida sem identificação numérica, pois foi treinada apenas pela metodologia sequencial.

Antes da aplicação dos modelos, foram realizados testes preliminares com diferentes configurações de parâmetros e arquiteturas, buscando identificar estruturas adequadas para estimativa das vazões especificas diárias. A partir desses testes, foram selecionadas as configurações que apresentaram melhor desempenho, as quais foram posteriormente mantidas nas diferentes bacias hidrográficas, cenários de falhas e repetições. As configurações finais adotadas para cada modelo são descritas nas subseções seguintes.

A adoção de duas estratégias foi necessária principalmente em função das características do modelo LSTM e da própria distribuição das falhas ao longo das séries. Como o objetivo do estudo consiste no preenchimento de dados ausentes localizados em diferentes posições da série temporal, o procedimento de modelagem deveria permitir a estimativa dessas falhas sem comprometer a estrutura temporal dos dados. Diferentemente dos modelos RF, XGBoost e MLP, que podem ser ajustados a partir de conjuntos de dados organizados de forma independente, a LSTM apresenta uma estrutura recorrente, desenvolvida para aprender dependências temporais existentes ao longo de uma sequência (Hochreiter; Schmidhuber, 1997; Kratzert et al., 2018).

A continuidade e a ordem cronológica das informações apresentam papel importante durante o treinamento, uma vez que os estados anteriores da série contribuem para a estimativa dos valores subsequentes. Essa característica exigiu a adoção de um procedimento de treinamento que respeitasse a progressão temporal da série, permitindo o preenchimento de falhas distribuídas em diferentes períodos sem utilizar, durante o ajuste do modelo, informações posteriores ao intervalo que estava sendo preenchido.

Na primeira estratégia, denominada como metodologia sequencial, o treinamento e o preenchimento foram realizados progressivamente ao longo da série temporal. Os modelos treinados por essa abordagem foram identificados como RF1, XGB1 e MLP1, além da LSTM. Inicialmente, identificou-se o primeiro bloco de falhas presente na série. Todos os dados disponíveis anteriormente ao início desse bloco foram então utilizados para o treinamento dos quatro modelos. Após o treinamento, os modelos foram empregados para estimar as vazões correspondentes aos dias que compunham o primeiro período de ausência de dados.

Concluído o preenchimento desse primeiro bloco, avançou-se cronologicamente na série até a identificação do segundo período de falhas. Nesse momento, uma nova base de treinamento foi construída utilizando toda a série disponível anteriormente ao início da segunda falha, incluindo os valores que haviam sido previamente estimados no primeiro bloco. A partir dessa nova série, outros modelos foram treinados e utilizados para o preenchimento do segundo bloco de falhas. O mesmo procedimento foi repetido sucessivamente para todos os blocos de falhas distribuídos ao longo da série temporal.

Assim, os modelos eram reajustados a cada novo período de falha, sempre utilizando somente as informações temporalmente anteriores ao bloco que seria preenchido. Consequentemente, à medida que se avançava na série, os valores estimados nos blocos anteriores passavam a integrar a base disponível para o treinamento dos modelos seguintes.

Esse procedimento pode ocasionar uma propagação parcial dos erros das estimativas anteriores, uma vez que valores simulados passam a ser utilizados juntamente com os valores

observados na construção dos modelos subsequentes. Entretanto, essa condição foi aplicada igualmente aos quatro modelos, permitindo que RF1, XGB1, MLP1 e LSTM fossem avaliados sob as mesmas condições de disponibilidade de informações. Portanto, a restrição decorrente da estrutura recorrente do modelo LSTM foi imposta também aos demais modelos.

Além dessa abordagem, foi adotada uma segunda estratégia, denominada metodologia tradicional, aplicada aos modelos identificados como RF2, XGB2 e MLP2. Essa abordagem é normalmente utilizada em trabalhos como os de (Kartal, 2024; Kedam et al., 2024; Kumar et al., 2023; Mushtaq et al., 2024; Z. Zhang et al., 2026). Nessa configuração, os dados destinados ao treinamento foram previamente separados e utilizados para o treino de um único modelo de cada algoritmo. Após o treinamento, os modelos ajustados foram mantidos e aplicados diretamente ao preenchimento dos diferentes blocos de falhas presentes no conjunto de teste, sem a necessidade de realizar um novo treinamento antes de cada período de ausência. Em seguida, as vazões estimadas para os dias originalmente considerados como falhos foram comparadas aos respectivos valores observados, que haviam sido temporariamente removidos apenas para fins de avaliação do desempenho.

A utilização das duas metodologias permitiu avaliar não apenas as diferenças entre os algoritmos, mas também o efeito da estratégia de treinamento sobre a capacidade de preenchimento das séries de vazão. Assim, a metodologia sequencial foi representada por RF1, XGB1, MLP1 e LSTM, enquanto a metodologia tradicional foi representada por RF2, XGB2 e MLP2. A LSTM não foi incluída na segunda abordagem em razão da necessidade de preservar a estrutura sequencial considerada em sua modelagem.

Para garantir a comparabilidade dos resultados, os períodos de falhas simulados foram mantidos exatamente nas mesmas posições para todos os modelos e para as duas metodologias. Assim, dentro de cada bacia hidrográfica, quantidade de dias consecutivos de falha e repetição, todos os algoritmos foram avaliados nos mesmos dias originalmente removidos da série. Portanto, as diferenças observadas no desempenho estiveram relacionadas às características dos modelos e às estratégias de treinamento, e não à utilização de diferentes períodos de ausência de dados.

Ressalta-se, ainda, que o preenchimento de cada bloco de falhas foi realizado de forma sequencial, em razão da utilização da vazão do dia anterior como uma das covariáveis de entrada. Para a estimativa do primeiro dia de cada bloco, foi utilizada como variável antecedente a vazão observada no dia imediatamente anterior ao início da falha. Após a estimativa da vazão do primeiro dia de falha, esse valor foi utilizado como vazão antecedente para estimar o dia seguinte, uma vez que a vazão observada correspondente não estava disponível, da mesma forma, a vazão estimada

para o segundo dia foi empregada na estimativa do terceiro, e assim sucessivamente, até o preenchimento completo do período de falha.

### 3.5.1. Random Forest (RF)

Para a construção do modelo Random Forest (RF) foram utilizadas 500 árvores de decisão, buscando obter maior estabilidade nas estimativas a partir da combinação das previsões produzidas pelas diferentes árvores que compõem o modelo. A profundidade máxima das árvores foi definida em 10, limitando seu crescimento e, consequentemente, controlando a complexidade do modelo. Além disso, a semente aleatória foi fixada em 42, permitindo a reprodutibilidade do procedimento de treinamento.

Durante o ajuste, as árvores foram construídas a partir das informações disponíveis no conjunto de treinamento, e a estimativa final da vazão foi obtida pela combinação das previsões individuais produzidas pelas árvores que constituem a floresta. Todos os modelos RF foram construídos utilizando a mesma configuração de parâmetros ao longo das diferentes bacias hidrográficas, cenários de falhas e repetições avaliadas.

### 3.5.2. Extreme Gradient Boosting (XGBoost – XGB)

Para a construção do modelo Extreme Gradient Boosting (XGBoost – XGB) foram também utilizadas 500 árvores de decisão, ajustadas de forma sequencial, de modo que cada nova árvore contribuísse para melhorar as estimativas produzidas pelo conjunto. A taxa de aprendizado foi definida em 0,05, controlando a contribuição de cada árvore adicionada ao modelo, enquanto a profundidade máxima foi fixada em 6, limitando o crescimento das árvores e, consequentemente, sua complexidade (Chen; Guestrin, 2016).

Durante o ajuste, foram utilizados 80% das observações (subsample = 0,8) e 80% das variáveis de entrada (colsample_bytree = 0,8) na construção de cada árvore. Para a função objetivo, foi adotado o erro quadrático para regressão (reg:squarederror), uma vez que a variável a ser estimada, a vazão, é contínua. A semente aleatória foi mantida fixa em 42, permitindo a reprodutibilidade do procedimento. Assim, os modelos XGB foram ajustados utilizando a mesma configuração de parâmetros ao longo das diferentes simulações realizadas.

### 3.5.3. Multilayer Perceptron (MLP)

Para a construção do modelo Perceptron Multicamadas (Multilayer Perceptron – MLP) foi utilizada uma arquitetura composta por duas camadas ocultas, contendo 128 e 64 neurônios, respectivamente. A função de ativação ReLU (Rectified Linear Unit) foi aplicada às camadas ocultas, permitindo ao modelo representar relações não lineares entre as variáveis de entrada e a vazão.

O ajuste dos pesos da rede foi realizado por meio do otimizador Adam, utilizado para atualizar os parâmetros do modelo ao longo do processo de aprendizagem. O número máximo de iterações foi definido em 500, estabelecendo o limite de ciclos disponíveis para a convergência do treinamento. Além disso, a semente aleatória foi fixada em 42, de forma a garantir a reprodutibilidade do ajuste. Assim, todos os modelos MLP foram construídos utilizando a mesma arquitetura e configuração de treinamento nas diferentes simulações realizadas.

### 3.5.4. Long Short-Term Memory

Para a construção do modelo Long Short-Term Memory (LSTM) foi utilizada uma arquitetura recorrente composta por uma camada LSTM com 64 unidades, responsável pelo processamento das informações ao longo da sequência temporal. Para cada estimativa, foram consideradas sequências correspondentes aos 30 dias anteriores (timesteps = 30), utilizando como informações de entrada as séries de vazão e precipitação. Dessa forma, o modelo foi estruturado para considerar as condições antecedentes na estimativa da vazão do dia seguinte.

Após a camada LSTM, foi aplicada uma camada de Dropout de 20%, com a finalidade de reduzir o risco de sobreajuste durante o treinamento. Em seguida, foi utilizada uma camada densa composta por 32 neurônios e função de ativação ReLU, seguida pela camada de saída, formada por um único neurônio também com ativação ReLU, responsável pela estimativa da vazão.

O treinamento da rede foi realizado utilizando o otimizador Adam, com taxa de aprendizado de 0,0005, e o erro quadrático médio (MSE) como função de perda. Foram definidas 50 épocas de treinamento, com tamanho de lote (batch size) igual a 32. Assim, todos os modelos LSTM foram construídos seguindo a mesma arquitetura e as mesmas configurações de treinamento ao longo das diferentes simulações realizadas.

## 3.6. Avaliação do Desempenho dos Modelos

### 3.6.1. Análise das métricas de desempenho entre as repetições

O desempenho dos modelos na transformação chuva–vazão foi avaliado exclusivamente no período de teste, por meio da comparação entre as vazões simuladas e observadas em cada bacia hidrográfica. A avaliação do desempenho dos modelos foi realizada por meio das seguintes métricas estatísticas: Raiz do Erro Quadrático Médio (RMSE), Erro Percentual Absoluto Médio Simétrico (SMAPE), Viés Percentual (PBIAS), Coeficiente de Eficiência de Kling–Gupta (KGE) e Coeficiente de Determinação ($R^2$).

O Erro Quadrático Médio (Root Mean Square Error – RMSE; Equação 14) foi utilizado para quantificar a magnitude dos erros entre as vazões estimadas e observadas. Por elevar os erros ao quadrado antes do cálculo da média, o RMSE atribui maior peso aos maiores desvios, tornando-se útil para identificar estimativas que apresentam erros de maior magnitude. Valores menores de RMSE indicam maior proximidade entre os valores estimados e observados, enquanto valores elevados representam maiores discrepâncias entre as séries (Hodson, 2022; Karunasingha, 2022; Willmott; Matsuura, 2005).

A análise do erro relativo foi realizada por meio do Erro Percentual Absoluto Médio Simétrico (Symmetric Mean Absolute Percentage Error – SMAPE; Equação 15). O SMAPE apresenta robustez a outliers e reduz a influência do ruído nos dados. Essa métrica foi escolhida por contornar limitações dos erros percentuais convencionais, especialmente em situações com valores observados nulos ou próximos de zero, comuns em séries de vazão durante períodos de seca. O SMAPE apresenta valores que vão de 0 a 200 (Armstrong; Forecasting, 1985; Tayman; Swanson, 1999).

O Viés Percentual (Percent Bias – PBIAS; Equação 16), foi utilizado para avaliar a tendência sistemática dos modelos em superestimar ou subestimar as vazões observadas. Essa métrica expressa, em termos percentuais, a diferença acumulada entre os valores simulados e observados, permitindo avaliar o viés associado ao volume de vazão reproduzido ao longo do período (Towler et al., 2023). Valores próximos de zero indicam menor viés, enquanto valores com maior magnitude representam maior discrepância entre as vazões simuladas e observadas. O sinal do PBIAS permite ainda identificar a direção do viés, indicando tendência de superestimação ou subestimação das vazões, conforme a convenção adotada para o cálculo (Evans et al., 2018). Com isso, o PBIAS fornece uma medida específica da tendência sistemática das estimativas, aspecto que não é diretamente representado por métricas baseadas apenas na magnitude dos erros.

O Coeficiente de Eficiência de Kling–Gupta (Kling–Gupta Efficiency – KGE; Equação 19), proposto por Gupta et al. (2009), por se tratar de uma métrica recomendada em avaliações hidrológicas que integra, em um único indicador, informações relacionadas à correlação linear, ao viés e à variabilidade relativa entre séries estimadas e observadas. Sua adoção se deu pelo fato de que métricas tradicionais podem apresentar bom desempenho na reprodução da dinâmica temporal da vazão, mas subestimar a variabilidade e distorcer o volume simulado, especialmente em bacias com diferentes regimes hidrológicos. Assim, o KGE fornece uma avaliação mais equilibrada do desempenho do modelo, sendo útil para identificar compensações entre os diferentes componentes do erro (Pool et al., 2018).

Por fim, incluiu-se o Coeficiente de Determinação ($R^2$), como descrito na equação, utilizado para avaliar o desempenho dos modelos na simulação das vazões, uma vez que quantifica a proporção da variabilidade dos dados observados explicada pelos valores simulados, indicando o grau de proximidade entre as vazões observadas e estimadas (Berggren, 2024; Eugénia Graça Martins, 2018). Embora valores elevados de $R^2$ indiquem maior correspondência entre a variação dos dados observados e estimados, essa métrica, quando utilizada isoladamente, não permite identificar a magnitude dos erros nem possíveis tendências de superestimação ou subestimação. Por esse motivo, o $R^2$ foi avaliado em conjunto com as demais métricas, possibilitando uma análise mais completa do desempenho dos modelos.

Além da avaliação dos valores das métricas, foram elaborados boxplots considerando os resultados obtidos nas diferentes repetições de cada combinação entre modelo, bacia e quantidade de falhas. Essa análise foi realizada com o objetivo de verificar a variabilidade do desempenho entre as repetições e identificar se os modelos mantiveram comportamento semelhante diante das diferentes posições dos períodos de falha nas séries. A amplitude das distribuições, a posição das medianas e a ocorrência de valores discrepantes foram utilizadas para avaliar a estabilidade dos resultados e a sensibilidade dos modelos às diferentes repetições.

$$RMSE = \sum_{i=1}^{n} \frac{\sqrt{(Yo - Ye)^2}}{n} \tag{14}$$

onde:

RMSE = erro médio absoluto (L $s^{-2}$ $m^{-2}$);
Yo = vazão observada (L $s^{-2}$ $m^{-2}$;

Ye = vazão simulada (L $s^{-2}$ $m^{-2}$);

N = número de observações, adimensional.

$$SMAPE = \frac{1}{n}\sum_{i=1}^{n}\frac{|Yo - Ye|}{\frac{|Yo| + |Ye|}{2}} \quad (15)$$

onde:

SMAPE = Erro percentual absoluto médio simétrico, adimensional.

$$PBIES = 100 * \frac{\sum_{i=1}^{n}(Yo - Ye)^2}{\sum_{i=1}^{n} Yo} \quad (16)$$

onde:

PBIES = Viés Percentual, adimensional.

Yo = vazão observada (L $s^{-2}$ $m^{-2}$;

Ye = vazão simulada (L $s^{-2}$ $m^{-2}$);

N = número de observações, adimensional.

$$\alpha = \frac{\sigma_{sim}}{\sigma_{obs}} \quad (17)$$

$$\beta = \frac{\mu_{sim}}{\mu_{obs}} \quad (18)$$

$$\text{KGE} = 1 - \sqrt{(r-1)^2 + (\alpha - 1)^2 + (\beta - 1)^2} \quad (19)$$

onde:

$\sigma_{obs}$ = Desvio padrão da vazão observada, (L $s^{-2}$ $m^{-2}$);

$\sigma_{sim}$ = Desvio padrão da vazão simulada (L $s^{-2}$ $m^{-2}$);

$\mu_{sim}$ = Vazão média simulada (L $s^{-2}$ $m^{-2}$);

$\mu_{obs}$ = Vazão édia observada (L $s^{-2}$ $m^{-2}$);

$r$ = Coeficíente de correlação, adimensional;

$\alpha$ = Razão entre os desvio padrão, adimensional;

$\beta$ = Razão entre as médias, adimensional; e

$KGE$ = Coeficiente de Eficiência de Kling–Gupta.

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(Yo - Ye)^2}{\sum_{i=1}^{n}(Yo - \overline{Ye})^2} \quad (20)$$

onde:

$R^2$ = coeficiente de determinação, adimensional; e

$\overline{Yo}$ = vazão média observada (L $s^{-2}$ $m^{-2}$).

$\overline{Ye}$ = vazão média simulada (L $s^{-2}$ $m^{-2}$).

### 3.6.2. Seleção das repetições representativas

Como os períodos de falha foram distribuídos aleatoriamente ao longo da série, cada combinação entre bacia e quantidade de dias consecutivos de falha tiveram 10 repetições, podendo resultar em diferentes níveis de desempenho em função da posição em que as falhas ocorreram. Para as análises seguintes, foi selecionada uma única repetição que representasse o comportamento médio de cada bacia e cenário de falhas, evitando escolher apenas aquela com melhor ou pior desempenho.

A seleção considerou conjuntamente os resultados dos modelos RF1, RF2, XGB1, XGB2, MLP1, MLP2 e LSTM, utilizando as métricas RMSE, SMAPE, PBIAS, KGE e R². A utilização conjunta das métricas foi adotada porque cada uma avalia diferentes características do erro e do ajuste entre vazões observadas e estimadas (M. P. Clark et al., 2017; Nearing et al., 2024).

Inicialmente, as métricas foram padronizadas em escores z (equação 21), subtraindo-se de cada valor a média da respectiva métrica e dividindo-se pelo seu desvio-padrão. Isso permitiu colocar as métricas em uma mesma escala, facilitando a comparação entre os indicadores e reduzindo a influência das diferenças de magnitude entre eles (Andrade, 2021).

Em seguida, cada repetição foi representada por um vetor contendo as métricas padronizadas de todos os modelos. Para cada combinação de bacia e quantidade de falhas, foi determinado um vetor central a partir da média das 10 repetições. A distância Euclidiana entre

cada repetição e esse centro foi então calculada como medida de proximidade no espaço multivariado (Gillard et al., 2022; Warton et al., 2012). A repetição com a menor distância foi considerada a mais representativa do comportamento médio do conjunto.

Assim, a repetição com a menor distância Euclidiana em relação ao vetor médio foi considerada representativa, por apresentar o conjunto de resultados mais próximo do comportamento médio das 10 repetições. As séries selecionadas foram utilizadas nas análises hidrológicas a seguir.

$$z = \frac{x - \mu}{\sigma} \tag{21}$$

onde:

z = escores z, adimensional;
x = valor a ser reescalado;
$\mu$ = média dos valores;
$\sigma$ = desvio padrão dos valores.

### 3.6.3. Análise das curvas de permanência

As curvas de permanência foram utilizadas para avaliar a capacidade dos modelos em representar o regime de vazões. Essas curvas relacionam a magnitude da vazão à frequência com que determinado valor é igualado ou excedido, permitindo analisar o comportamento das vazões ao longo de diferentes faixas do regime hidrológico (Leong; Yokoo, 2021).

As curvas foram construídas utilizando as repetições representativas selecionadas para cada bacia e para os períodos de 10, 30, 60, 90 e 120 dias de falha. Em cada condição, as curvas geradas a partir das vazões estimadas foram comparadas à curva observada, permitindo verificar a capacidade dos modelos em reproduzir a distribuição das vazões.

A partir das curvas de permanência, foram determinadas as vazões de referência $Q_{10}$, $Q_{50}$, $Q_{90}$, $Q_{95}$ e $Q_{99}$, correspondentes às vazões igualadas ou excedidas em 10%, 50% 90%, 95% e 99% do tempo, respectivamente. Para isso, as vazões foram ordenadas em ordem decrescente e associadas às respectivas probabilidades de excedência, permitindo obter as vazões correspondentes a cada percentual da curva de permanência (Rudra; Alam, 2023; Sarigil et al., 2024).

Posteriormente, para cada bacia e período de falha, foi calculada a razão entre a vazão de referência estimada e a observa (Qest/Qobs). Valores próximos de 1 indicam maior concordância com as observações, enquanto valores superiores ou inferiores a 1 indicam, respectivamente, superestimação ou subestimação.

### 3.6.4. Regime hidrológico mensal

O regime hidrológico mensal foi analisado utilizando as vazões observadas e as vazões preenchidas pelos modelos. Para cada combinação entre bacia hidrográfica e período de falha (10, 30, 60, 90 e 120 dias), foram calculadas as vazões médias mensais a partir da repetição representativa previamente selecionada.

A agregação dos dados em escala mensal teve como objetivo caracterizar o comportamento sazonal das vazões e verificar a capacidade dos modelos em reproduzir a variação intra-anual observada, especialmente a alternância entre os períodos de maior e menor disponibilidade hídrica. A comparação entre as séries observadas e estimadas foi realizada considerando tanto a magnitude das vazões médias quanto o período de ocorrência dos máximos e mínimos sazonais.

### 3.6.5. Teste de Friedman – Comparação entre os modelos

Para verificar se as diferenças de desempenho entre os modelos foram estatisticamente significativas, foi aplicado o teste não paramétrico de Friedman. Esse teste é indicado para a comparação de três ou mais grupos relacionados, baseando-se na ordenação dos resultados em postos e apresentando menor dependência de pressupostos relacionados à normalidade dos dados (Friedman, 1937; Kim et al., 2026).

O teste foi aplicado separadamente para cada bacia hidrográfica e período de falha, comparando os modelos RF1, RF2, XGB1, XGB2, MLP1, MLP2 e LSTM com base nos resíduos entre as vazões observadas e estimadas. Os resíduos foram obtidos pela diferença entre os valores observados e estimados, permitindo avaliar o comportamento dos erros produzidos pelos modelos e identificar possíveis desvios sistemáticos nas estimativas (Z. Zhang et al., 2026).

Para cada condição analisada, os resultados dos modelos foram organizados de forma pareada e convertidos em postos. A hipótese nula ($H_0$) estabeleceu que não existiam diferenças significativas entre os modelos, enquanto a hipótese alternativa ($H_1$) considerou que pelo menos um deles apresentava comportamento distinto. Adotou-se nível de significância de 5% ($\alpha = 0,05$), sendo $H_0$ rejeitada quando $p < 0,05$.

Quando o teste de Friedman indicou diferença significativa, foram realizadas comparação par a par pelo teste de Wilcoxon para amostras pareadas, permitindo identificar quais modelos diferiram entre si. Nessa etapa, os valores de *p* foram ajustados para múltiplas comparações, reduzindo o efeito do aumento do número de testes sobre a ocorrência de resultados significativos (Wilcoxon et al., 1963).

### 3.6.6. Hidrogramas dos modelos mais representativos

Para complementar a avaliação dos resultados, foram elaborados hidrogramas considerando os modelos que apresentaram os melhores desempenhos nas análises anteriores. A seleção de um número reduzido de modelos teve como objetivo facilitar a comparação visual e permitir uma análise mais clara da capacidade de cada método em reproduzir a magnitude e a variação das vazões observadas.

Os hidrogramas foram construídos utilizando somente os valores correspondentes aos períodos em que as falhas foram simuladas, ou seja, os dias efetivamente preenchidos pelos modelos. Como esses blocos estavam distribuídos em diferentes posições da série de teste, sua representação utilizando diretamente as datas resultaria em trechos descontínuos e separados no gráfico. Dessa forma, os valores estimados foram organizados sequencialmente, mantendo sua ordem de ocorrência ao longo da série, e apresentados lado a lado.

Assim, o eixo x não representa uma sequência contínua de datas, mas a posição dos dias preenchidos ao longo do período de teste, contada desde o início até o final da série.

# 4. RESULTADOS E DISCUSSÃO

## 4.1. Avaliação da influência das repetições feitas em cada bacia e quantidade de dias seguidos de falhas nas métricas de desempenho

A análise das repetições realizadas para cada combinação entre bacia, modelo e quantidade de dias seguidos de falhas da falha (Figuras 3, 4, 5, 6 e 7) mostrou baixa variabilidade no desempenho dos modelos na maior parte dos casos, principalmente para as falhas de 10 e 30 dias consecutivos. Para as falhas de 60, 90 e 120 dias, algumas combinações apresentaram maior amplitude interquartil, indicando que a posição em que a falha foi inserida na série pode influenciar o desempenho dos modelos. Ainda assim, mesmo para as falhas mais longas, vários modelos apresentaram resultados pouco dispersos entre as repetições. Portanto, o aumento da extensão da falha não resultou, necessariamente, em maior variabilidade dos resultados entre as diferentes posições avaliadas.

Essas variações podem estar relacionadas ao comportamento temporal das vazões e ao período hidrológico atingido pela falha. Blocos localizados durante a recessão do hidrograma tendem a apresentar maior persistência temporal, uma vez que a redução das vazões ocorre de forma mais gradual e está associada à liberação da água armazenada na bacia e à contribuição do escoamento de base. Em contraste, falhas que abrangem a ascensão do hidrograma e eventos de cheia são mais difíceis de reconstruir, pois envolvem variações rápidas da vazão e maior amplitude entre dias consecutivos. Além disso, eventos de elevada vazão são menos frequentes na série, aumentando a dificuldade de representação desses padrões pelos modelos (Arriagada et al., 2021).

Esse comportamento também interfere diretamente nas métricas de erro. O RMSE atribui maior peso aos maiores resíduos, conforme destacado por Hodson (2022), de modo que falhas coincidentes com períodos de elevada vazão podem aumentar consideravelmente os valores dessa métrica quando os modelos não reproduzem adequadamente os picos do hidrograma, conforme discutido por Wang e Lu (2018). O SMAPE, por sua vez, expressa os erros em termos relativos e pode apresentar maior sensibilidade quando os valores observados e estimados possuem pequena magnitude (Chicco et al., 2021). Portanto, a variação observada entre as repetições não representa apenas instabilidade dos modelos, mas também diferenças nas condições hidrológicas atingidas pelas falhas.

Apesar dessa influência, houve um padrão consistente de desempenho. O XGB2 apresentou os resultados mais próximos das observações, com menores valores de RMSE e SMAPE, maiores $R^2$ e KGE e baixa dispersão entre as repetições. O RF2 apresentou, de maneira

geral, o segundo melhor desempenho, enquanto a LSTM ocupou uma posição intermediária. Em contraste, o MLP1 apresentou maiores erros e menores valores de $R^2$ e KGE. A diferença entre as estratégias tornou-se mais evidente com o aumento da extensão das falhas, principalmente entre XGB1 e XGB2. Nos modelos sequenciais, a vazão estimada no dia anterior é utilizada na estimativa subsequente, consequentemente, erros produzidos no início do bloco podem ser incorporados às entradas dos dias seguintes e propagados ao longo do período de falha.

O PBIAS evidenciou principalmente a direção dos erros entre os modelos. O XGB2 permaneceu próximo de zero na maior parte das combinações, indicando menor viés sistemático (Moriasi et al., 2007). RF1, XGB1 e MLP1 apresentaram maior ocorrência de valores negativos, indicando tendência de subestimação, enquanto o MLP2 apresentou maior tendência de superestimação. A subestimação pode estar relacionada à dificuldade dos modelos em reproduzir a magnitude das maiores vazões, principalmente quando os períodos de falha coincidem com eventos de maior resposta hidrológica.

Apesar dessas diferenças, a maioria dos valores de PBIAS permaneceu dentro de ±25%, faixa considerada satisfatória para simulações de vazão (Mei et al., 2022). Assim, embora existam diferenças na direção e magnitude do viés, grande parte das estimativas apresentou erros percentuais dentro de limites considerados aceitáveis.

O KGE reforçou esse padrão, com valores mais elevados e concentrados para o XGB2 e, em seguida, para o RF2. Como essa métrica considera conjuntamente componentes relacionados à correlação, ao viés e à variabilidade, seu comportamento indica que esses modelos apresentaram maior capacidade de preservar diferentes características da série observada, e não apenas reduzir a magnitude dos erros (Pool et al., 2018).

O $R^2$ mostrou que o XGB2 apresentou maior capacidade de acompanhar a variabilidade das vazões observadas, mantendo valores elevados mesmo com o aumento dos períodos de falha. O RF2 apresentou comportamento semelhante, enquanto a LSTM mostrou maior redução nos blocos mais extensos. Os menores valores foram observados para os modelos MLP, indicando menor capacidade de representar as variações da vazão ao longo da série. Em conjunto, os resultados indicam maior estabilidade e proximidade das observações para o XGB2, seguido pelo RF2 e pela LSTM, enquanto os demais modelos apresentaram maior sensibilidade ao aumento da extensão e à posição das falhas.

**Figura 3:** Variação dos valores de RMSE das 10 repetições feitas em cada bacia e quantidade de dias seguidos de falhas.

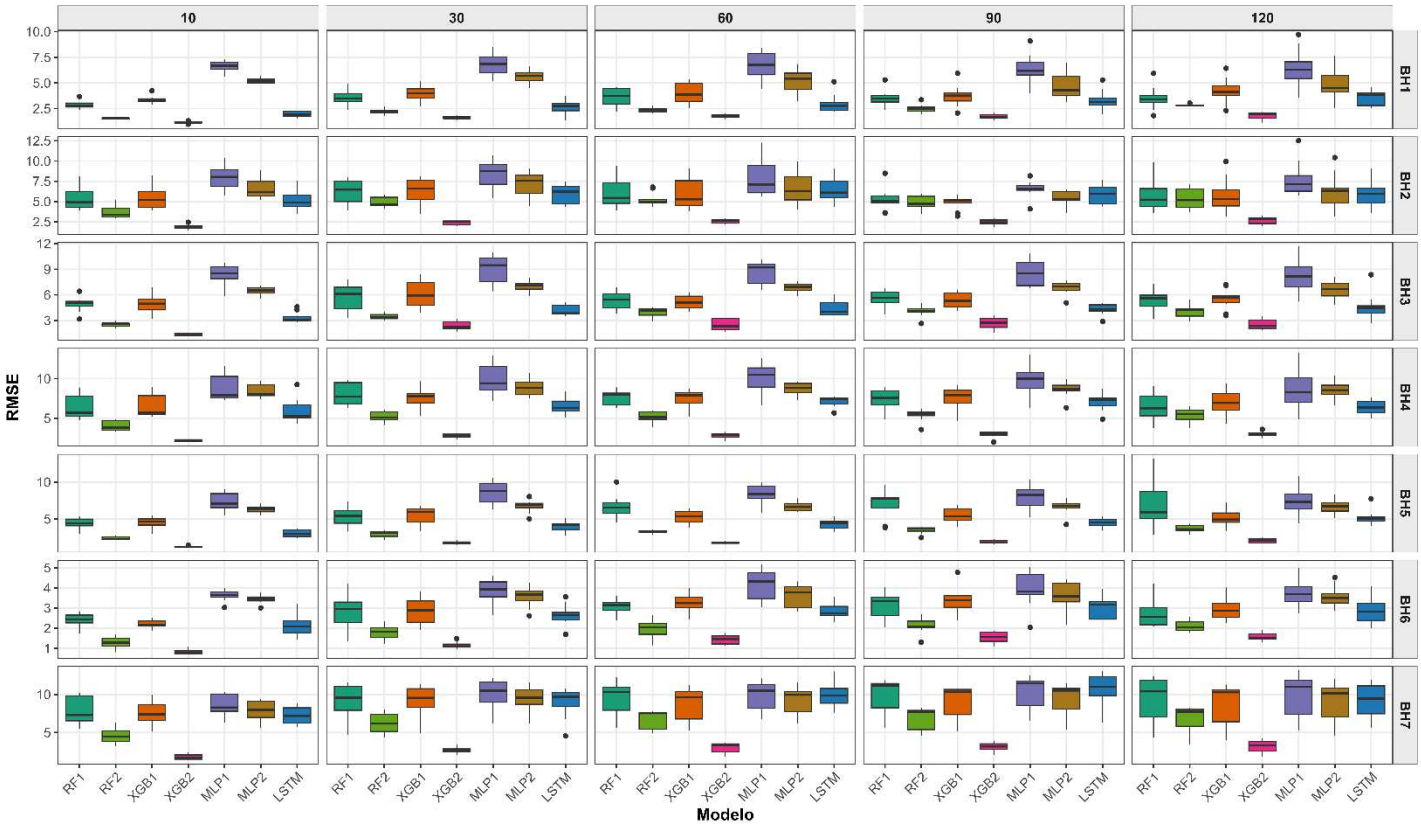


**Figura 4:** Variação dos valores de PBIAS das 10 repetições feitas em cada bacia e quantidade de dias seguidos de falhas

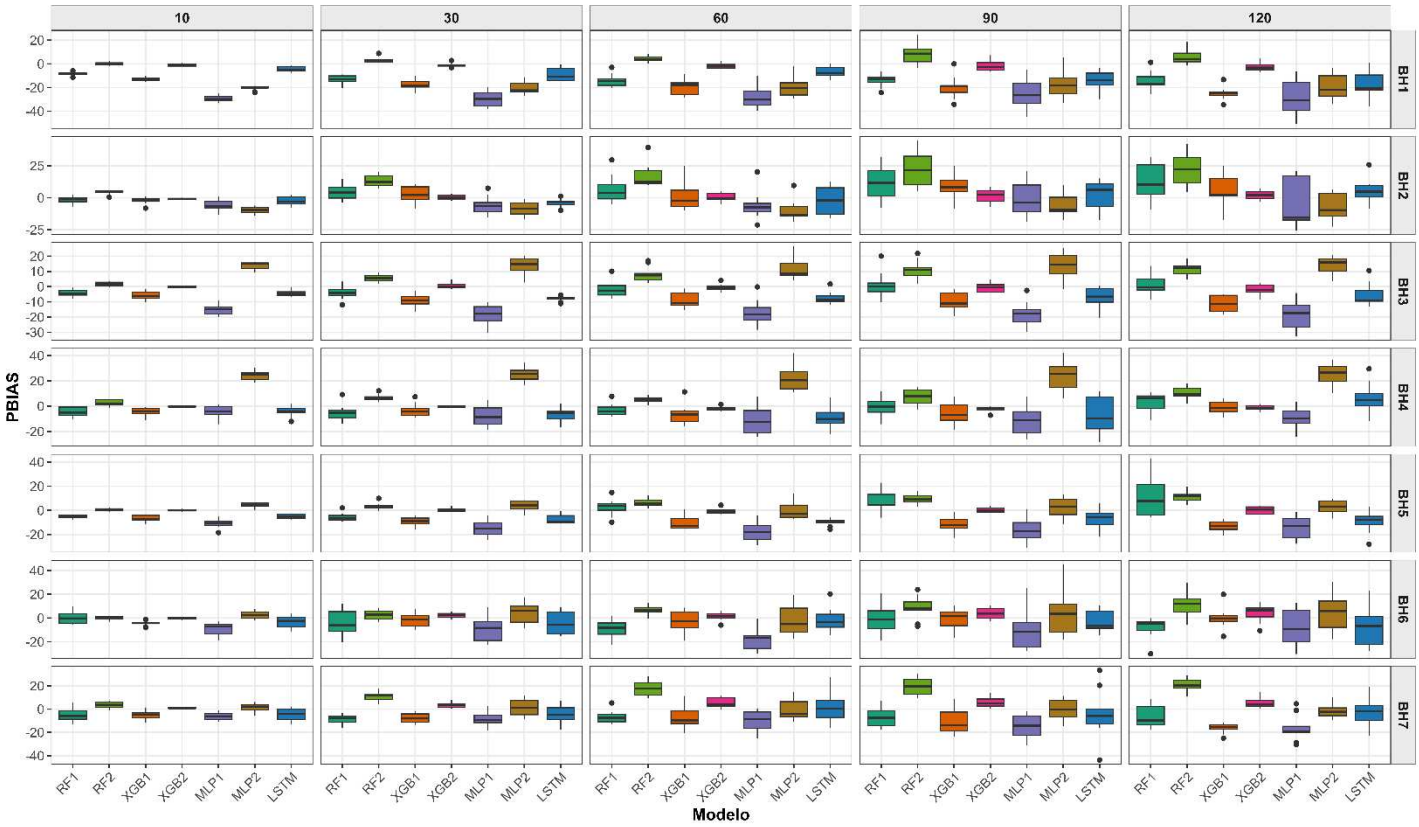


**Figura 5:** Variação dos valores de SMAPE das 10 repetições feitas em cada bacia e quantidade de dias seguidos de falhas

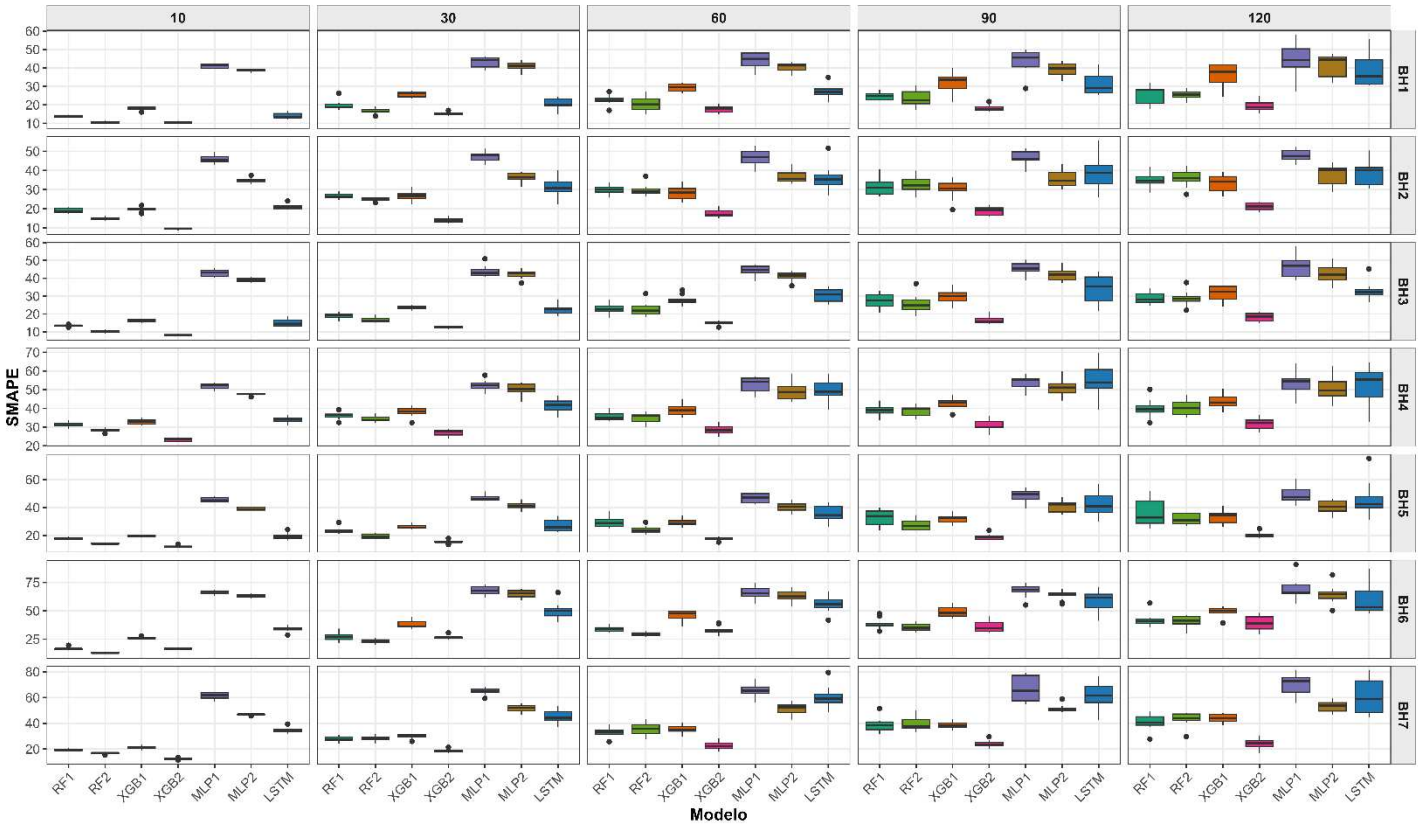


**Figura 6:** Variação dos valores de KGE das 10 repetições feitas em cada bacia e quantidade de dias seguidos de falhas

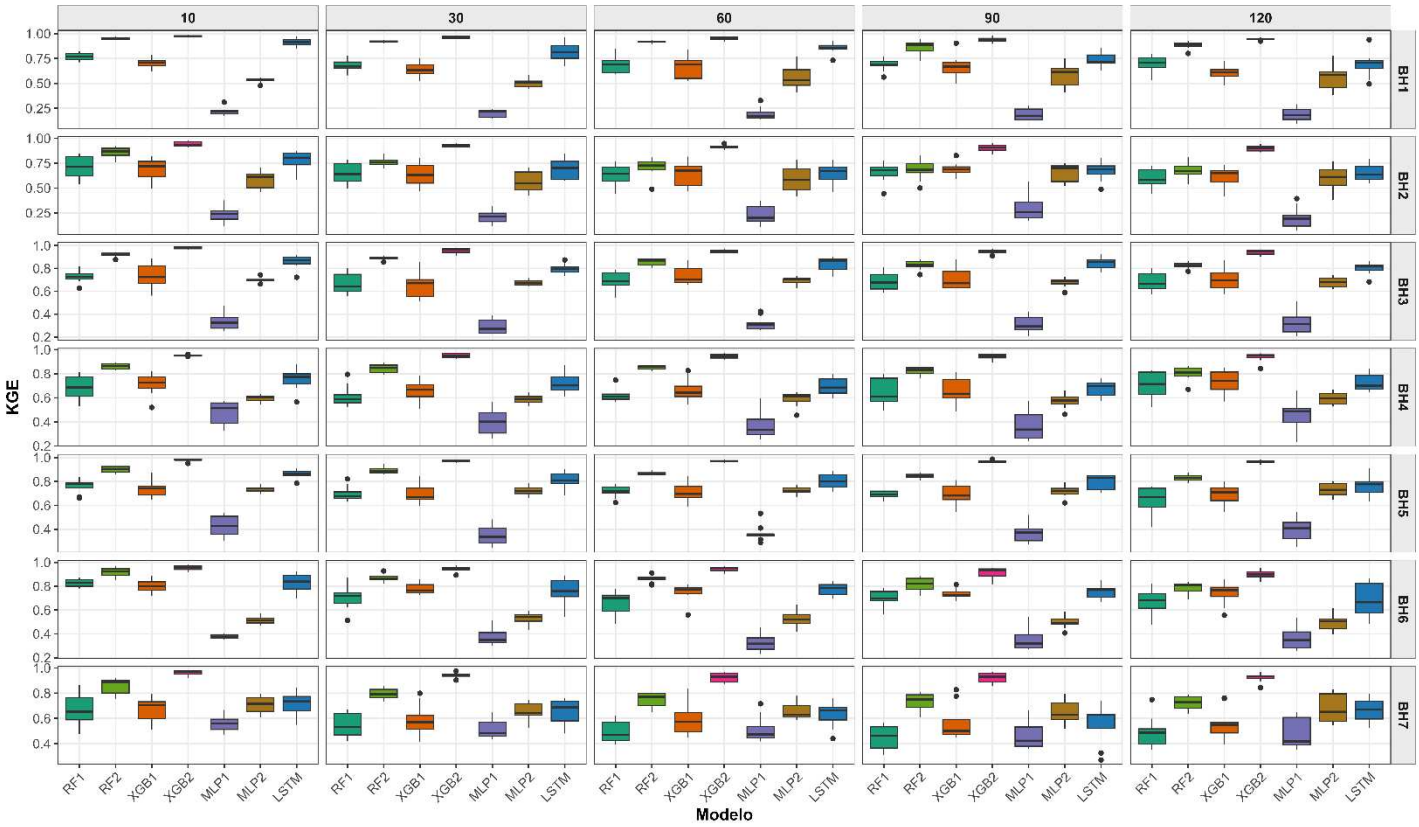
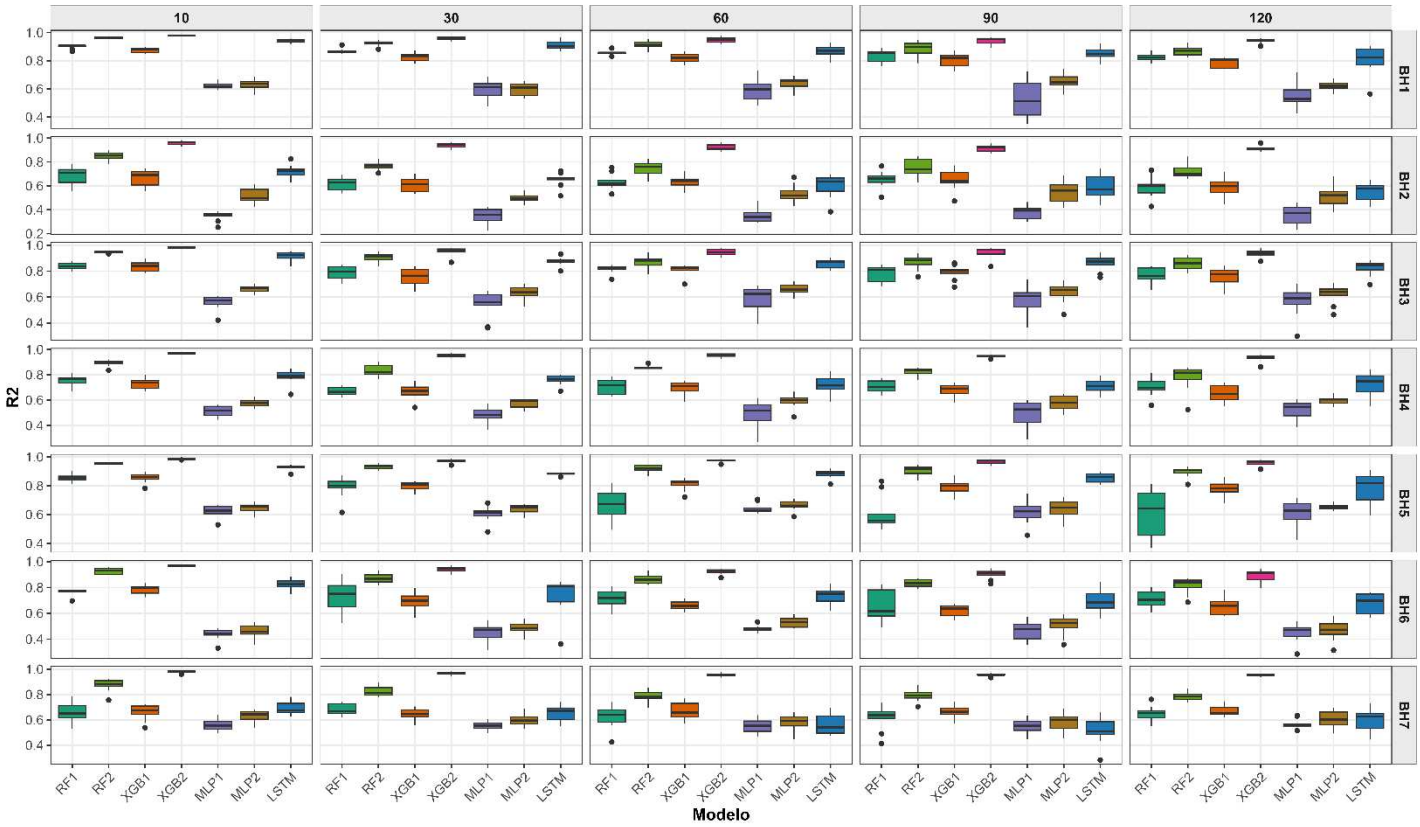


**Figura 7:** Variação dos valores de $R^2$ das 10 repetições feitas em cada bacia e quantidade de dias seguidos de falhas




## 4.2. Importância das Covariáveis nos Modelos

A importância das variáveis de entrada foi avaliada para os modelos Random Forest (RF), XGBoost (XGB) e Multilayer Perceptron (MLP), permitindo identificar a contribuição de cada covariável para as estimativas de vazão. Para a LSTM, essa análise não foi realizada, pois sua arquitetura recorrente processa as variáveis ao longo das sequências temporais e incorpora suas dependências na estrutura interna da rede, o que dificulta a obtenção de uma medida de importância das variáveis diretamente comparável àquela calculada para os demais modelos (Kratzert et al., 2018).

De forma geral, observou-se um comportamento consistente entre os modelos, bacias, quantidades de falhas e repetições. A vazão do dia anterior apresentou a maior importância relativa em praticamente todas as combinações avaliadas, seguida pelas precipitações de um e dois dias anteriores. Em contrapartida, a precipitação do dia da previsão e as precipitações de três, quatro e cinco dias antecedentes apresentaram baixa contribuição para a estimativa das vazões.

Resultados semelhantes foram encontrados por (Vilaseca et al., 2023), que identificaram a vazão antecedente como a variável mais importante para a previsão de vazões diárias utilizando Random Forest, seguida pelas variáveis de precipitação antecedente. Segundo os autores, a inclusão da vazão antecedente foi determinante para o desempenho do modelo.

A importância relativa das variáveis manteve-se praticamente constante entre as diferentes repetições e quantidades de falhas consecutivas. Embora tenham ocorrido pequenas variações nos valores de importância, a hierarquia das variáveis permaneceu praticamente inalterada, com a vazão do dia anterior ocupando a primeira posição, seguida pelas precipitações de um e dois dias anteriores.

## 4.3.Avaliação das Curvas de permanência nos diferentes cenários de ausência de dados

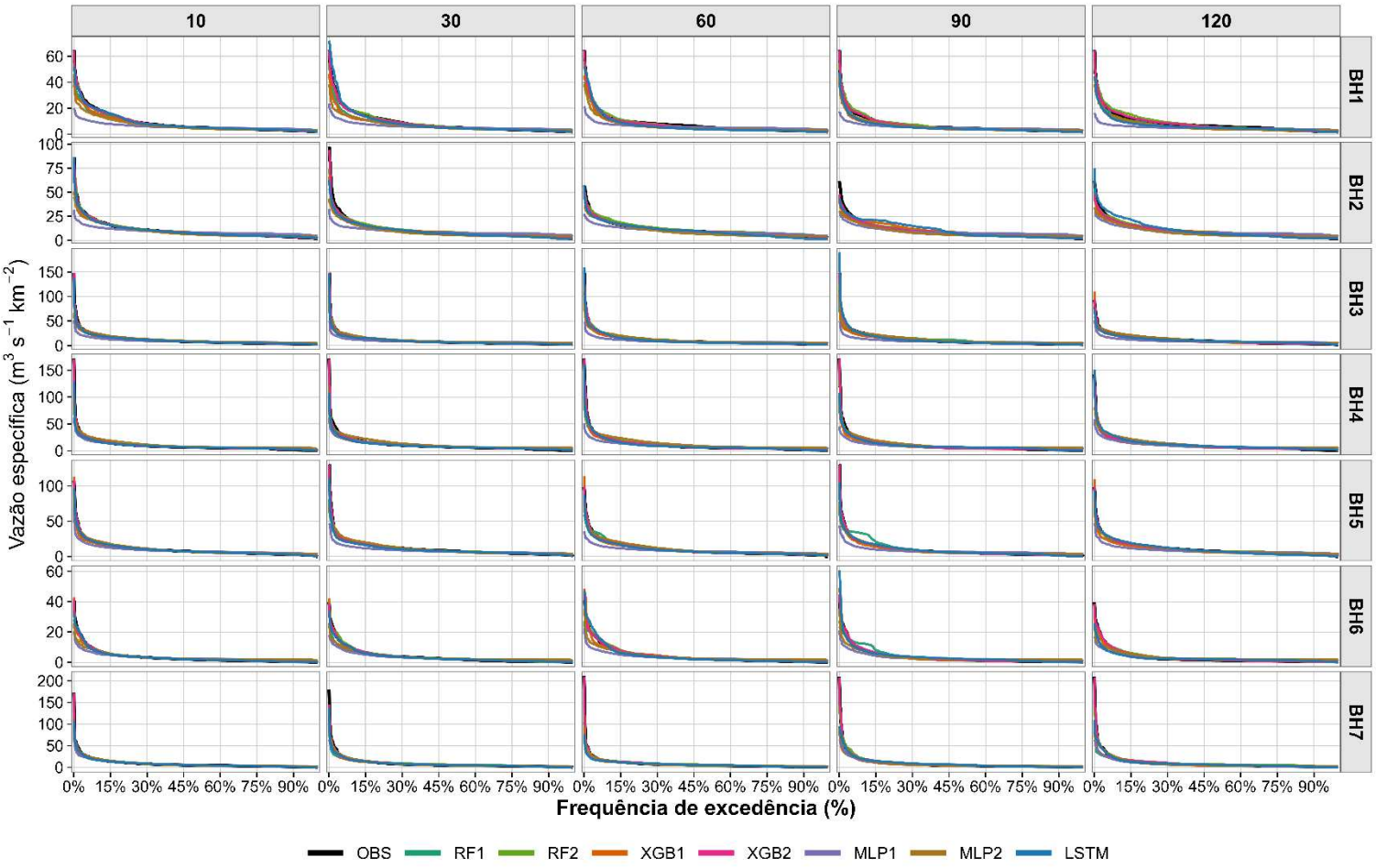
As curvas de permanência obtidas em cada bacia e quantidade de dias seguidos de falhas (Figura 8) mostraram que os modelos preservaram a distribuição das vazões específicas observadas nas diferentes bacias e períodos de falha. A proximidade entre as curvas indica que os valores preenchidos preservaram o comportamento global do regime de vazões, uma vez que a curva de permanência sintetiza a frequência e a variabilidade das vazões ao longo da série (Persiano et al., 2022).

As principais diferenças ocorreram nas frequências de excedência inferiores a, aproximadamente, 10%, correspondentes às maiores vazões. Essa faixa está associada à resposta mais rápida da bacia, sendo influenciada pela precipitação, pelas condições antecedentes de armazenamento e pela geração e propagação do escoamento. Por apresentarem maior magnitude e variabilidade temporal, esses eventos tendem a ser mais difíceis de reconstruir quando coincidem com os períodos de falha (Ghotbi et al., 2020). Nas frequências intermediárias e elevadas, por outro lado, houve maior aproximação entre as curvas estimadas e observadas, refletindo a maior persistência das componentes mais lentas do escoamento (Leong;  Yokoo, 2022).

Entre os modelos, o XGB2 apresentou a concordância mais consistente com a curva observada na maioria das bacias e períodos de falha. RF2 também acompanhou adequadamente a distribuição das vazões em diferentes cenários, enquanto os modelos sequenciais apresentaram maior tendência de subestimação das maiores vazões. A LSTM apresentou comportamento variável entre as bacias e os períodos de falha, sem tendência clara de aumento dos desvios com a duração das falhas. Em geral, manteve-se próxima à curva observada, com diferenças pontuais em algumas bacias e faixas de excedência.

O aumento do período de falha, de 10 para 120 dias consecutivos, não resultou em uma perda gradual de desempenho em todas as bacias. Mesmo nos cenários com falhas mais extensas, vários modelos mantiveram curvas próximas às observadas. Esse comportamento sugere que a capacidade de reconstrução das vazões está relacionada não apenas à duração da falha, mas também às características hidrológicas do período em que os dados foram removidos.

**Figura 8:** Curva de permanência para todos os modelos em cada bacia e quantidade de falhas




### 4.3.1. Análise dos valores de $Q_{10}$, $Q_{50}$, $Q_{90}$, $Q_{95}$ e $Q_{99}$

Diferentemente do comportamento geral observado nas curvas de permanência, a análise das vazões de referência mostrou que os desvios entre os valores estimados e observados aumentaram com a extensão das falhas (Figuras 9, 10, 11, 12 e 13). De modo geral, os blocos de 10 e 30 dias apresentaram maior proximidade com as vazões observadas, enquanto as falhas de

60, 90 e 120 dias resultaram em diferenças mais acentuadas, indicando perda de precisão dos modelos à medida que aumenta o período consecutivo sem dados.

Nas vazões máximas ($Q_{10}$), o XGB2 apresentou os melhores resultados, seguido pela LSTM e pelo RF2. Na $Q_{50}$, o XGB2 também se manteve próximo dos valores observados, enquanto a LSTM apresentou bom desempenho. Nas baixas vazões, principalmente em $Q_{95}$ e $Q_{99}$, as diferenças ficaram mais claras. O XGB2 permaneceu mais estável, enquanto RF1, RF2 e XGB1 tenderam a superestimar as vazões. Os MLPs apresentaram os maiores desvios. Já a LSTM mostrou um comportamento diferente, com tendência de subestimação em vários cenários, principalmente nos maiores períodos de falha.

A estabilidade do XGB2 pode estar relacionada à forma como o algoritmo aprende. No boosting, as árvores são construídas de maneira sucessiva, e cada nova árvore busca corrigir parte dos erros das anteriores (Chen; Guestrin, 2016). Esse processo pode ter favorecido a representação das diferentes condições de vazão e ajuda a explicar por que o modelo manteve resultados mais próximos dos observados desde $Q_{10}$ até $Q_{99}$.

O RF2 também apresentou bom desempenho, principalmente nas vazões altas e intermediárias, mas passou a superestimar com maior frequência as baixas vazões. No Random Forest, a estimativa final resulta da combinação das previsões de várias árvores (Breiman, 2001). Essa característica reduz a variabilidade das previsões, mas pode suavizar respostas mais extremas, o que pode ter contribuído para os desvios encontrados nas maiores permanências.

Entre os modelos treinados pela metodologia sequencial, a LSTM apresentou os melhores resultados, especialmente em $Q_{10}$ e $Q_{50}$. Esse comportamento era esperado devido à sua capacidade de aprender dependências temporais e utilizar informações anteriores da sequência, característica importante para representar a memória hidrológica da bacia (S. R. Clark et al., 2024; Kratzert et al., 2018). Entretanto, seu desempenho diminuiu em $Q_{90}$, $Q_{95}$ e $Q_{99}$, com subestimação das baixas vazões em vários cenários. Como a vazão estimada de um dia é utilizada para estimar o dia seguinte, os erros podem se acumular ao longo da falha, principalmente nos blocos mais longos.

Esse efeito também pode explicar o desempenho inferior de RF1 e XGB1 em relação a RF2 e XGB2. Na metodologia sequencial, esses modelos passam a utilizar valores anteriormente estimados, fazendo com que as incertezas sejam carregadas para os dias seguintes. Os resultados sugerem que esse problema se torna mais importante nas falhas longas e nas baixas vazões. Mesmo submetida a essa condição, a LSTM apresentou maior estabilidade entre os modelos sequenciais, provavelmente por sua estrutura ser mais adequada ao tratamento de sequências temporais.

Os MLPs apresentaram as maiores limitações, resultado semelhante ao observado por (Rahimzad et al., 2021), que identificaram pior desempenho desse modelo, principalmente na

representação dos picos de vazão. O MLP1 tendeu a subestimar $Q_{10}$, enquanto ambos os MLPs superestimaram as baixas vazões ($Q_{90}$, $Q_{95}$ e $Q_{99}$). Esse comportamento pode estar relacionado à ausência de mecanismos internos de memória temporal, tornando o modelo mais dependente das variáveis de entrada e menos eficiente na representação tanto dos extremos quanto da persistência das baixas vazões.

De forma geral, considerando todas as vazões de referência, o XGB2 apresentou o melhor desempenho, seguido pelo RF2, LSTM, XGB1, RF1, MLP2 e MLP1, embora essa ordem tenha variado conforme a faixa de vazão analisada.

**Figura 9:** Razão entre os valores de $Q_{10}$ estimados e observados nos diferentes cenários de ausência de dados

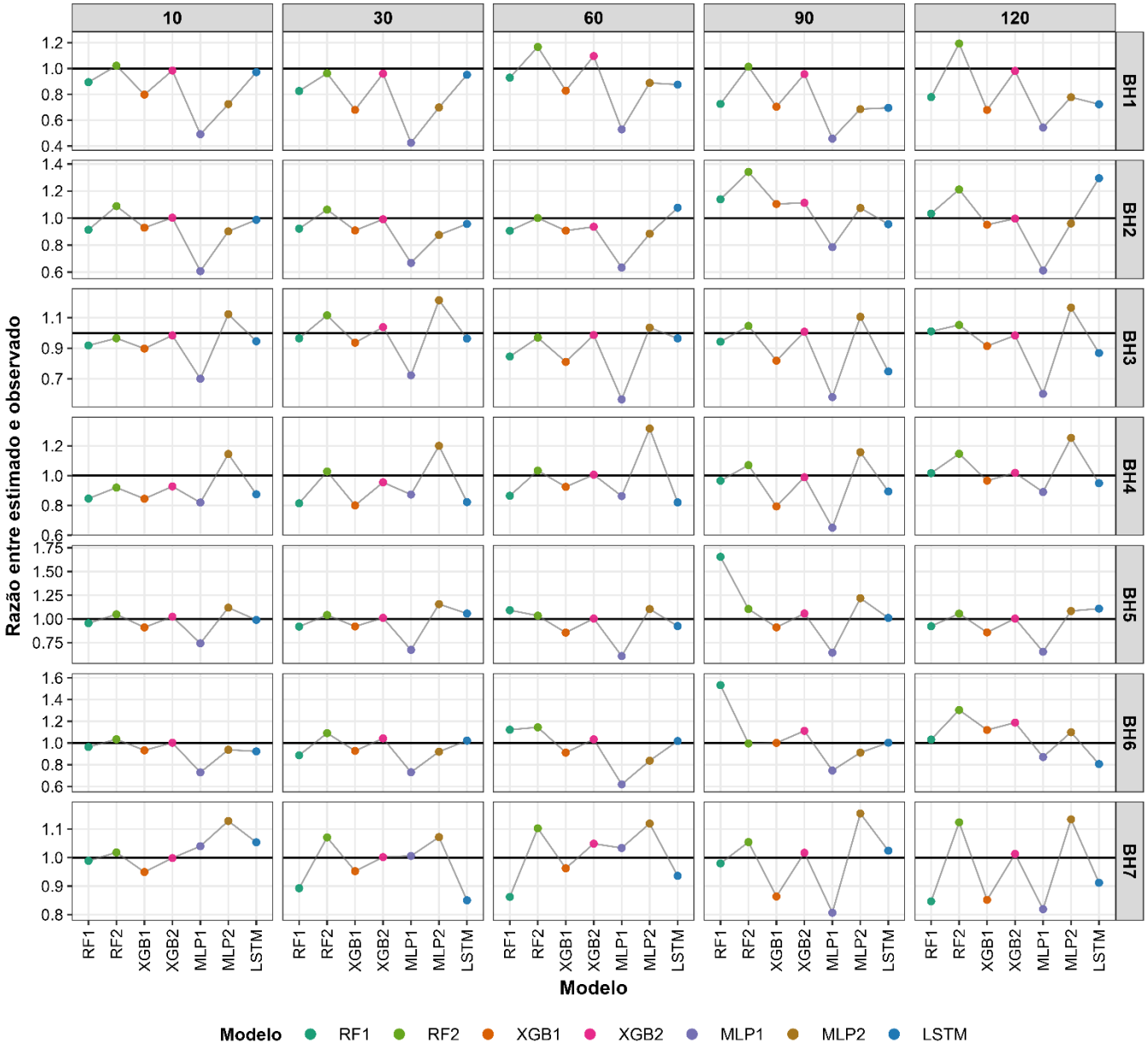


**Figura 10:** Razão entre os valores de $Q_{50}$ estimados e observados nos diferentes cenários de ausência de dados

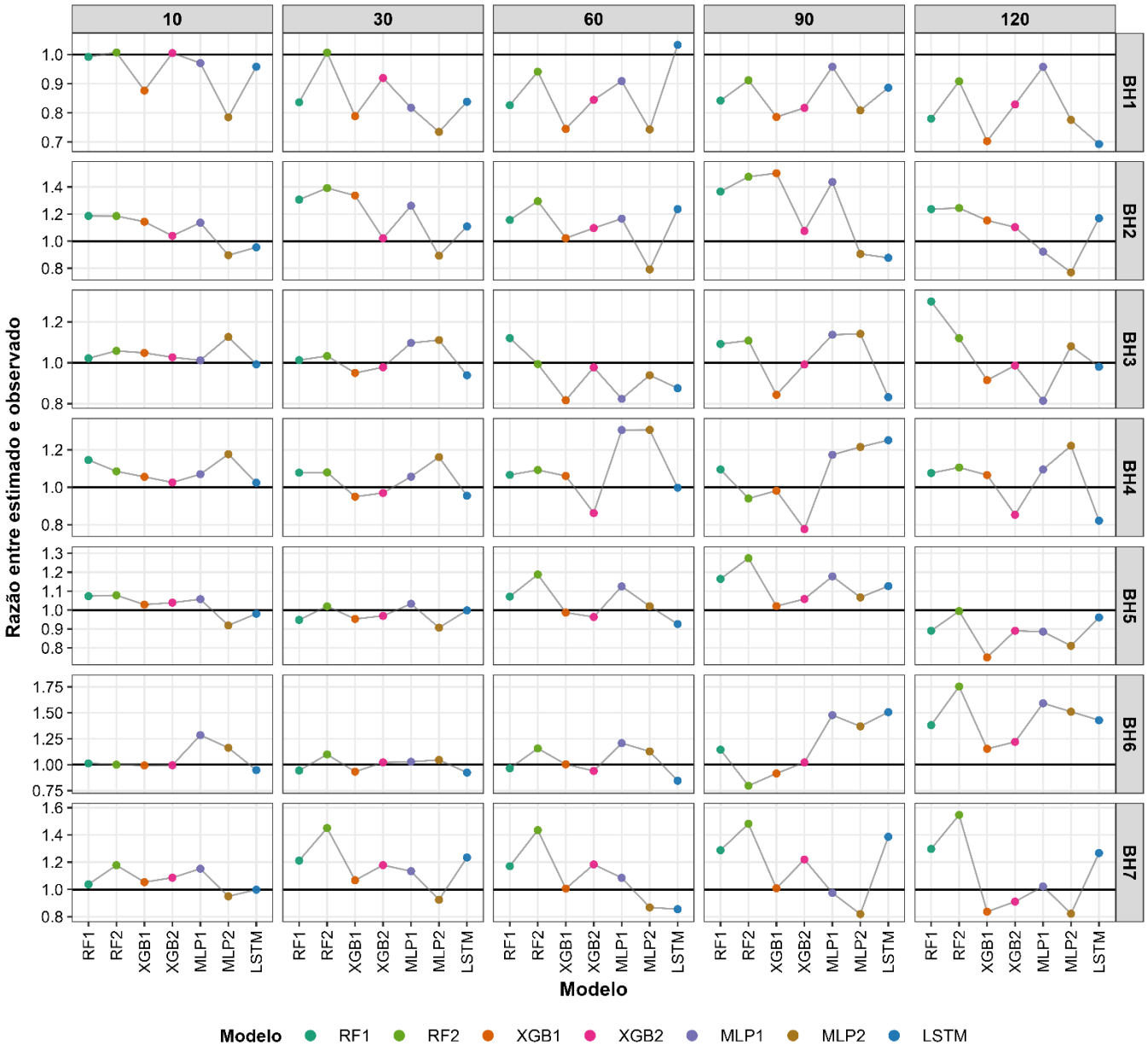


**Figura 11:** Razão entre os valores de $Q_{90}$ estimados e observados nos diferentes cenários de ausência de dados

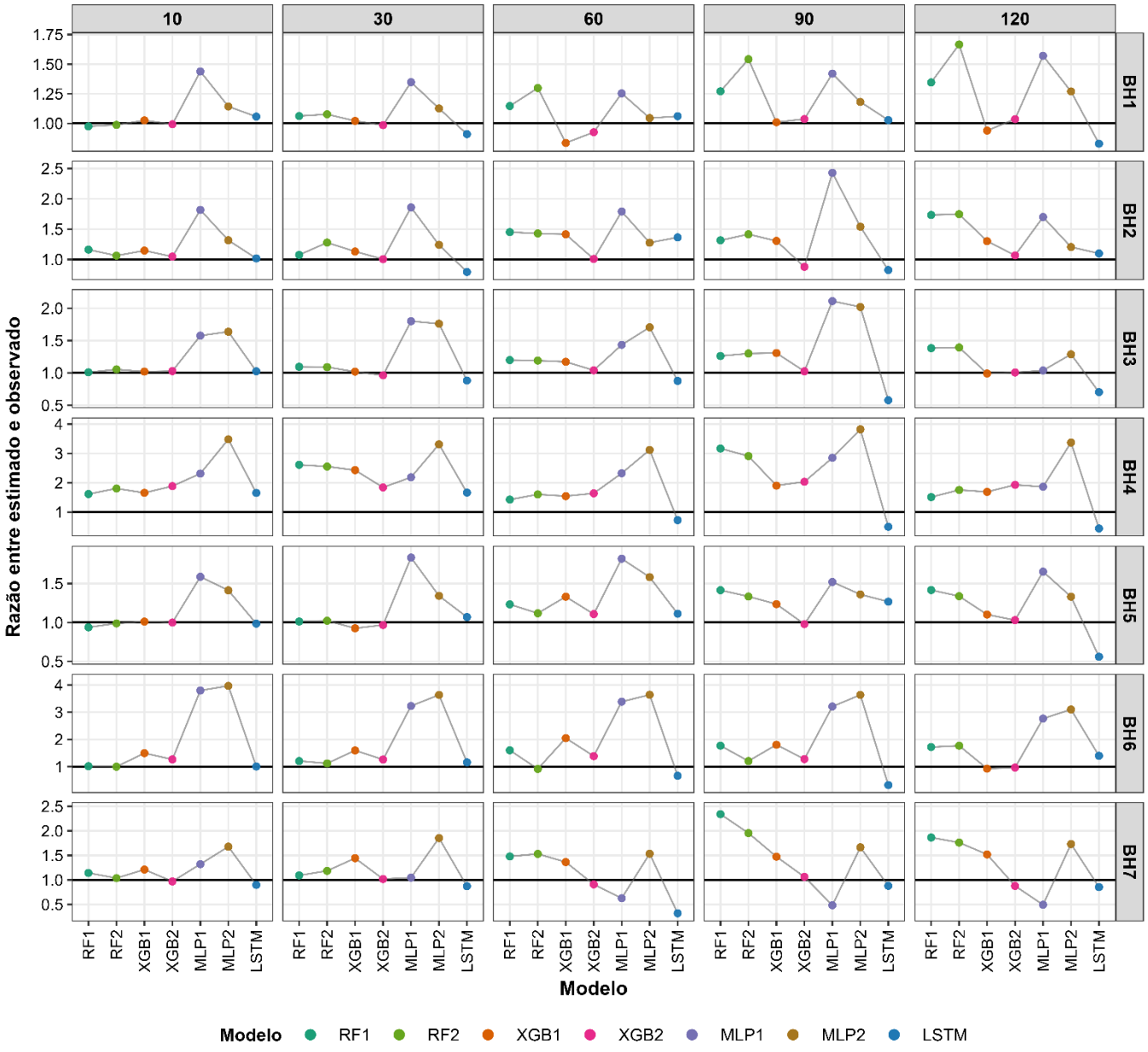


**Figura 12:** Razão entre os valores de $Q_{95}$ estimados e observados nos diferentes cenários de ausência de dados

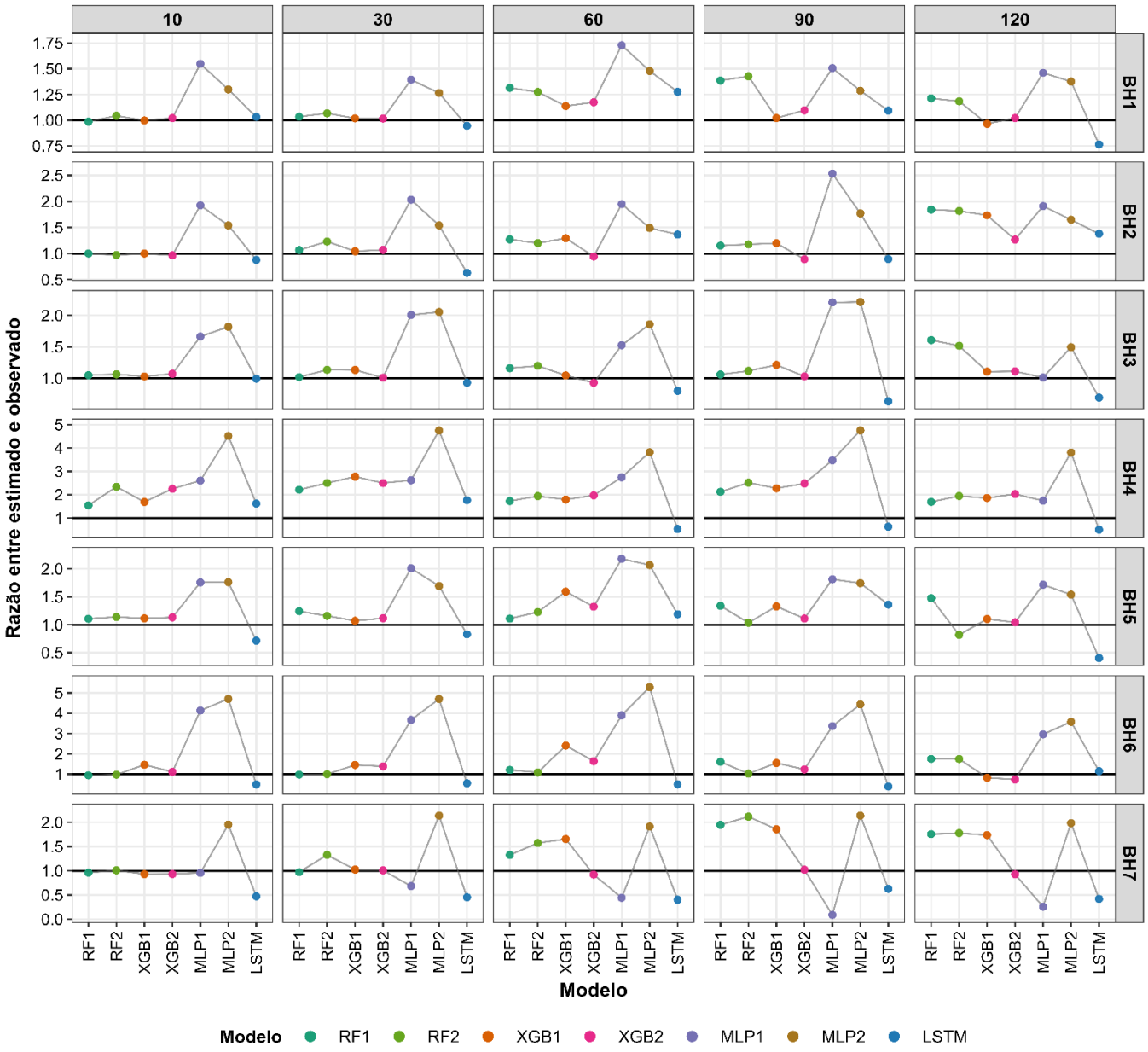
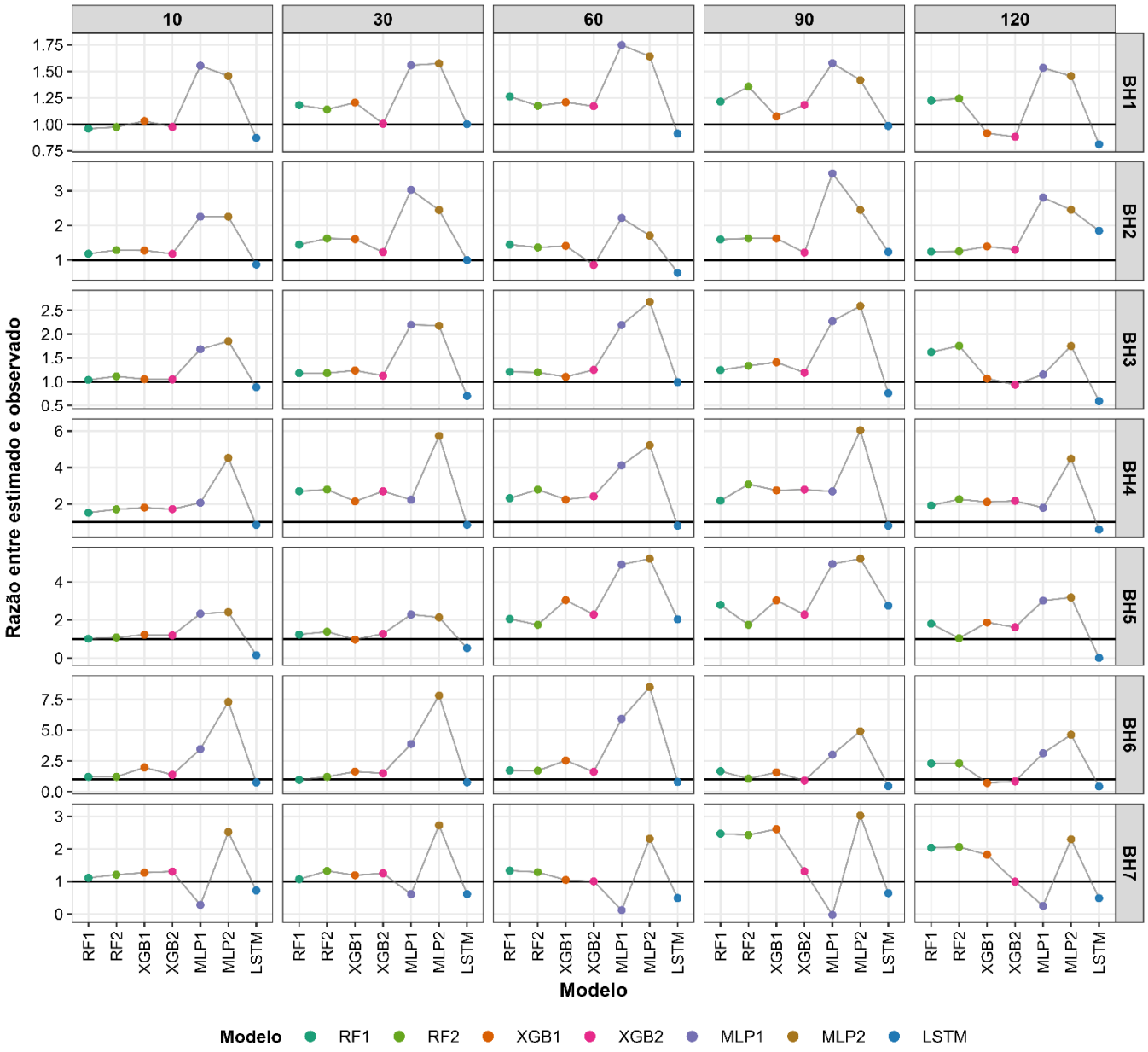


**Figura 13:** Razão entre os valores de $Q_{99}$ estimados e observados nos diferentes cenários de ausência de dados




## 4.4.Regime hidrológico mensal

A Figura 14 apresenta o regime hidrológico médio mensal das sete bacias nos diferentes cenários de falhas consecutivas. Os modelos preservaram a sazonalidade observada, acompanhando os períodos de maior vazão, a recessão durante o período seco e a recuperação no início do período úmido. As maiores diferenças ocorreram nos meses de maior vazão, principalmente na magnitude dos picos, enquanto no período de estiagem as curvas ficaram mais próximas, em parte devido à menor magnitude das vazões.

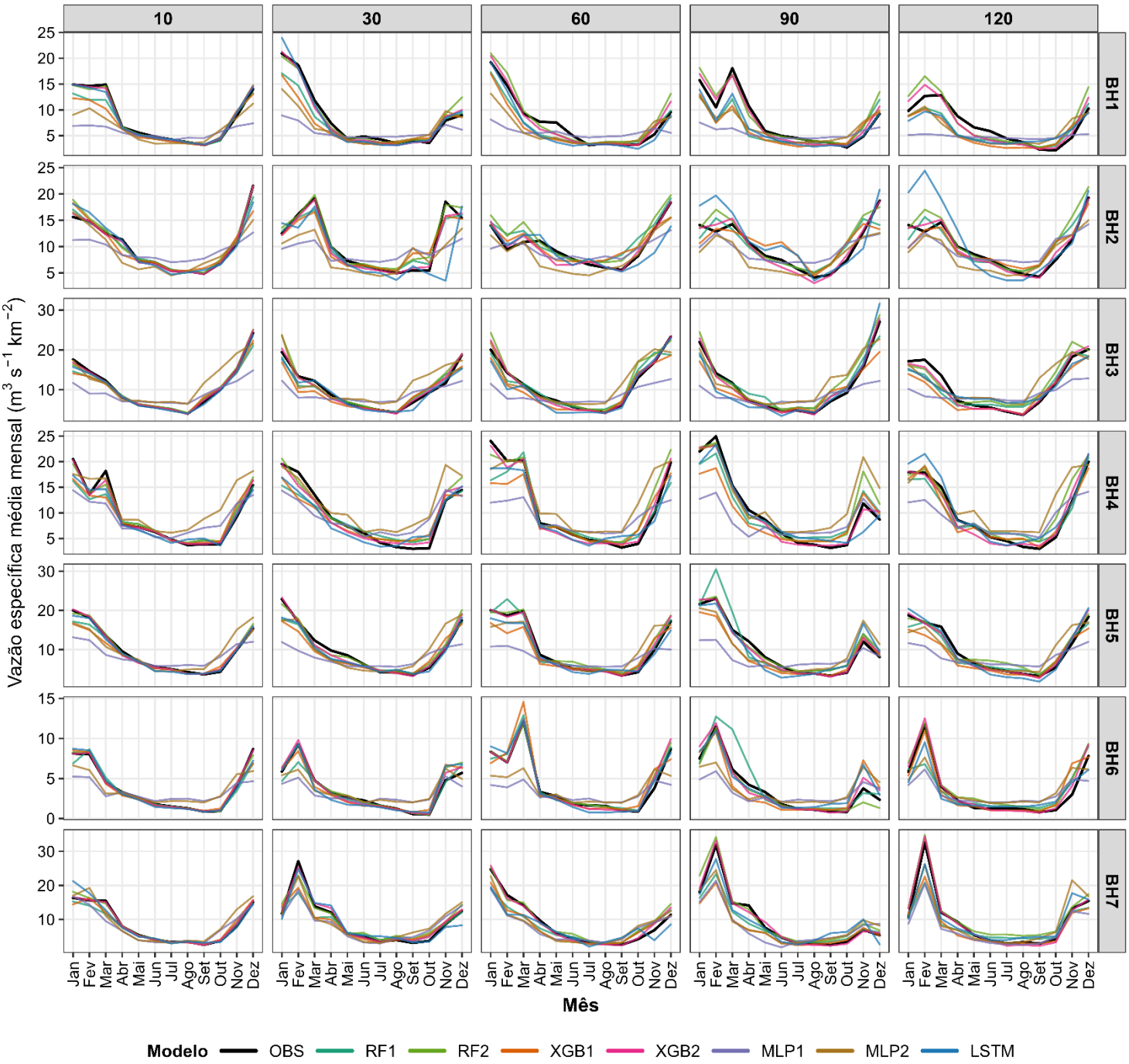
Embora as bacias apresentem diferenças de área de drenagem, forma, posição geográfica e organização da rede de drenagem (Figura 2), essas características não parecem ter alterado o desempenho relativo dos modelos. O padrão observado nas análises anteriores foi mantido na escala mensal: XGB2 apresentou, em geral, maior proximidade com as vazões médias mensais observadas, seguido por RF2 e LSTM. Algumas exceções ocorreram, como RF2 na BH3 com 60 dias e na BH6 com 60 e 90 dias de falha, além da LSTM na BH1 com 10 dias. Essas situações, entretanto, foram pontuais e não indicam uma associação consistente entre determinado modelo e uma bacia específica.

Isso fica mais evidente nas bacias BH6 e BH7. Mesmo localizadas em diferentes regiões e apresentando características físicas e regimes mensais distintos, ambas apresentaram pequena dispersão entre algumas curvas estimadas e observadas, inclusive nos cenários com falhas mais extensas. Da mesma forma, bacias com maior área ou redes de drenagem mais ramificadas não apresentaram uma mudança sistemática na ordem de desempenho dos modelos. Assim, dentro das condições avaliadas, as características das bacias parecem influenciar mais a forma e a magnitude do regime hidrológico do que determinar qual modelo apresenta melhor desempenho.

O efeito da extensão das falhas foi mais evidente. Com 10 dias consecutivos, houve maior sobreposição entre as vazões médias observadas e estimadas. À medida que os blocos aumentaram, os desvios ficaram mais pronunciados, principalmente nos meses de maior vazão. Ainda assim, mesmo com 90 e 120 dias de falha, a estrutura sazonal foi mantida, indicando que os modelos tiveram maior facilidade em reproduzir o período de ocorrência das fases de cheia e estiagem do que a magnitude exata das vazões mensais.

Entre os demais modelos, RF1 e XGB1 mantiveram o comportamento sazonal, mas apresentaram maior variação na magnitude das vazões. Os MLPs mostraram menor concordância, principalmente o MLP1, que produziu curvas mais suavizadas e apresentou maior dificuldade em representar a amplitude sazonal e os picos. Esse comportamento também acompanha os resultados obtidos nas análises anteriores, reforçando que as diferenças de desempenho estão mais associadas à estrutura e à estratégia de treinamento dos modelos do que a uma condição específica de determinada bacia.

**Figura 14:** Vazões médias mensais observadas e estimadas por cada modelo de machine learning nos diferentes cenários de ausência de dados (10 a 120 dias)




4.5.Comparação entre os modelos de Machine Learning

A Figura 15 apresenta a distribuição dos resíduos e as diferenças entre os modelos de Machine Learning usados para a geração das séries de dados de vazão diária, avaliadas pelo teste de Friedman seguido do teste de Wilcoxon a 5% de significância. Na maioria das bacias e cenários

de falha, os modelos diferiram significativamente, mesmo quando apresentaram medianas próximas de zero. Isso ocorre porque o teste de Wilcoxon considera a direção e a magnitude das diferenças entre observações pareadas, e não apenas a proximidade entre seus valores centrais (Wilcoxon et al., 1963). Assim, diferenças estatísticas podem ocorrer mesmo quando a diferença absoluta entre os modelos é pequena, como observado por (Gauch et al., 2021).

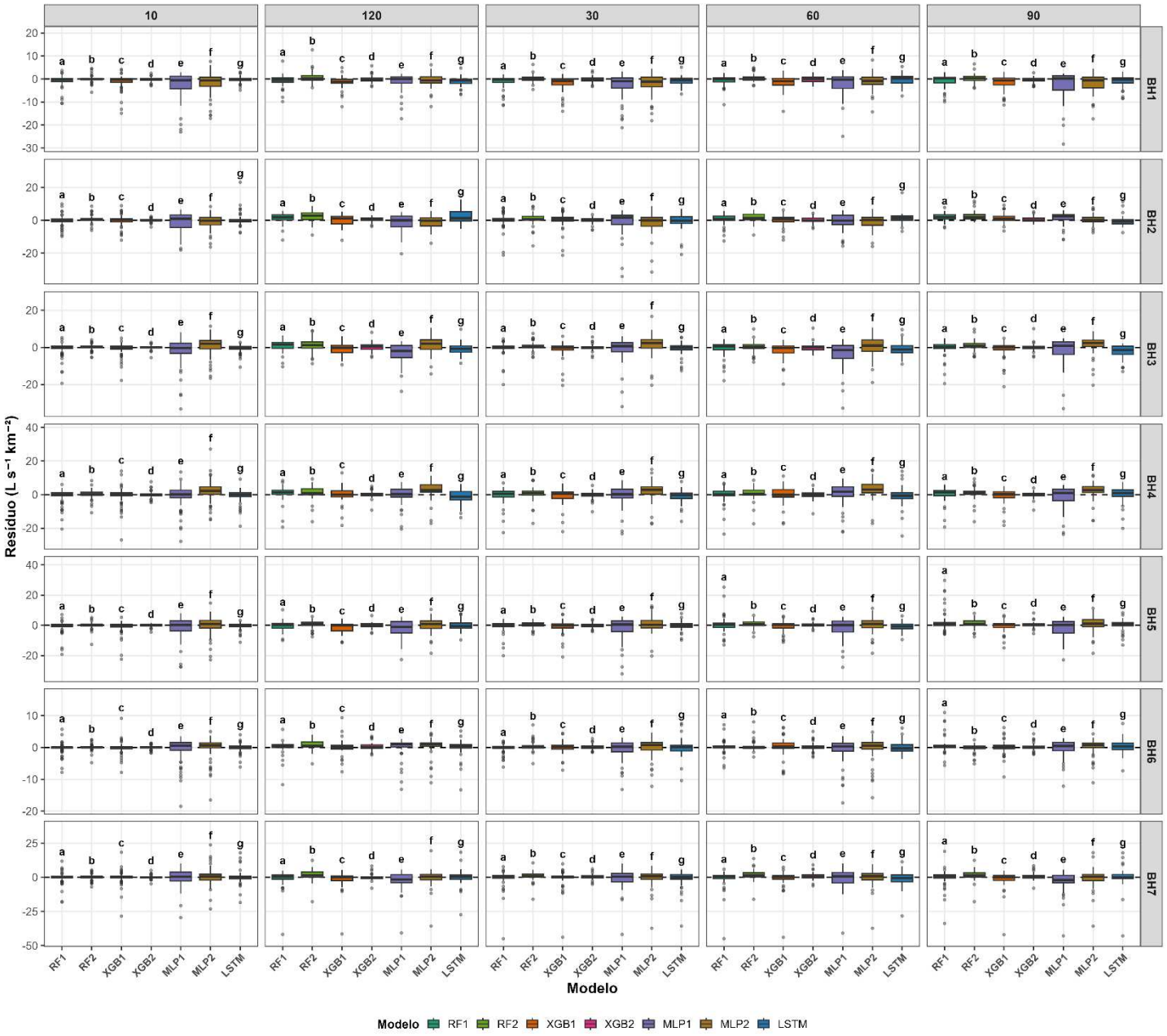
Como já visto, o XGB2 apresentou resíduos próximos de zero e baixa dispersão, mantendo o bom desempenho observado nas análises anteriores. O RF2 também apresentou resíduos relativamente concentrados, embora com tendência de superestimação em algumas situações. Entre os modelos sequenciais, a LSTM apresentou comportamento mais equilibrado, enquanto RF1 e XGB1 alternaram entre superestimações e subestimações dependendo da bacia e da extensão da falha.

As maiores variações foram observadas nos MLPs. O MLP1 apresentou maior dispersão e resíduos negativos mais extremos em várias bacias, indicando maior ocorrência de subestimações. O MLP2 também apresentou maior variabilidade, mas com resíduos positivos mais frequentes em algumas condições. Isso mostra uma maior sensibilidade desses modelos às variações das séries e aos períodos em que as falhas foram inseridas.

A dispersão dos resíduos também variou entre as bacias, mas não aumentou continuamente com o número de dias de falha. Em algumas situações, blocos menores apresentaram resíduos mais extremos que blocos maiores. Portanto, o erro não depende apenas da extensão da falha, mas também do período hidrológico em que ela ocorre.

De forma geral, a análise dos resíduos confirma o padrão encontrado nas avaliações anteriores: XGB2 e RF2 apresentaram maior estabilidade, enquanto a LSTM se destacou entre os modelos sequenciais. Mesmo com diferenças entre bacias e períodos de falha, o comportamento relativo dos modelos permaneceu semelhante ao observado nas demais análises.

**Figura 15:** Comparação entre os resíduos das vazões observadas e estimadas por cada modelo de machine leanrnig nos diferentes cenários de ausência de dados, pelo teste de Friedman e Wilcoxon




4.6. Hidrogramas das vazões simuladas para os modelos XGB2, RF2 e LSTM

As Figuras 16 a 22 apresentam, para as sete bacias hidrográficas avaliadas, os hidrogramas dos três modelos com melhor desempenho nas análises anteriores: XGB2, RF2 e LSTM. Os três acompanharam o comportamento geral das vazões, mas apresentaram diferenças principalmente

na representação dos picos. O XGB2 permaneceu mais próximo da série observada na maioria das bacias e cenários de falha, seguido pelo RF2 e pela LSTM.

O melhor desempenho do XGB2 também foi observado por (Ding et al., 2025; Hameed et al., 2025; Terzi et al., 2023), que encontraram resultados superiores do XGBoost em relação ao RF e à LSTM na estimativa diária de vazões. Essa vantagem está relacionada à capacidade do algoritmo de representar relações não lineares e corrigir progressivamente os erros durante o treinamento. Entretanto, esse comportamento não é geral.

O RF2 acompanhou bem a série observada, mas apresentou maior dificuldade em alguns picos. Comportamento semelhante foi encontrado por Schoppa et al. (2020), que observaram bom desempenho do Random Forest para vazões de baixa e média magnitude, com redução da precisão nos eventos de maior vazão. A combinação das previsões de várias árvores favorece a estabilidade do RF, mas pode suavizar valores extremos.

A LSTM também reproduziu bem a dinâmica das vazões, embora tenha apresentado maior variação nos picos. Resultados satisfatórios para esse modelo também foram encontrados em outros estudos (Clark et al., 2024; Lees et al., 2021; Rahimzad et al., 2021). Neste estudo, seu desempenho pode ter sido limitado pelo treinamento sequencial com menor quantidade de dados e pela propagação dos erros ao longo das falhas.

Além disso, mesmo quando apresenta bom desempenho geral, a LSTM pode ter dificuldade em reproduzir a magnitude dos eventos extremos. Clark et al. (2024) observaram subestimação de picos em parte das simulações, mostrando que um bom ajuste global não garante a representação exata das maiores vazões.

O aumento das falhas não alterou de forma clara a ordem dos modelos. Mesmo nos blocos de 90 e 120 dias, XGB2 e RF2 permaneceram mais próximos das vazões observadas, enquanto a LSTM continuou sendo o melhor modelo sequencial. Assim, os hidrogramas confirmam os resultados anteriores e mostram que a principal diferença entre os três modelos está na representação dos picos. O melhor desempenho de XGB2 e RF2 pode estar relacionado tanto às características desses algoritmos quanto à maior quantidade de dados disponível, além de não possuírem informações que carregam erros decorrente de ajustes anteriores.

**Figura 16:** Hidrogramas das vazões observadas e simuladas pelos modelos XGB2, RF2 e LSTM, em cada cenário de ausência de dados, para a bacia BH1

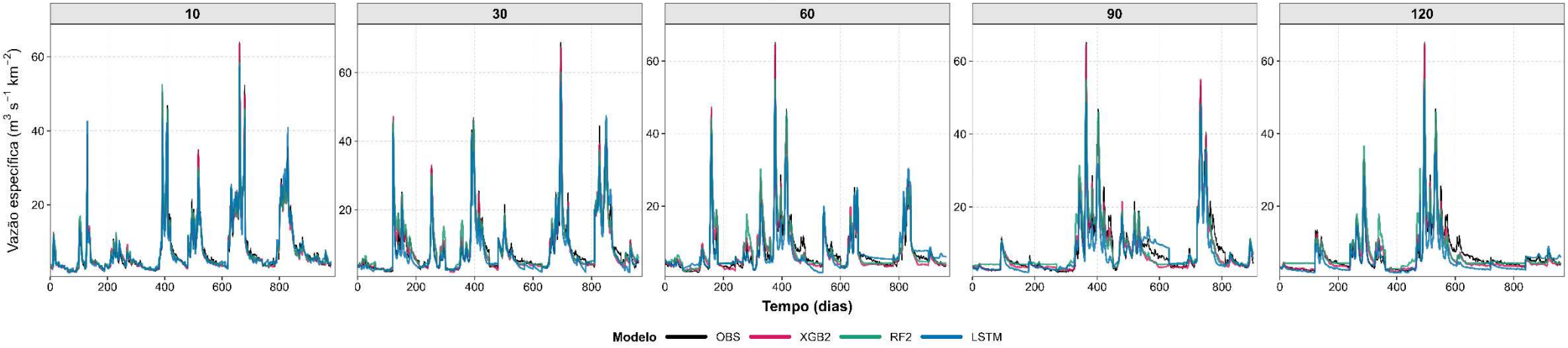



**Figura 17:** Hidrogramas das vazões observadas e simuladas pelos modelos XGB2, RF2 e LSTM, em cada cenário de ausência de dados, para a bacia BH2

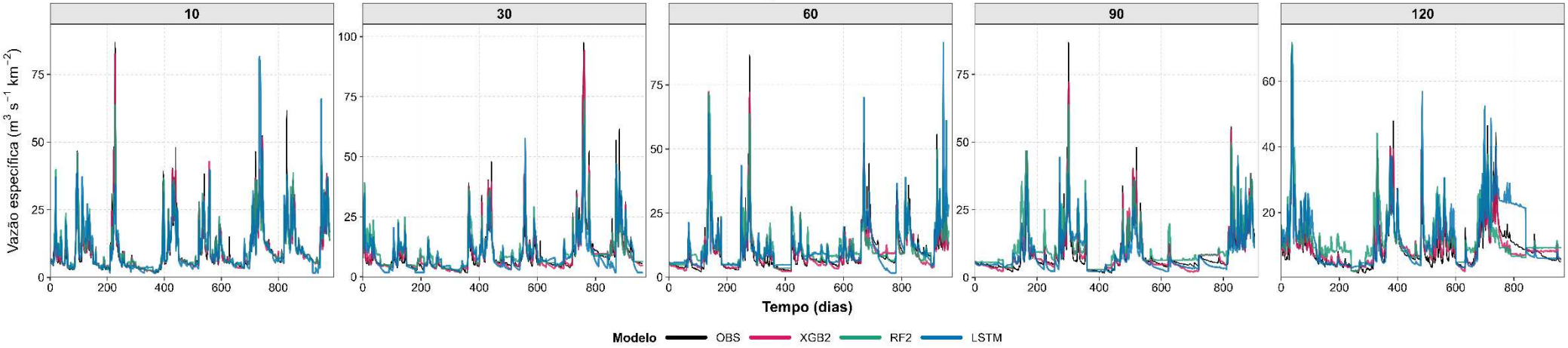


**Figura 18:** Hidrogramas das vazões observadas e simuladas pelos modelos XGB2, RF2 e LSTM, em cada cenário de ausência de dados, para a bacia BH3

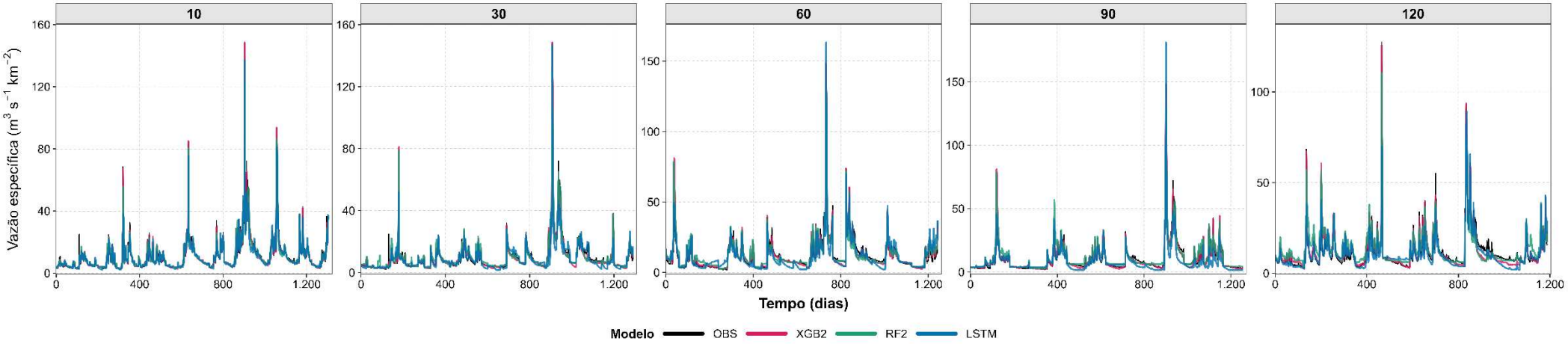



**Figura 19:** Hidrogramas das vazões observadas e simuladas pelos modelos XGB2, RF2 e LSTM, em cada cenário de ausência de dados, para a bacia BH4

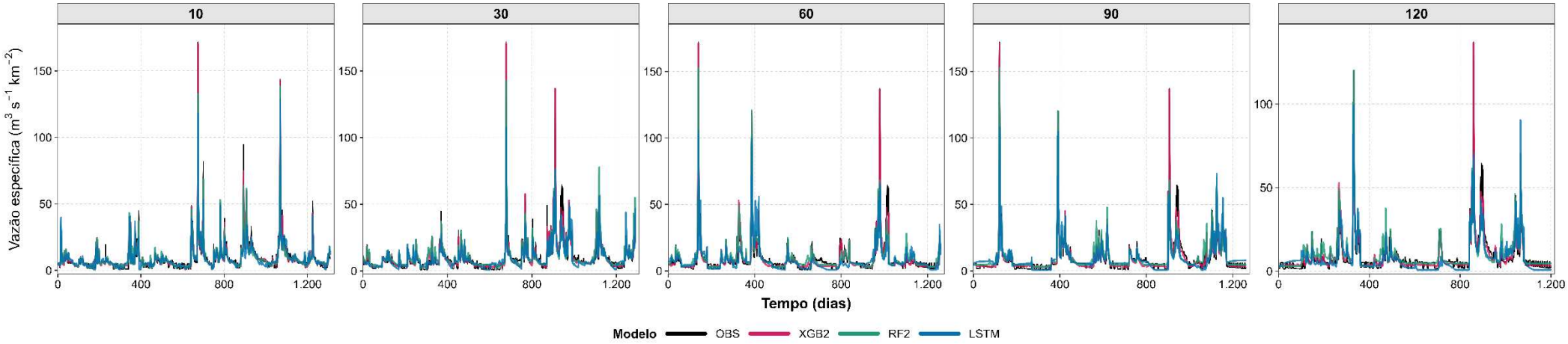


**Figura 20:** Hidrogramas das vazões observadas e simuladas pelos modelos XGB2, RF2 e LSTM, em cada cenário de ausência de dados, para a bacia BH5

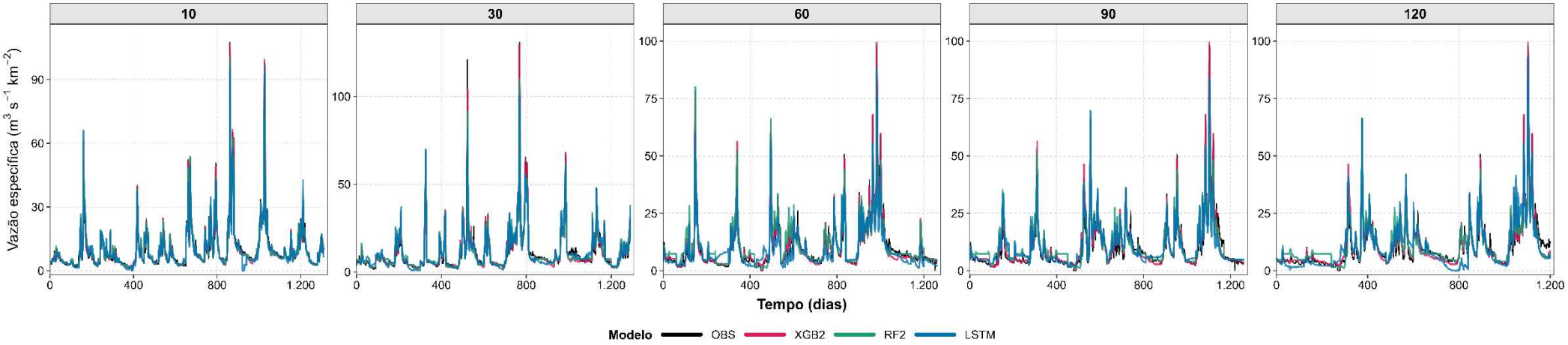



**Figura 21:** Hidrogramas das vazões observadas e simuladas pelos modelos XGB2, RF2 e LSTM, em cada cenário de ausência de dados, para a bacia BH6

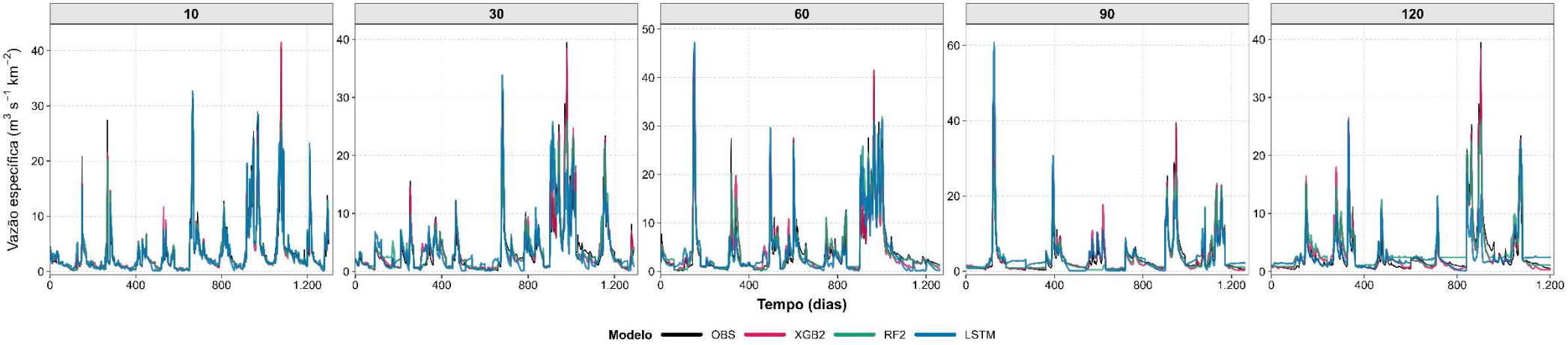
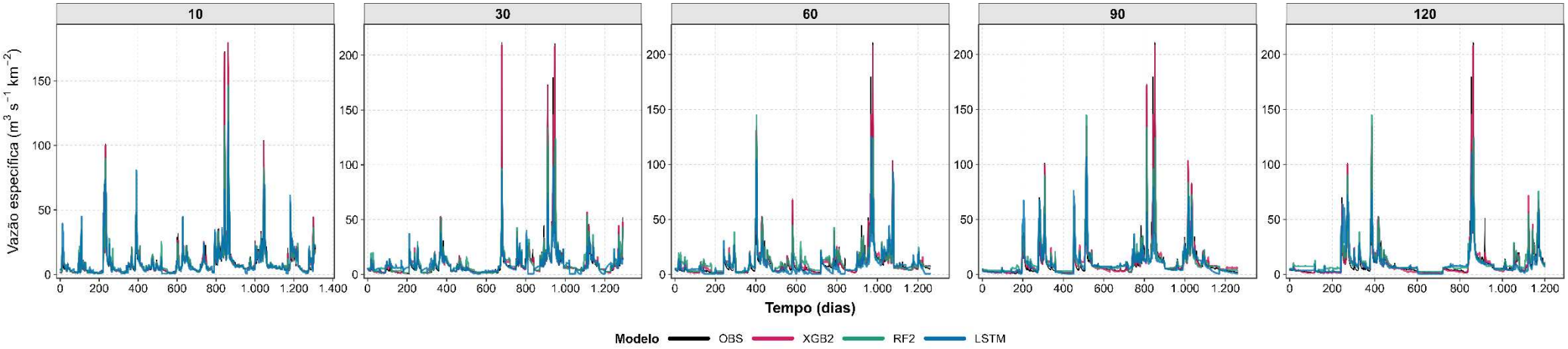


**Figura 22:** Hidrogramas das vazões observadas e simuladas pelos modelos XGB2, RF2 e LSTM, em cada cenário de ausência de dados, para a bacia BH7



# 5. CONCLUSÕES

Com base nos resultados obtidos, pode-se concluir que:

- Os modelos RF, XGB, MLP e LSTM apresentaram melhor capacidade de reconstrução das vazões nos blocos de 10 dias consecutivos de falhas, com redução do desempenho à medida que as falhas aumentaram. Entretanto, a partir de 60 dias, as diferenças entre os períodos de falha tornaram-se menos evidentes;

- A estratégia de treinamento influenciou o desempenho dos modelos, sendo observados melhores resultados para os modelos baseados em árvores quando treinados pela metodologia tradicional;

- Nas condições avaliadas neste estudo, as diferenças fisiográficas entre as sub-bacias não resultaram em alteração sistemática do padrão relativo de desempenho dos modelos

- O XGB2 treinado pela metodologia tradicional apresentou o melhor desempenho geral, seguido pelo RF2, mantendo maior precisão na reconstrução das séries e melhor representação das diferentes faixas de vazão;

- A LSTM apresentou o melhor desempenho entre os modelos avaliados pela metodologia sequencial, demonstrando capacidade de reproduzir a dinâmica temporal das séries, embora com desempenho geral inferior ao XGB2 e ao RF2 treinados pela metodologia tradicional;

- Os modelos RF1 e XGB1 apresentaram desempenho inferior às respectivas versões treinadas pela metodologia tradicional, enquanto os modelos MLP apresentaram, de modo geral, os resultados menos satisfatórios entre os algoritmos avaliados;

- Os modelos apresentaram maior dificuldade na representação das condições extremas do regime hidrológico, principalmente na reprodução da magnitude dos picos de vazão e das condições mais críticas de baixa vazão.

- A importância das covariáveis variou entre os modelos, com maior contribuição da vazão do dia anterior e das precipitações ocorridas dois e três dias antes.

# REFERÊNCIAS


Agência Nacional de Águas e Saneamento Básico. (January). *Hidroweb: Sistema de Informações Hidrológicas*. Sistema Nacional de Informações sobre Recursos Hídricos (SNIRH); Agência Nacional de Águas e Saneamento Básico (ANA). https://www.snirh.gov.br/hidroweb/apresentacao

Ali Abd Al-Hameed, K. (2022). Spearman's correlation coefficient in statistical analysis. *International Journal of Nonlinear Analysis and Applications*, *13*(1), 3249–3255.

Ali, S., Ghosh, N. C., & Singh, R. (2010). Rainfall–runoff simulation using a normalized antecedent precipitation index. *Hydrological Sciences Journal*, *55*(2), 266–274. https://doi.org/10.1080/02626660903546175

Almeida, C. A., Oliveira, A. F. de, Pacheco, A. A., Lopes, R. P., Neves, A. A., & Lopes Ribeiro de Queiroz, M. E. (2018). Characterization and evaluation of sorption potential of the iron mine waste after Samarco dam disaster in Doce River basin – Brazil. *Chemosphere*, *209*, 411–420. https://doi.org/10.1016/j.chemosphere.2018.06.071

Althoff, D., & Rodrigues, L. N. (2021). Goodness-of-fit criteria for hydrological models: Model calibration and performance assessment. *Journal of Hydrology*, *600*, 126674. https://doi.org/10.1016/j.jhydrol.2021.126674

Andrade, C. (2021). Z Scores, Standard Scores, and Composite Test Scores Explained. *Indian Journal of Psychological Medicine*, *43*(6), 555–557. https://doi.org/10.1177/02537176211046525

Armstrong, J. S., & Forecasting, L.-R. (1985). From crystal ball to computer. *New York Ua*, *348*, 553.

Arriagada, P., Karelovic, B., & Link, O. (2021). Automatic gap-filling of daily streamflow time series in data-scarce regions using a machine learning algorithm. *Journal of Hydrology*, *598*, 126454. https://doi.org/10.1016/j.jhydrol.2021.126454

Ashfaq, S., Waseem, S., Tufail, M., & Tariq, A. (2026). Evaluating the Accuracy of Machine Learning Models with Short-Term Lag Inputs for Univariate Daily Streamflow time Series Forecasting. *Water Resources Management*, *40*(4), 177. https://doi.org/10.1007/s11269-025-04409-x

Bennett, B., Leonard, M., Deng, Y., & Westra, S. (2018). An empirical investigation into the effect of antecedent precipitation on flood volume. *Journal of Hydrology*, *567*, 435–445. https://doi.org/10.1016/j.jhydrol.2018.10.025

Bentéjac, C., Csörgő, A., & Martínez-Muñoz, G. (2021). A comparative analysis of gradient

boosting algorithms. *Artificial Intelligence Review*, *54*(3), 1937–1967. https://doi.org/10.1007/s10462-020-09896-5

Berggren, M. (2024). Coefficients of determination measured on the same scale as the outcome: Alternatives to $R^2$ that use standard deviations instead of explained variance. *Psychological Methods*. https://doi.org/10.1037/met0000681

Biau, G., & Scornet, E. (2016). A random forest guided tour. *TEST*, *25*(2), 197–227. https://doi.org/10.1007/s11749-016-0481-7

Birgand, F., Lellouche, G., & Appelboom, T. W. (2013). Measuring flow in non-ideal conditions for short-term projects: Uncertainties associated with the use of stage-discharge rating curves. *Journal of Hydrology*, *503*, 186–195. https://doi.org/10.1016/j.jhydrol.2013.09.007

Bombardi, R. J., & Carvalho, L. M. V. (2011). The South Atlantic dipole and variations in the characteristics of the South American Monsoon in the WCRP-CMIP3 multi-model simulations. *Climate Dynamics*, *36*(11–12), 2091–2102. https://doi.org/10.1007/s00382-010-0836-9

Breiman, L. (2001). Random Forests. *Machine Learning*, *45*(1), 5–32. https://doi.org/10.1023/A:1010933404324

Brunner, P., & Simmons, C. T. (2012). HydroGeoSphere: a fully integrated, physically based hydrological model. *Ground Water, Wiley, 2012/50/2/170-176*.

Cabral Júnior, J. B., & Lucena, R. L. (2020). ANALYSIS OF PRECIPITATIONS BY NON-PARAMETRIC TESTS OF MANN-KENDALL AND KRUSKAL-WALLIS. *Mercator*, *19*(1), 1–14. https://doi.org/10.4215/rm2020.e19001

Carvalho, D. F. de, Rocha, H. S. da, Bonomo, R., & Souza, A. P. de. (2015). Estimativa da evapotranspiração de referência a partir de dados meteorológicos limitados. *Pesquisa Agropecuária Brasileira*, *50*(1), 1–11. https://doi.org/10.1590/S0100-204X2015000100001

Cavalcanti, I. F. A. (2016). *Tempo e clima no Brasil*. Oficina de textos.

CBH-Doce, C. da B. H. do R. D. (2026). *A Bacia*. https://cbhdoce.org.br/institucional/a-bacia

Chagas, V. B. P., Chaffe, P. L. B., Addor, N., Fan, F. M., Fleischmann, A. S., Paiva, R. C. D., & Siqueira, V. A. (2020). CAMELS-BR: hydrometeorological time series and landscape attributes for 897 catchments in Brazil. *Earth System Science Data*, *12*(3), 2075–2096. https://doi.org/10.5194/essd-12-2075-2020

Chen, T., & Guestrin, C. (2016). XGBoost. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 785–794.

https://doi.org/10.1145/2939672.2939785

Chicco, D., Warrens, M. J., & Jurman, G. (2021). The coefficient of determination R-squared is more informative than SMAPE, MAE, MAPE, MSE and RMSE in regression analysis evaluation. *PeerJ Computer Science*, *7*, e623. https://doi.org/10.7717/peerj-cs.623

Cho, K., & Kim, Y. (2022). Improving streamflow prediction in the WRF-Hydro model with LSTM networks. *Journal of Hydrology*, *605*, 127297. https://doi.org/10.1016/j.jhydrol.2021.127297

Clark, M. P., Bierkens, M. F. P., Samaniego, L., Woods, R. A., Uijlenhoet, R., Bennett, K. E., Pauwels, V. R. N., Cai, X., Wood, A. W., & Peters-Lidard, C. D. (2017). The evolution of process-based hydrologic models: historical challenges and the collective quest for physical realism. *Hydrology and Earth System Sciences*, *21*(7), 3427–3440. https://doi.org/10.5194/hess-21-3427-2017

Clark, S. R., Lerat, J., Perraud, J.-M., & Fitch, P. (2024). Deep learning for monthly rainfall–runoff modelling: a large-sample comparison with conceptual models across Australia. *Hydrology and Earth System Sciences*, *28*(5), 1191–1213. https://doi.org/10.5194/hess-28-1191-2024

Collins, E. L., David, C. H., Riggs, R., Allen, G. H., Pavelsky, T. M., Lin, P., Pan, M., Yamazaki, D., Meentemeyer, R. K., & Sanchez, G. M. (2024). Global patterns in river water storage dependent on residence time. *Nature Geoscience*, *17*(5), 433–439. https://doi.org/10.1038/s41561-024-01421-5

Cupolillo, F., Abreu, M. L. de, & Vianello, R. L. (2022). Climatologia da Bacia do Rio Doce e sua relação com a topografia local. *Revista Geografias*, *4*(2), 45–60. https://doi.org/10.35699/2237-549X..13251

D. N. Moriasi, J. G. Arnold, M. W. Van Liew, R. L. Bingner, R. D. Harmel, & T. L. Veith. (2007). Model Evaluation Guidelines for Systematic Quantification of Accuracy in Watershed Simulations. *Transactions of the ASABE*, *50*(3), 885–900. https://doi.org/10.13031/2013.23153

Dakhlaoui, H., Ruelland, D., Tramblay, Y., & Bargaoui, Z. (2017). Evaluating the robustness of conceptual rainfall-runoff models under climate variability in northern Tunisia. *Journal of Hydrology*, *550*, 201–217. https://doi.org/10.1016/j.jhydrol.2017.04.032

Darbandi, S., & Pourhosseini, F. A. (2018). River flow simulation using a multilayer perceptron-firefly algorithm model. *Applied Water Science*, *8*(3), 85. https://doi.org/10.1007/s13201-018-0713-y

Dastour, H., & Hassan, Q. K. (2023). A Machine-Learning Framework for Modeling and

Predicting Monthly Streamflow Time Series. *Hydrology*, *10*(4), 95. https://doi.org/10.3390/hydrology10040095

Depetris, P. J. (2021). The Importance of Monitoring River Water Discharge. *Frontiers in Water*, *3*. https://doi.org/10.3389/frwa.2021.745912

Dias, T. L., Cataldi, M., & Ferreira, V. H. (2016). Aplicação de técnicas de redes neurais e modelagem atmosférica para elaboração de previsões de vazão na Bacia do Rio Grande (MG). *Engenharia Sanitaria e Ambiental*, *22*(1), 169–178. https://doi.org/10.1590/s1413-41522016158015

Ding, B., Yu, X., & Jia, G. (2025). Exploring the Controlling Factors of Watershed Streamflow Variability Using Hydrological and Machine Learning Models. *Water Resources Research*, *61*(5). https://doi.org/10.1029/2024WR039734

Dobriyal, P., Badola, R., Tuboi, C., & Hussain, S. A. (2017). A review of methods for monitoring streamflow for sustainable water resource management. *Applied Water Science*, *7*(6), 2617–2628. https://doi.org/10.1007/s13201-016-0488-y

Ďurigová, M., Hlavčová, K., & Poórová, J. (2020). Detection of Changes in Hydrological Time Series During Recent Decades. *Slovak Journal of Civil Engineering*, *28*(2), 56–62. https://doi.org/10.2478/sjce-2020-0016

Ejiohuo, O., Onyeaka, H., Akinsemolu, A., Nwabor, O. F., Siyanbola, K. F., Tamasiga, P., & Al-Sharify, Z. T. (2025). Ensuring water purity: Mitigating environmental risks and safeguarding human health. *Water Biology and Security*, *4*(2), 100341. https://doi.org/10.1016/j.watbs.2024.100341

Erikson, C. M., Renshaw, C. E., & Magilligan, F. J. (2024). Spatial variation in drainage area — Runoff relationships and implications for bankfull geometry scaling. *Geomorphology*, *446*, 108998. https://doi.org/10.1016/j.geomorph.2023.108998

Eugénia Graça Martins, M. (2018). Coeficiente de determinação. *Revista de Ciência Elementar*, *6*(1). https://doi.org/10.24927/rce2018.024

Evans, C. M., Dritschel, D. G., & Singer, M. B. (2018). Modeling Subsurface Hydrology in Floodplains. *Water Resources Research*, *54*(3), 1428–1459. https://doi.org/10.1002/2017WR020827

Falkenmark, M., & Lannerstad, M. (2005). Consumptive water use to feed humanity - curing a blind spot. *Hydrology and Earth System Sciences*, *9*(1/2), 15–28. https://doi.org/10.5194/hess-9-15-2005

Fasipe, O. A., & Izinyon, O. C. (2021). Exponent determination in a poorly gauged basin system in Nigeria based on flow characteristics investigation and regionalization method.

*SN Applied Sciences*, *3*(3), 319. https://doi.org/10.1007/s42452-021-04302-3

Fathi, M. M., Al Mehedi, M. A., Smith, V., Fernandes, A. M., Hren, M. T., & Terry, D. O. (2025). Evaluation of LSTM vs. conceptual models for hourly rainfall runoff simulations with varied training period lengths. *Scientific Reports*, *15*(1), 15820. https://doi.org/10.1038/s41598-025-96577-4

Fischer, S., Dallan, E., Fiori, A., Grimaldi, S., Kochanek, K., Prieto, C., Reis, D. S., & Volpi, E. (2025). Hydrological design in the HELPING decade – inspiring the community to innovate the hydrological design concept. *Hydrological Sciences Journal*, *70*(3), 375–389. https://doi.org/10.1080/02626667.2024.2436634

Fortesa, J., García-Comendador, J., Calsamiglia, A., López-Tarazón, J. A., Latron, J., Alorda, B., & Estrany, J. (2019). Comparison of stage/discharge rating curves derived from different recording systems: Consequences for streamflow data and water management in a Mediterranean island. *Science of The Total Environment*, *665*, 968–981. https://doi.org/10.1016/j.scitotenv.2019.02.158

Fowler, K., Coxon, G., Freer, J., Peel, M., Wagener, T., Western, A., Woods, R., & Zhang, L. (2018). Simulating Runoff Under Changing Climatic Conditions: A Framework for Model Improvement. *Water Resources Research*, *54*(12), 9812–9832. https://doi.org/10.1029/2018WR023989

Friedman, M. (1937). The Use of Ranks to Avoid the Assumption of Normality Implicit in the Analysis of Variance. *Journal of the American Statistical Association*, *32*(200), 675–701. https://doi.org/10.1080/01621459.1937.10503522

Gacu, J. G., Monjardin, C. E. F., Mangulabnan, R. G. T., & Mendez, J. C. F. (2025). Application of Artificial Intelligence in Hydrological Modeling for Streamflow Prediction in Ungauged Watersheds: A Review. *Water*, *17*(18), 2722. https://doi.org/10.3390/w17182722

Gauch, M., Kratzert, F., Klotz, D., Nearing, G., Lin, J., & Hochreiter, S. (2021). Rainfall–runoff prediction at multiple timescales with a single Long Short-Term Memory network. *Hydrology and Earth System Sciences*, *25*(4), 2045–2062. https://doi.org/10.5194/hess-25-2045-2021

Gholami, V., & Sahour, H. (2022). Simulation of rainfall-runoff process using an artificial neural network (ANN) and field plots data. *Theoretical and Applied Climatology*, *147*(1–2), 87–98. https://doi.org/10.1007/s00704-021-03817-4

Ghotbi, S., Wang, D., Singh, A., Blöschl, G., & Sivapalan, M. (2020). A New Framework for Exploring Process Controls of Flow Duration Curves. *Water Resources Research*, *56*(1).

https://doi.org/10.1029/2019WR026083

Gillard, J., O'Riordan, E., & Zhigljavsky, A. (2022). Simplicial and Minimal-Variance Distances in Multivariate Data Analysis. *Journal of Statistical Theory and Practice*, *16*(1), 9. https://doi.org/10.1007/s42519-021-00227-7

Girihagama, L., Naveed Khaliq, M., Lamontagne, P., Perdikaris, J., Roy, R., Sushama, L., & Elshorbagy, A. (2022). Streamflow modelling and forecasting for Canadian watersheds using LSTM networks with attention mechanism. *Neural Computing and Applications*, *34*(22), 19995–20015. https://doi.org/10.1007/s00521-022-07523-8

Gravelle, R. (2015). 3.3. 5. Discharge Estimation: Techniques and Equipment. *Geomorphological Techniques*.

Grayson, R. B., Moore, I. D., & McMahon, T. A. (1992). Physically based hydrologic modeling: 2. Is the concept realistic? *Water Resources Research*, *28*(10), 2659–2666. https://doi.org/10.1029/92WR01259

Gupta, H. V., Kling, H., Yilmaz, K. K., & Martinez, G. F. (2009). Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modelling. *Journal of Hydrology*, *377*(1–2), 80–91. https://doi.org/10.1016/j.jhydrol.2009.08.003

Hamed, K. H., & Ramachandra Rao, A. (1998). A modified Mann-Kendall trend test for autocorrelated data. *Journal of Hydrology*, *204*(1–4), 182–196. https://doi.org/10.1016/S0022-1694(97)00125-X

Hameed, M. M., Masood, A., Hamid, A., Elbeltagi, A., Razali, S. F. M., & Salem, A. (2025). Forecasting monthly runoff in a glacierized catchment: A comparison of extreme gradient boosting (XGBoost) and deep learning models. *PLOS One*, *20*(5), e0321008. https://doi.org/10.1371/journal.pone.0321008

Hao, R., & Yan, H. (2025). Towards Interpreting Machine-Learning Models for Multi-Step Ahead Daily Streamflow Forecasting. *Hydrological Processes*, *39*(5). https://doi.org/10.1002/hyp.70163

Herath, H. M. V. V., Chadalawada, J., & Babovic, V. (2021). Hydrologically informed machine learning for rainfall–runoff modelling: towards distributed modelling. *Hydrology and Earth System Sciences*, *25*(8), 4373–4401. https://doi.org/10.5194/hess-25-4373-2021

Hochreiter, S., & Schmidhuber, J. (1997). Long Short-Term Memory. *Neural Computation*, *9*(8), 1735–1780. https://doi.org/10.1162/neco.1997.9.8.1735

Hodson, T. O. (2022). Root-mean-square error (RMSE) or mean absolute error (MAE): when

to use them or not. *Geoscientific Model Development*, *15*(14), 5481–5487. https://doi.org/10.5194/gmd-15-5481-2022

Hu, J., & Szymczak, S. (2023). A review on longitudinal data analysis with random forest. *Briefings in Bioinformatics*, *24*(2). https://doi.org/10.1093/bib/bbad002

Hu, Y., Zhou, Y., Wang, Y., Lu, F., Xiao, W., Hou, B., Yu, Y., Liu, J., & Xue, W. (2022). Impacts of Precipitation Type Variations on Runoff Changes in the Source Regions of the Yangtze and Yellow River Basins in the Past 40 Years. *Water*, *14*(24), 4115. https://doi.org/10.3390/w14244115

Hunt, K. M. R., Matthews, G. R., Pappenberger, F., & Prudhomme, C. (2022). Using a long short-term memory (LSTM) neural network to boost river streamflow forecasts over the western United States. *Hydrology and Earth System Sciences*, *26*(21), 5449–5472. https://doi.org/10.5194/hess-26-5449-2022

Jacomazzi, M. A., Zuffo, A. C., Imteaz, M. A., Boulomytis, V. T. G., Galbetti, M. V., & Shinma, T. A. (2022). Maximum Extreme Flow Estimations in Historical Hydrological Series under the Influence of Decadal Variations. *Hydrology*, *9*(8), 130. https://doi.org/10.3390/hydrology9080130

Jaiswal, R. K., Ali, S., & Bharti, B. (2020). Comparative evaluation of conceptual and physical rainfall–runoff models. *Applied Water Science*, *10*(1), 48. https://doi.org/10.1007/s13201-019-1122-6

Jesus, E. T. de, Amorim, J. da S., Junqueira, R., Viola, M. R., & Mello, C. R. de. (2020). Meteorological and hydrological drought from 1987 to 2017 in Doce River Basin, Southeastern Brazil. *RBRH*, *25*. https://doi.org/10.1590/2318-0331.252020190181

Jeung, M., Baek, S., Beom, J., Cho, K. H., Her, Y., & Yoon, K. (2019). Evaluation of random forest and regression tree methods for estimation of mass first flush ratio in urban catchments. *Journal of Hydrology*, *575*, 1099–1110. https://doi.org/10.1016/j.jhydrol.2019.05.079

Jung, M. S., da Silva, J. A. G., Fachinetto, J. M., Carvalho, I. R., Lucchese, O. A., Basso, N. C. F., Copetti, C. M., & da Silva, L. G. (2023). Water: A fundamental resource for ensuring sustainability. *Revista de Gestão Social e Ambiental*, *17*(7), 1–17.

Kahya, E., & Kalaycı, S. (2004). Trend analysis of streamflow in Turkey. *Journal of Hydrology*, *289*(1–4), 128–144. https://doi.org/10.1016/j.jhydrol.2003.11.006

Kartal, V. (2024). Machine learning-based streamflow forecasting using <scp>CMIP6</scp> scenarios: Assessing performance and improving hydrological projections and climate change. *Hydrological Processes*, *38*(6). https://doi.org/10.1002/hyp.15204

Karunasingha, D. S. K. (2022). Root mean square error or mean absolute error? Use their ratio as well. *Information Sciences*, *585*, 609–629. https://doi.org/10.1016/j.ins.2021.11.036

Kedam, N., Tiwari, D. K., Kumar, V., Khedher, K. M., & Salem, M. A. (2024). River stream flow prediction through advanced machine learning models for enhanced accuracy. *Results in Engineering*, *22*, 102215. https://doi.org/10.1016/j.rineng.2024.102215

Kendall, M. G. (1948). *Rank correlation methods.*

Kim, D., Jang, W., Han, H., Shin, H.-S., Kim, H., & Jang, C. (2026). Multi-Model Comparison of Hydrologic Simulation Performance Using DWAT, PRMS, and TANK Models. *Water*, *18*(2), 145. https://doi.org/10.3390/w18020145

Kratzert, F., Klotz, D., Brenner, C., Schulz, K., & Herrnegger, M. (2018). Rainfall–runoff modelling using Long Short-Term Memory (LSTM) networks. *Hydrology and Earth System Sciences*, *22*(11), 6005–6022. https://doi.org/10.5194/hess-22-6005-2018

Küçüktopcu, E., Cemek, E., Cemek, B., & Simsek, H. (2023). Hybrid Statistical and Machine Learning Methods for Daily Evapotranspiration Modeling. *Sustainability*, *15*(7), 5689. https://doi.org/10.3390/su15075689

Kumar, V., Kedam, N., Sharma, K. V., Mehta, D. J., & Caloiero, T. (2023). Advanced Machine Learning Techniques to Improve Hydrological Prediction: A Comparative Analysis of Streamflow Prediction Models. *Water*, *15*(14), 2572. https://doi.org/10.3390/w15142572

Lakew, H. B., Haile, A. T., Taye, M. T., Asfaw, W., & Tedla, H. Z. (2026). Recovering daily streamflow data gaps in the Lake Tana sub-basin: leveraging machine learning and remote sensing techniques in Ethiopia. *Journal of Water and Climate Change*. https://doi.org/10.2166/wcc.2026.319

Lee, H., McIntyre, N., Wheater, H., & Young, A. (2005). Selection of conceptual models for regionalisation of the rainfall-runoff relationship. *Journal of Hydrology*, *312*(1–4), 125–147. https://doi.org/10.1016/j.jhydrol.2005.02.016

Lees, T., Buechel, M., Anderson, B., Slater, L., Reece, S., Coxon, G., & Dadson, S. J. (2021). Benchmarking data-driven rainfall–runoff models in Great Britain: a comparison of long short-term memory (LSTM)-based models with four lumped conceptual models. *Hydrology and Earth System Sciences*, *25*(10), 5517–5534. https://doi.org/10.5194/hess-25-5517-2021

Leong, C., & Yokoo, Y. (2021). A step toward global-scale applicability and transferability of flow duration curve studies: A flow duration curve review (2000–2020). *Journal of*

*Hydrology*, *603*, 126984. https://doi.org/10.1016/j.jhydrol.2021.126984

Leong, C., & Yokoo, Y. (2022). A multiple hydrograph separation technique for identifying hydrological model structures and an interpretation of dominant process controls on flow duration curves. *Hydrological Processes*, *36*(4). https://doi.org/10.1002/hyp.14569

Li, W., Kiaghadi, A., & Dawson, C. (2021). High temporal resolution rainfall–runoff modeling using long-short-term-memory (LSTM) networks. *Neural Computing and Applications*, *33*(4), 1261–1278. https://doi.org/10.1007/s00521-020-05010-6

Lima, J. M., & Cupolillo, F. (2018). ANÁLISE ESPAÇO-TEMPORAL DAS CHUVAS PERSISTENTES NA REGIÃO DO PARQUE ESTADUAL DO RIO DOCE–PERD, SOB INFLUÊNCIAS DAS ZCAS E ZCOU (OUT/2015 A DEZ/2016). *Revista Brasileira de Climatologia*, *1*. https://doi.org/10.5380/abclima.v1i0.59298

Liu, H., Yan, H., & Guan, M. (2025). Evaluating the effects of topography and land use change on hydrological signatures: a comparative study of two adjacent watersheds. *Hydrology and Earth System Sciences*, *29*(8), 2109–2132. https://doi.org/10.5194/hess-29-2109-2025

Liu, W., Hao, H., Song, X., Wang, Q., Hao, Y., Zhang, J., Huang, X., & Yan, X. (2025). Investigating hydrological processes using explainable deep-learning models. *Hydrological Sciences Journal*, *70*(1), 110–126. https://doi.org/10.1080/02626667.2024.2423050

Liu, X., Zhang, X., Kong, X., & Shen, Y.-J. (2022). Random Forest Model Has the Potential for Runoff Simulation and Attribution. *Water*, *14*(13), 2053. https://doi.org/10.3390/w14132053

Lopes Martins, L., Martins, W. A., Rodrigues, I. C. de A., Freitas Xavier, A. C., Moraes, J. F. L. de, & Blain, G. C. (2023). Gap-filling of daily precipitation and streamflow time series: a method comparison at random and sequential gaps. *Hydrological Sciences Journal*, *68*(1), 148–160. https://doi.org/10.1080/02626667.2022.2145200

Loritz, R., Hrachowitz, M., Neuper, M., & Zehe, E. (2021). The role and value of distributed precipitation data in hydrological models. *Hydrology and Earth System Sciences*, *25*(1), 147–167. https://doi.org/10.5194/hess-25-147-2021

Lu, D., Konapala, G., Painter, S. L., Kao, S.-C., & Gangrade, S. (2021). Streamflow simulation in data-scarce basins using Bayesian and physics-informed machine learning models. *Journal of Hydrometeorology*. https://doi.org/10.1175/JHM-D-20-0082.1

Lukas, P., Melesse, A. M., & Kenea, T. T. (2024). Predicting reservoir sedimentation using multilayer perceptron – Artificial neural network model with measured and forecasted

hydrometeorological data in Gibe-III reservoir, Omo-Gibe River basin, Ethiopia. *Journal of Environmental Management*, *359*, 121018. https://doi.org/10.1016/j.jenvman.2024.121018

Maggiotto, S. R., Dias, N. L., Vissotto, S., Prevedello, C. L., Nascimento, E. L., Grodzki, L., & de Souza, M. F. (2007). Diretrizes para redes automáticas e telemétricas de superfície. *Revista Brasileira de Recursos Hídricos*, *12*, 225–240.

Magyari-Sáska, Z., Haidu, I., & Magyari-Sáska, A. (2025). Experimental Comparative Study on Self-Imputation Methods and Their Quality Assessment for Monthly River Flow Data with Gaps: Case Study to Mures River. *Applied Sciences*, *15*(3), 1242. https://doi.org/10.3390/app15031242

Mann, H. B. (1945). Nonparametric Tests Against Trend. *Econometrica*, *13*(3), 245. https://doi.org/10.2307/1907187

Marasini, U., & Pokhrel, M. (2024). Comparative analysis of rainfall-runoff simulation using a long short-term memory (LSTM) deep learning model and a conceptual hydrological model, HEC-HMS: a case study of the mountainous river basin of Nepal. *Discover Civil Engineering*, *1*(1), 78. https://doi.org/10.1007/s44290-024-00084-w

Marcos Junior, A. D., Silveira, C. da S., Costa, J. M. F. da, & Gonçalves, S. T. N. (2024). Combining traditional hydrological models and machine learning for streamflow prediction. *RBRH*, *29*. https://doi.org/10.1590/2318-0331.292420230105

MARCUZZO, F. F. N., Romero, V., CARDOSO, M. R. D., & PINTO FILHO, R. de F. (2011). *Detalhamento hidromorfológico da bacia do rio Doce*.

Martins, F. B., Gonzaga, G., Dos Santos, D. F., & Reboita, M. S. (2018). CLASSIFICAÇÃO CLIMÁTICA DE KÖPPEN E DE THORNTHWAITE PARA MINAS GERAIS: CENÁRIO ATUAL E PROJEÇÕES FUTURAS. *Revista Brasileira de Climatologia*, *1*. https://doi.org/10.5380/abclima.v1i0.60896

Materia, S., Ardilouze, C., Prodhomme, C., Donat, M. G., Benassi, M., Doblas-Reyes, F. J., Peano, D., Caron, L.-P., Ruggieri, P., & Gualdi, S. (2022). Summer temperature response to extreme soil water conditions in the Mediterranean transitional climate regime. *Climate Dynamics*, *58*(7–8), 1943–1963. https://doi.org/10.1007/s00382-021-05815-8

Mazi, K., Koussis, A. D., Lykoudis, S., Psiloglou, B. E., Vitantzakis, G., Kappos, N., Katsanos, D., Rozos, E., Koletsis, I., & Kopania, T. (2023). Establishing and Operating (Pilot Phase) a Telemetric Streamflow Monitoring Network in Greece. *Hydrology*, *10*(1), 19. https://doi.org/10.3390/hydrology10010019

Mei, X., Smith, P. K., Li, J., & Li, B. (2022). Hydrological evaluation of gridded climate datasets in a texas urban watershed using soil and water assessment tool and artificial neural network. *Frontiers in Environmental Science*, *10*. https://doi.org/10.3389/fenvs.2022.905774

Mendonça, L., Gomide, I., Sousa, J., & Blanco, C. (2021). Modelagem chuva-vazão via redes neurais artificiais para simulação de vazões de uma bacia hidrográfica da Amazônia. *Revista de Gestão de Água Da América Latina*, *18*(1), 2–0. https://doi.org/10.21168/rega.v18e2

Miranda, C. C., Maia, J. L., Silva, J. P. O., & Vieira, E. M. (2021). Análise da cobrança pelo uso de recursos hídricos em águas de domínio da União: Estudo de caso da bacia hidrográfica do rio Doce (MG). *Research, Society and Development*, *10*(4), e5610413785. https://doi.org/10.33448/rsd-v10i4.13785

Montini, T. L., Jones, C., & Carvalho, L. M. V. (2019). The South American Low-Level Jet: A New Climatology, Variability, and Changes. *Journal of Geophysical Research: Atmospheres*, *124*(3), 1200–1218. https://doi.org/10.1029/2018JD029634

Moreira, G. de B., Maia, J. L., & Silva, J. P. O. (2022). Disponibilidade hídrica superficial da bacia hidrográfica do rio Suaçuí: análise por diferentes metodologias de definição de vazão de referência. *Research, Society and Development*, *11*(3), e0211325987. https://doi.org/10.33448/rsd-v11i3.25987

Moreira, R. M., dos Santos, B. C., Biggs, T., de Sales, F., & Sieber, S. (2024). Identifying clusters of precipitation for the Brazilian Legal Amazon based on magnitude of trends and its correlation with sea surface temperature. *Scientific Reports*, *14*(1), 14067. https://doi.org/10.1038/s41598-024-63583-x

Muhammad Ali, P. J. (2022). Investigating the Impact of Min-Max Data Normalization on the Regression Performance of K-Nearest Neighbor with Different Similarity Measurements. *ARO-THE SCIENTIFIC JOURNAL OF KOYA UNIVERSITY*, *10*(1), 85–91. https://doi.org/10.14500/aro.10955

Mushtaq, H., Akhtar, T., Hashmi, M. Z. ur R., Masood, A., & Saeed, F. (2024). Hydrologic interpretation of machine learning models for 10-daily streamflow simulation in climate sensitive upper Indus catchments. *Theoretical and Applied Climatology*, *155*(6), 5525–5542. https://doi.org/10.1007/s00704-024-04932-8

Mwaguni, S., Ruwa, R. K., Odhiambo-Ochiewo, J., & Osore, M. (2016). *Integrated Water Resources Management in a Changing Climate: The Implication of Anthropogenic Activities on the Tana and Athi/Sabaki Rivers Water System for Sustainable*

*Development* (pp. 111–129). https://doi.org/10.1007/978-3-319-25370-1_7

Nazeri Tahroudi, M., Ahmadi, F., & Mirabbasi, R. (2023). Performance comparison of IHACRES, random forest and copula-based models in rainfall-runoff simulation. *Applied Water Science*, *13*(6), 134. https://doi.org/10.1007/s13201-023-01929-y

Nearing, G., Cohen, D., Dube, V., Gauch, M., Gilon, O., Harrigan, S., Hassidim, A., Klotz, D., Kratzert, F., Metzger, A., Nevo, S., Pappenberger, F., Prudhomme, C., Shalev, G., Shenzis, S., Tekalign, T. Y., Weitzner, D., & Matias, Y. (2024). Global prediction of extreme floods in ungauged watersheds. *Nature*, *627*(8004), 559–563. https://doi.org/10.1038/s41586-024-07145-1

Ni, L., Wang, D., Wu, J., Wang, Y., Tao, Y., Zhang, J., & Liu, J. (2020). Streamflow forecasting using extreme gradient boosting model coupled with Gaussian mixture model. *Journal of Hydrology*, *586*, 124901. https://doi.org/10.1016/j.jhydrol.2020.124901

Nicolete, D. A. P., & Piroli, E. L. (2024). *Geoprocessamento no Estudo de Bacias Hidrográficas* (A. A. da N. da A. P. ANAP (ed.)). https://www.researchgate.net/profile/Edson-Piroli/publication/385943297_Geoprocessamento_no_Estudo_de_Bacias_Hidrograficas_Praticas_com_QGIS/links/673d1839b94c451c11667d22/Geoprocessamento-no-Estudo-de-Bacias-Hidrograficas-Praticas-com-QGIS.pdf

Nourani, V., Gökçekuş, H., & Gichamo, T. (2021). Ensemble data-driven rainfall-runoff modeling using multi-source satellite and gauge rainfall data input fusion. *Earth Science Informatics*, *14*(4), 1787–1808. https://doi.org/10.1007/s12145-021-00615-4

Okkan, U., Ersoy, Z. B., Ali Kumanlioglu, A., & Fistikoglu, O. (2021). Embedding machine learning techniques into a conceptual model to improve monthly runoff simulation: A nested hybrid rainfall-runoff modeling. *Journal of Hydrology*, *598*, 126433. https://doi.org/10.1016/j.jhydrol.2021.126433

Ouali, D., Chebana, F., & Ouarda, T. B. M. J. (2017). Fully nonlinear statistical and machine-learning approaches for hydrological frequency estimation at ungauged sites. *Journal of Advances in Modeling Earth Systems*, *9*(2), 1292–1306. https://doi.org/10.1002/2016MS000830

Parisouj, P., Jun, C., Bateni, S. M., Heggy, E., & Band, S. S. (2024). Daily runoff forecasting using novel optimized machine learning methods. *Results in Engineering*, *24*, 103319. https://doi.org/10.1016/j.rineng.2024.103319

Pathiraja, S., Anghileri, D., Burlando, P., Sharma, A., Marshall, L., & Moradkhani, H. (2018).

Time-varying parameter models for catchments with land use change: the importance of model structure. *Hydrology and Earth System Sciences*, *22*(5), 2903–2919. https://doi.org/10.5194/hess-22-2903-2018

Penna, D., Tromp-van Meerveld, H. J., Gobbi, A., Borga, M., & Dalla Fontana, G. (2011). The influence of soil moisture on threshold runoff generation processes in an alpine headwater catchment. *Hydrology and Earth System Sciences*, *15*(3), 689–702. https://doi.org/10.5194/hess-15-689-2011

Pereira, D. R., Oliveira, A. R., Costa, M. S., Ramos, T. B., Rollnic, M., & Neves, R. J. J. (2024). Evaluation of precipitation products in a Brazilian watershed: Tocantins-Araguaia watershed case study. *Theoretical and Applied Climatology*, *155*(8), 7845–7865. https://doi.org/10.1007/s00704-024-05091-6

Persiano, S., Pugliese, A., Aloe, A., Skøien, J. O., Castellarin, A., & Pistocchi, A. (2022). Streamflow data availability in Europe: a detailed dataset of interpolated flow-duration curves. *Earth System Science Data*, *14*(9), 4435–4443. https://doi.org/10.5194/essd-14-4435-2022

Pettitt, A. N. (1979). A Non-Parametric Approach to the Change-Point Problem. *Applied Statistics*, *28*(2), 126. https://doi.org/10.2307/2346729

Piotrowski, A. P., & Napiorkowski, J. J. (2013). A comparison of methods to avoid overfitting in neural networks training in the case of catchment runoff modelling. *Journal of Hydrology*, *476*, 97–111. https://doi.org/10.1016/j.jhydrol.2012.10.019

Pool, S., Vis, M., & Seibert, J. (2018). Evaluating model performance: towards a non-parametric variant of the Kling-Gupta efficiency. *Hydrological Sciences Journal*, *63*(13–14), 1941–1953. https://doi.org/10.1080/02626667.2018.1552002

Pulgarín-Morales, L., Schwenzer, J., & Krueger, T. (2026). Machine learning-based streamflow forecasting in Colombian hydropower basins at daily and monthly scales. *Hydrology Research*, *57*(4), 804–821. https://doi.org/10.2166/nh.2026.236

Rahimzad, M., Moghaddam Nia, A., Zolfonoon, H., Soltani, J., Danandeh Mehr, A., & Kwon, H.-H. (2021). Performance Comparison of an LSTM-based Deep Learning Model versus Conventional Machine Learning Algorithms for Streamflow Forecasting. *Water Resources Management*, *35*(12), 4167–4187. https://doi.org/10.1007/s11269-021-02937-w

Rahmad Ramadhan, L., & Anne Mudya, Y. (2024). A comparative study of Z-score and min-max normalization for rainfall classification in Pekanbaru. *Journal of Data Science*, *2024*(04), 1–8.

Reboita, M. S., Ambrizzi, T., Crespo, N. M., Dutra, L. M. M., Ferreira, G. W. de S., Rehbein, A., Drumond, A., da Rocha, R. P., & Souza, C. A. de. (2021). Impacts of teleconnection patterns on South America climate. *Annals of the New York Academy of Sciences*, *1504*(1), 116–153. https://doi.org/10.1111/nyas.14592

Roza dos Santos, J. G., Vespucci, A. G., & Bayer, M. (2017). Qualificação dos dados hidrológicos disponíveis na base HIDROWEB/ANA: estações fluviométricas do estado de Goiás. *Ateliê Geográfico*, *10*(3), 89. https://doi.org/10.5216/ag.v10i3.35554

Rudra, A. K., & Alam, A. K. M. R. (2023). Streamflow characteristics of Sangu-Matamuhuri watershed in the southeastern part of Bangladesh. *Heliyon*, *9*(3), e14559. https://doi.org/10.1016/j.heliyon.2023.e14559

Rybski, D., & Neumann, J. (2011). A Review on the Pettitt Test. In *In Extremis* (pp. 202–213). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-14863-7_10

Saadi, M., Oudin, L., & Ribstein, P. (2019). Random Forest Ability in Regionalizing Hourly Hydrological Model Parameters. *Water*, *11*(8), 1540. https://doi.org/10.3390/w11081540

Sabzipour, B., Arsenault, R., Troin, M., Martel, J.-L., Brissette, F., Brunet, F., & Mai, J. (2023). Comparing a long short-term memory (LSTM) neural network with a physically-based hydrological model for streamflow forecasting over a Canadian catchment. *Journal of Hydrology*, *627*, 130380. https://doi.org/10.1016/j.jhydrol.2023.130380

Sahoo, A., Samantaray, S., & Ghose, D. K. (2022). Multilayer perceptron and support vector machine trained with grey wolf optimiser for predicting floods in Barak river, India. *Journal of Earth System Science*, *131*(2), 85. https://doi.org/10.1007/s12040-022-01815-2

Sánchez, H. G., Loaiza Usuga, J. C., & Vélez Upégui, J. I. (2024). Hydrological Response to Predominant Land Use and Land Cover in the Colombian Andes at the Micro-Watershed Scale. *Land*, *13*(8), 1140. https://doi.org/10.3390/land13081140

Sarigil, G., Cavus, Y., Aksoy, H., & Eris, E. (2024). Frequency curves of high and low flows in intermittent river basins for hydrological analysis and hydraulic design. *Stochastic Environmental Research and Risk Assessment*, *38*(8), 3079–3092. https://doi.org/10.1007/s00477-024-02732-0

Sauer, V. B., & Turnipseed, D. P. (2010). *Stage measurement at gaging stations*. US Geological Survey.

Scaife, C. I., & Band, L. E. (2017). Nonstationarity in threshold response of stormflow in southern <scp>A</scp> ppalachian headwater catchments. *Water Resources Research*, *53*(8), 6579–6596. https://doi.org/10.1002/2017WR020376

Schoener, G., & Stone, M. C. (2019). Impact of antecedent soil moisture on runoff from a semiarid catchment. *Journal of Hydrology*, *569*, 627–636. https://doi.org/10.1016/j.jhydrol.2018.12.025

Schoppa, L., Disse, M., & Bachmair, S. (2020). Evaluating the performance of random forest for large-scale flood discharge simulation. *Journal of Hydrology*, *590*, 125531. https://doi.org/10.1016/j.jhydrol.2020.125531

Sedighi, F., Vafakhah, M., & Javadi, M. R. (2016). Rainfall–Runoff Modeling Using Support Vector Machine in Snow-Affected Watershed. *Arabian Journal for Science and Engineering*, *41*(10), 4065–4076. https://doi.org/10.1007/s13369-016-2095-5

Seibert, J. (1999). Regionalisation of parameters for a conceptual rainfall-runoff model. *Agricultural and Forest Meteorology*, *98–99*, 279–293. https://doi.org/10.1016/S0168-1923(99)00105-7

Senthil Kumar, A. R., Sudheer, K. P., Jain, S. K., & Agarwal, P. K. (2005). Rainfall-runoff modelling using artificial neural networks: comparison of network types. *Hydrological Processes*, *19*(6), 1277–1291. https://doi.org/10.1002/hyp.5581

Shen, Y., Ruijsch, J., Lu, M., Sutanudjaja, E. H., & Karssenberg, D. (2022). Random forests-based error-correction of streamflow from a large-scale hydrological model: Using model state variables to estimate error terms. *Computers & Geosciences*, *159*, 105019. https://doi.org/10.1016/j.cageo.2021.105019

Shin, M.-J., & Jung, Y. (2025). Comparison of Multilayer Perceptron with an Optimal Activation Function and Long Short-Term Memory for Rainfall-Runoff Simulations and Ungauged Catchment Runoff Prediction. *Water Resources Management*, *39*(6), 2479–2502. https://doi.org/10.1007/s11269-024-04074-6

Silva, L. A., Silva, A. M., Coelho, G., Pinto, L. C., & Eduardo, E. N. (2017). Vazões mínimas e de referência e rendimento específico para o estado de Minas Gerais. *Revista Brasileira de Ciências Agrárias - Brazilian Journal of Agricultural Sciences*, *12*(4), 543–549. https://doi.org/10.5039/agraria.v12i4a5467

Smakhtin, V. . (2001). Low flow hydrology: a review. *Journal of Hydrology*, *240*(3–4), 147–186. https://doi.org/10.1016/S0022-1694(00)00340-1

Sneyers, R. (1991). *On the statistical analysis of series of observations.* (Issue 143).

Solanki, H., & Mishra, V. (2026). Filling Streamflow Data Gaps in Indian Catchments Using Machine Learning and K-Means Clustering. *Water Resources Research*, *62*(3). https://doi.org/10.1029/2025WR042210

Solomatine, D. P., & Ostfeld, A. (2008). Data-driven modelling: some past experiences and

new approaches. *Journal of Hydroinformatics*, *10*(1), 3–22. https://doi.org/10.2166/hydro.2008.015

Sondermann, M., Chou, S. C., Lyra, A., Latinovic, D., Siqueira, G. C., Junior, W. C., Giornes, E., & Leite, F. P. (2022). Climate change projections and impacts on the eucalyptus plantation around the Doce River basin, in Minas Gerais, Brazil. *Climate Services*, *28*, 100327. https://doi.org/10.1016/j.cliser.2022.100327

Song, S., & Wang, W. (2019). Impacts of Antecedent Soil Moisture on the Rainfall-Runoff Transformation Process Based on High-Resolution Observations in Soil Tank Experiments. *Water*, *11*(2), 296. https://doi.org/10.3390/w11020296

Sumith, K. V. (2025). Comprehensive Evaluation of Satellite-Based Rainfall Measurements Through Rain Gauge Validation Using Advanced Statistical Regression and Machine Learning Models by Using Python. *Water Resources Management*, *39*(9), 4563–4587. https://doi.org/10.1007/s11269-025-04168-9

Tayman, J., & Swanson, D. A. (1999). On the validity of MAPE as a measure of population forecast accuracy. *Population Research and Policy Review*, *18*(4), 299–322. https://doi.org/10.1023/A:1006166418051

Terzi, Ö., Küçüksille, E. U., Baykal, T., & Taylan, E. D. (2023). Deep and machine learning for daily streamflow estimation: a focus on LSTM, RFR and XGBoost. *Water Practice & Technology*, *18*(10), 2401–2414. https://doi.org/10.2166/wpt.2023.144

Thom, H. C. S. (1966). *Some methods of climatological analysis* (Vol. 81). Secretariat of the World Meteorological Organization Geneva.

Tongal, H., & Booij, M. J. (2018). Simulation and forecasting of streamflows using machine learning models coupled with base flow separation. *Journal of Hydrology*, *564*, 266–282. https://doi.org/10.1016/j.jhydrol.2018.07.004

Towler, E., Foks, S. S., Dugger, A. L., Dickinson, J. E., Essaid, H. I., Gochis, D., Viger, R. J., & Zhang, Y. (2023). Benchmarking high-resolution hydrologic model performance of long-term retrospective streamflow simulations in the contiguous United States. *Hydrology and Earth System Sciences*, *27*(9), 1809–1825. https://doi.org/10.5194/hess-27-1809-2023

Tuladhar, D., Dewan, A., Kuhn, M., & J. Corner, R. (2019). The Influence of Rainfall and Land Use/Land Cover Changes on River Discharge Variability in the Mountainous Catchment of the Bagmati River. *Water*, *11*(12), 2444. https://doi.org/10.3390/w11122444

Valdés-Pineda, R., Pizarro, R., García-Chevesich, P., Valdés, J. B., Olivares, C., Vera, M.,

Balocchi, F., Pérez, F., Vallejos, C., Fuentes, R., Abarza, A., & Helwig, B. (2014). Water governance in Chile: Availability, management and climate change. *Journal of Hydrology*, *519*, 2538–2567. https://doi.org/10.1016/j.jhydrol.2014.04.016

Valles, J. (2024). *Application of a Multilayer Perceptron Artificial Neural Network (MLP-ANN) in Hydrological Forecasting in El Salvador* (pp. 213–239). https://doi.org/10.1002/9781119639268.ch8

Vaze, J., Jordan, P., Beecham, R., Frost, A., & Summerell, G. (2011). Guidelines for rainfall-runoff modelling: towards best practice model application. *Australia: EWater Cooprative Research Centre*.

Velásquez, N., & Krajewski, W. F. (2024). Effect of streamflow measurement error on flood frequency estimation. *Stochastic Environmental Research and Risk Assessment*, *38*(7), 2903–2910. https://doi.org/10.1007/s00477-024-02707-1

Vilanova, M. R. N., Simoes, S. J. C., & Trannin, I. C. de B. (2012). Interpolação geoespacial da evapotranspiracao de referencia (ETo) em regioes com escassez de dados: estudo de caso no Sul de Minas Gerais, Brasil. *Ambiente e Agua - An Interdisciplinary Journal of Applied Science*, *7*(2), 179–194. https://doi.org/10.4136/ambi-agua.763

Vilaseca, F., Castro, A., Chreties, C., & Gorgoglione, A. (2023). Assessing influential rainfall–runoff variables to simulate daily streamflow using random forest. *Hydrological Sciences Journal*, *68*(12), 1738–1753. https://doi.org/10.1080/02626667.2023.2232356

Wang, S., Peng, H., Hu, Q., & Jiang, M. (2022). Analysis of runoff generation driving factors based on hydrological model and interpretable machine learning method. *Journal of Hydrology: Regional Studies*, *42*, 101139. https://doi.org/10.1016/j.ejrh.2022.101139

Wang, W., & Lu, Y. (2018). Analysis of the Mean Absolute Error (MAE) and the Root Mean Square Error (RMSE) in Assessing Rounding Model. *IOP Conference Series: Materials Science and Engineering*, *324*, 012049. https://doi.org/10.1088/1757-899X/324/1/012049

Warton, D. I., Wright, S. T., & Wang, Y. (2012). Distance-based multivariate analyses confound location and dispersion effects. *Methods in Ecology and Evolution*, *3*(1), 89–101. https://doi.org/10.1111/j.2041-210X.2011.00127.x

Wilby, R. L., Clifford, N. J., De Luca, P., Harrigan, S., Hillier, J. K., Hodgkins, R., Johnson, M. F., Matthews, T. K. R., Murphy, C., Noone, S. J., Parry, S., Prudhomme, C., Rice, S. P., Slater, L. J., Smith, K. A., & Wood, P. J. (2017). The 'dirty dozen' of freshwater science: detecting then reconciling hydrological data biases and errors. *WIREs Water*, *4*(3). https://doi.org/10.1002/wat2.1209

Wilcoxon, F., Katti, S. K., & Wilcox, R. A. (1963). *Critical values and probability levels for the Wilcoxon rank sum test and the Wilcoxon signed rank test* (Vol. 1). American Cyanamid Pearl River, NY.

Willmott, C., & Matsuura, K. (2005). Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance. *Climate Research*, *30*, 79–82. https://doi.org/10.3354/cr030079

Witt, T. J. (2007). Using the autocorrelation function to characterize time series of voltage measurements. *Metrologia*, *44*(3), 201–209. https://doi.org/10.1088/0026-1394/44/3/006

Xie, M., Ren, Z., Li, Z., Zhang, X., Ma, X., Li, P., & Shen, Z. (2023). Evolution of the precipitation–stream runoff relationship in different precipitation scenarios in the Yellow River Basin. *Urban Climate*, *51*, 101609. https://doi.org/10.1016/j.uclim.2023.101609

Xiong, M., Liu, P., Cheng, L., Deng, C., Gui, Z., Zhang, X., & Liu, Y. (2019). Identifying time-varying hydrological model parameters to improve simulation efficiency by the ensemble Kalman filter: A joint assimilation of streamflow and actual evapotranspiration. *Journal of Hydrology*, *568*, 758–768. https://doi.org/10.1016/j.jhydrol.2018.11.038

Ye, S., Liu, L., Li, J., Pan, H., Li, W., & Ran, Q. (2023). From rainfall to runoff: The role of soil moisture in a mountainous catchment. *Journal of Hydrology*, *625*, 130060. https://doi.org/10.1016/j.jhydrol.2023.130060

Zhang, Y., Zhang, Y., Zhai, X., Xia, J., Tang, Q., Wang, W., Wu, J., Niu, X., & Han, B. (2025). Understanding meteorological and physio-geographical controls of variability of flood event classes in headstream catchments of China. *Hydrology and Earth System Sciences*, *29*(14), 3257–3275. https://doi.org/10.5194/hess-29-3257-2025

Zhang, Z., Xiao, Y., Chen, R., Long, K., Deng, H., Zheng, Z., & Liao, J. (2026). Comparative assessment of machine learning models for daily streamflow prediction in a subtropical monsoon watershed. *Scientific Reports*, *16*(1), 7341. https://doi.org/10.1038/s41598-026-38969-8