

## XXVI SIMPÓSIO BRASILEIRO DE RECURSOS HÍDRICOS

### **ANÁLISE DA PREVISÃO DE VAZÕES UTILIZANDO MODELAGEM ESTOCÁSTICA NO RIO SEPOTUBA**

*José Rodrigues da Silva Neto<sup>1</sup>; Jadson Vinícius Melo Santos<sup>2</sup> & Frederico Carlos Martins de Menezes Filho<sup>3</sup>*

**Abstract:** Streamflow forecasting is a widely used and essential technique for assessing the future availability of water resources, serving as an indicator of the consequences of unsustainable consumption. In this context, stochastic models—particularly suitable for data-scarce basins—offer effective predictive capabilities without requiring large volumes of detailed input data. Among these, the SARIMA and SARIMAX models were employed in this study. Both are seasonal autoregressive integrated moving average models; however, SARIMAX incorporates an exogenous regressor. This study aimed to evaluate streamflow forecasts using ARIMA-family models in a sub-basin of the Sepotuba River, located in the state of Mato Grosso, Brazil. One seasonal model (SARIMA) and three SARIMAX models—with precipitation, temperature, and both as exogenous variables—were developed. The models were trained using 11 years of data and tested over one year. Results indicated that the SARIMAX model with temperature as the exogenous variable performed best, achieving a Nash-Sutcliffe Efficiency (NSE) of 0.83, Kling-Gupta Efficiency (KGE) of 0.86, and Root Mean Square Error (RMSE) of  $23.74 \text{ m}^3 \cdot \text{s}^{-1}$  during the training period, and NSE of 0.92, KGE of 0.36, and RMSE of  $18.58 \text{ m}^3 \cdot \text{s}^{-1}$  during testing. These findings confirm the applicability of ARIMA-family models for streamflow forecasting under varying data availability, with SARIMAX, incorporating temperature as an exogenous variable, achieving the best performance.

**Resumo:** A previsão de vazões é uma técnica amplamente utilizada e importante para a verificação da disponibilidade futura de recursos hídricos, sinalizando as consequências de um consumo predatório de tais riquezas. Neste sentido, modelos estocásticos, principalmente para bacias com dados escassos, oferecem capacidades preditivas eficazes sem a necessidade de grandes volumes de dados de entrada. Dentre estes, os modelos SARIMA e SARIMAX foram empregados neste estudo. Ambos são modelos sazonais, autorregressivos, integrados com média móvel, porém o SARIMAX possui uma variável regressora. Assim, este trabalho objetivou avaliar as previsões de vazões utilizando modelos da família ARIMA em uma sub-bacia do Rio Sepotuba, no estado do Mato Grosso. Para isso, utilizou-se de um modelo sazonal e três modelos sazonais com variáveis regressoras, a saber precipitação, temperatura e ambas. Utilizou-se um tempo de treino de 11 anos, e 1 ano de teste, assim obtendo os modelos SARIMA e SARIMAX. Constatou-se que o modelo SARIMAX com a temperatura como variável exógena foi superior, obtendo NSE de 0,83; KGE de 0,86; RMSE de  $23,74 \text{ m}^3 \cdot \text{s}^{-1}$  para o período de treino e NSE de 0,92; KGE de 0,36; RMSE de  $18,58 \text{ m}^3 \cdot \text{s}^{-1}$  para o teste. Tais resultados confirmam a aplicabilidade dos modelos da família ARIMA para a previsão de vazões em condições variadas de disponibilidade de dados, sendo que o modelo SARIMAX, com a inclusão da temperatura como variável exógena, apresentou o melhor desempenho.

1) Universidade Federal de Viçosa, Campus Rio Paranaíba - Rodovia BR 230 KM 7, 38810-000, Rio Paranaíba – MG, jose.r.neto@ufv.br

2) Universidade Federal de Viçosa, Campus Rio Paranaíba - Rodovia BR 230 KM 7, 38810-000, Rio Paranaíba – MG, jadson.santos@ufv.br

3) Universidade Federal de Viçosa, Campus Rio Paranaíba - Rodovia BR 230 KM 7, 38810-000, Rio Paranaíba – MG, frederico.menezes@ufv.br

**Palavras-Chave:** Séries temporais, ARIMA, SARIMA.

## INTRODUÇÃO

O monitoramento da vazão de corpos hídricos é imprescindível para a compreensão do comportamento dos rios e assim definir as condições em que eles se encontram. De acordo com o Relatório Mundial das Nações Unidas sobre o Desenvolvimento dos Recursos Hídricos de 2025, existem mais de 2 bilhões de pessoas que vivem em países com alto nível de estresse hídrico, e que eventos climáticos severos estão cada vez mais frequentes e significativos (UNESCO, 2025).

A partir do acompanhamento das águas fluviais é possível obter informações sequenciais ao longo do tempo, o que pode ser caracterizado como uma série temporal (HYNDMAN e ATHANASOPOULOS, 2021). Em posse dos dados do monitoramento, é realizada a análise das informações para assim implementar nas configurações dos modelos e realizar projeções para os cursos d'água, como a vazão. O modelo ARIMA foi difundido pelo trabalho de Box e Jenkins (1976), no qual analisa as previsões para uma série temporal contendo ou não a estacionariedade.

Conforme Almeida e Serra (2017) existem 2 tipos de modelos hidrológicos, a saber: modelos determinísticos ou conceituais, sendo aqueles que utilizam de informações e características físicas da bacia para determinar a vazão, ou, modelos estocásticos, aqueles que utilizam de relações hipotéticas entre os dados. Como exemplo respectivamente, tem-se o SWAT (Soil and Water Assessment Tool), usado no trabalho de Deus *et al.* (2020) para a previsão de vazão dos rios da bacia hidrográfica Tocantins-Araguaia, e o trabalho de Caminha *et al.* (2025) previsão de precipitação e vazão a partir do modelo SARIMA no Rio Caratinga, sub-bacia do Rio Doce.

Caixeta *et al.* (2021) utilizaram o modelo SARIMA para a previsão de vazão em uma sub-bacia do Rio Paranaíba em Minas Gerais com área 7.640 km<sup>2</sup> para dados de 12 anos, obtendo uma boa modelagem para o período de teste, com NSE de 0,85 e para o treino -3,34 devido à estiagem no ano de previsão. Costa *et al.* (2023) compararam o modelo SARIMA e SARIMAX, sendo este último superior em todos os 5 períodos, a saber: treino, 3,6,9 e 12 meses de teste, em uma sub-bacia do Rio Paranaíba em Minas Gerais com área de 3.190,5 km<sup>2</sup> para dados de 9 anos. Já Junior *et al.* (2024) avaliaram 6 modelos hidrológicos, dentre eles o SARIMA e o SARIMAX. Ambos possuíram métricas válidas, principalmente nos meses de junho a agosto, no qual o SARIMAX resultou em melhor desempenho para uma bacia hidrográfica do reservatório da hidrelétrica de Três Marias em Minas Gerais com área de 52000 km<sup>2</sup>, com dados de 28 anos.

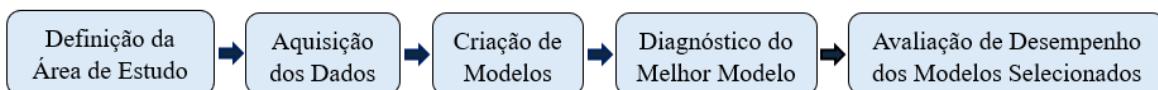
No estado do Mato Grosso se encontra a bacia do Rio Sepotuba que faz parte da Bacia do Alto Paraguai e é um afluente para as regiões alagadas do Pantanal. No trabalho de Silva *et al.* (2021), se demonstram expressivos processos antropomórficos visando o desenvolvimento agropecuário tais como supressão da vegetação nativa. Além do crescimento urbano principalmente no município de Tangará da Serra, a antropização ocorre ao longo de toda a bacia do Rio Sepotuba, interferindo diretamente na vazão do rio devido ao consumo de água na agricultura e no abastecimento público.

Deste modo, este trabalho objetivou a avaliação da previsão de vazões utilizando modelos da família ARIMA em uma sub-bacia do Rio Sepotuba, no estado do Mato Grosso, tão quanto, avaliar o desempenho da consideração de variáveis exógenas, por meio dos modelos SARIMAX.

## METODOLOGIA

No fluxograma descrito na Figura 1, apresenta-se o procedimento para a modelagem da vazão empregada neste trabalho.

Figura 1 – Fluxograma Metodológico



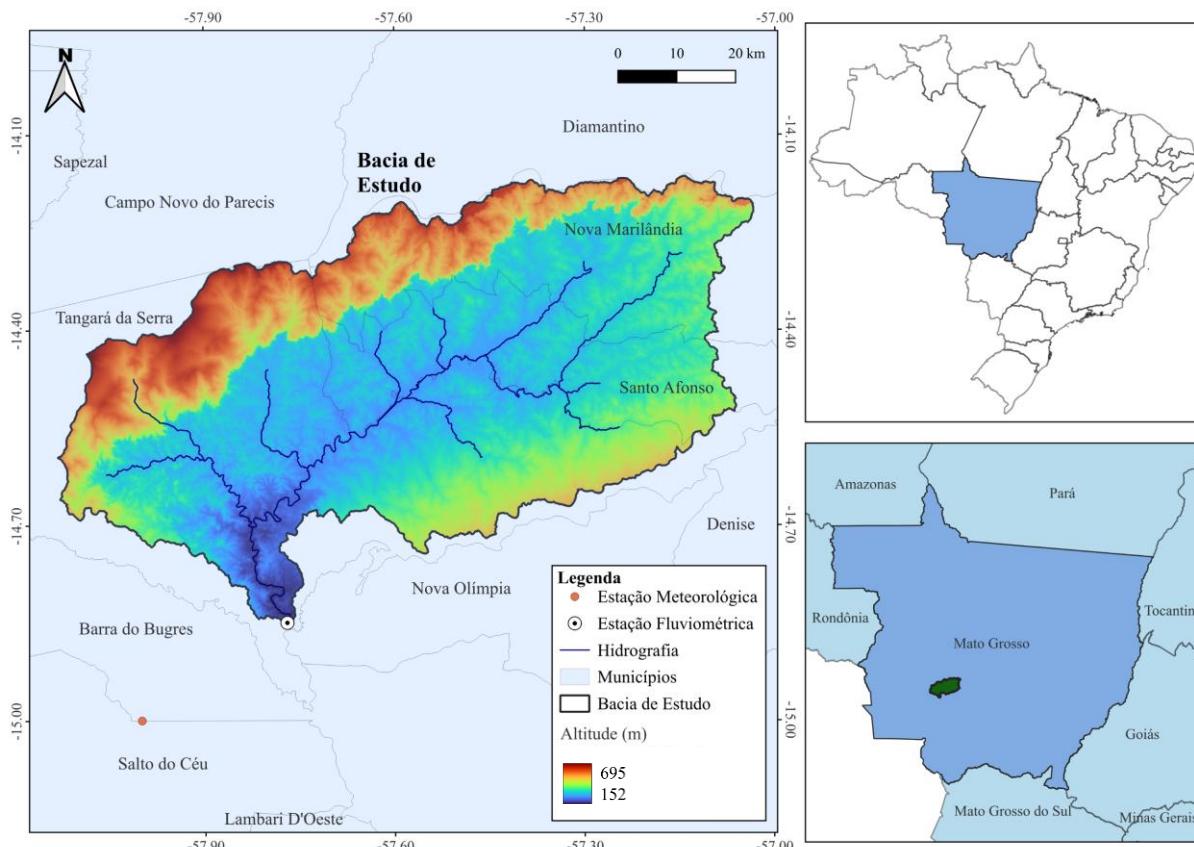
Fonte: Autores.

## Área de estudo

O estudo foi realizado na Bacia Hidrográfica do Rio Sepotuba, tendo como exutório a estação fluviométrica da Serra do Tapirapuã, localizada na latitude de  $-14,85^{\circ}$ , longitude de  $-57,77^{\circ}$  e altitude de 212 m, no município de Tangará da Serra – MT. A área drenagem é de aproximadamente 5.290 km<sup>2</sup>, o que representa cerca de 54% da área total da bacia do Rio Sepotuba, contendo parte significativa da bacia nas cidades de Tangará da Serra, Santo Afonso, Nova Marilândia, Campo Novo do Parecis e Diamantino, como pode ser observado na Figura 2.

A bacia hidrográfica estudada integra a Bacia Hidrográfica do Rio Sepotuba, importante afluente do Rio Paraguai, pertencente à Bacia Platina. Sua área de drenagem é de aproximadamente 9.840 km<sup>2</sup>, o que representa 1% da área do estado, sendo composta pelos biomas cerrado, floresta amazônica e pantanal (Gamero *et al.*, 2023). Os usos da água na bacia compreendem a dessedentação animal, agricultura e abastecimento urbano de 10 municípios, além da geração de energia elétrica em pequenas e médias hidroelétricas, principalmente ao longo do Rio Juba (Leandro e Rocha, 2019).

Figura 2 – Mapa da Área de Estudo da Bacia do Rio Sepotuba.



Fonte: Autores (2025).

## Aquisição dos dados

Os dados utilizados para modelagem da série histórica de vazões e precipitações foram obtidos da Agência Nacional de Águas e Saneamento Básico (ANA) por meio de sua plataforma, Hidroweb e os dados de temperatura foram obtidos do Instituto Nacional de Pesquisas Espaciais (INPE). As informações das estações estão dispostas na Tabela 1.

Tabela 1 – Informações da Estações utilizadas.

Código	Nome	Tipo de Estação	Coordenadas Geográficas		Altitude	Início
			Latitude	Longitude		
1457000	Tapirapuã	Pluviométrica	-14,85	-57,77	212,00	06/1971
66050000	Tapirapuã	Fluviométrica	-14,85	-57,77	212,00	06/1971
A936	Salto do Céu	Meteorológica	-15,12	-58,13	300,83	01/2008

Fonte: ANA e INPE (2025)

Realizou-se então o pré-processamento dos dados, por meio da identificação de erros, como discordância de valores ou falhas para a devida correção, caso houvesse necessidade. Deste modo, definiu-se o período temporal para as séries compreendendo o intervalo de janeiro de 2008 a dezembro de 2019, totalizando 144 observações. Buscou-se também avaliar o desempenho dos modelos propostos para distintos modelos, a saber o SARIMA apenas com os dados da vazão, e os modelos SARIMAX, incluindo a precipitação e a temperatura como variáveis regressoras, e ambas totalizando 4 modelos. Para a criação dos modelos, os dados foram divididos em dois períodos, a etapa de treino abrangendo 2008 a 2018 e o tempo de teste para o ano de 2019.

### Modelo de previsão SARIMA e SARIMAX

Depois da obtenção dos dados, utilizou-se a metodologia de Box e Jenkins (1976) para a criação dos modelos SARIMA e SARIMAX, sabendo que nesta metodologia existem quatro etapas, a saber: identificação do modelo; estimativa dos parâmetros; verificação e avaliação do modelo selecionado e previsão.

Em conformidade a Bayer e Souza (2010) o modelo AR, existe quando o processo estocástico possui somente a parte autorregressiva, ou modelos MA, quando existe interação somente de médias móveis. Quando os modelos apresentam ambas as componentes, são denominados de modelos ARMA. Em séries não estacionárias, após o processo de diferenciação, o modelo passa a possuir a parte integrada I, portanto, denominado ARIMA. Quando se adiciona a sazonalidade o modelo da classe ARIMA podem ser estendidos para modelos da classe SARIMA.

Em termos matemáticos o modelo ARIMA ( $p, d, q$ ) é composto apenas por uma variável autorregressiva ( $p$ ), uma variável integradora ( $d$ ) e uma de média móvel ( $q$ ). Já para o modelo SARIMA ( $p, d, q$ ) ( $P, D, Q$ ), sendo o mesmo determinado pela Equação 1. Um modelo SARIMA ( $p, d, q$ ) ( $P, D, Q$ ) possui as ordens de autorregressão ( $p, P$ ), as ordens de diferenciação ( $d, D$ ) e as ordens de média móvel ( $q, Q$ ), sendo as variáveis ( $p, d, q$ ) não sazonais, e ( $P, D, Q$ ) sazonais.

$$\phi(B)\Phi(B^S)(1 - B^S)^D(1 - B)^d Z_t = \theta(B)\Theta(B^S)\varepsilon_t \quad (1)$$

Na Equação 1, ( $p, d, q$ ) são ordens do modelo referentes à dinâmica ordinal, ( $P, D, Q$ ) são ordens do modelo sazonal,  $s$  é a periodicidade sazonal,  $\phi(B)$  é o polinômio autorregressivo,  $\theta(B)$  é o polinômio de médias móveis,  $\Phi(B^S)$  é o parâmetro autorregressivo sazonal,  $\Theta(B^S)$  é o parâmetro de médias móveis sazonais, sendo determinados pelas Equações 2 e 3.

$$\Phi(B^S) = 1 - \Phi_1 B^S - \Phi_2 B^{S2} - \dots - \Phi_p B^{Sp} \quad (2)$$

$$\Theta(B^S) = 1 - \Theta_1 B^S - \Theta_2 B^{S2} - \dots - \Theta_p B^{Sp} \quad (3)$$

O modelo SARIMAX compreende a aplicação de uma variável regressora (X) ao modelo, com a finalidade de prever com valores futuros da série temporal selecionada utilizando relações lineares dos valores prévios observados dos dados sequenciais, informações secundárias fornecidas pelas variáveis exógenas e termos de erro (Fazla et al., 2023). A sua equação típica só de diferencia da Equação 1, pelo acréscimo da variável regressora, representado pelo termo somatório do lado direito que modela a relação entre a sequência de dados observados e o vetor das informações explanatórias como descrito pela Equação 4.

$$\phi(B)\Phi(B^S)(1-B^S)^D(1-B)^d Z_t = \theta(B)\theta(B^S)\varepsilon_t + \sum_{i=1}^m \beta_{iS}t_i \quad (4)$$

Todo o processo para obtenção dos modelos foi realizado em R.

### Métricas para avaliação dos modelos

A fim de escolher os modelos preditivos de vazões mensais, aplicou-se o comando `auto.arima()`, função pertencente ao pacote `forecast`. Nesse comando a seleção do modelo é automatizada, com o objetivo de identificar o modelo ARIMA mais adequado, combinando testes de raiz unitária, minimização do Critério de Informação de Akaike (AIC) e estimação por máxima verossimilhança (MLE) (Hyndman e Athanasopoulos, 2021).

Após a seleção do modelo, é necessário passar por uma avaliação dos dados simulados. Neste caso, foram utilizadas as métricas do coeficiente de Nash-Sutcliffe (NSE), o coeficiente de Kling-Gupta (KGE) e a raiz quadrada do erro médio (RMSE), como estão descritos nas Equações 5, 6 e 7, respectivamente.

$$NSE = 1 - \left[ \frac{\sum_{i=1}^n (Q_{i,obs} - Q_{i,sim})^2}{\sum_{i=1}^n (Q_{i,obs} - \bar{Q}_{obs})^2} \right] \quad (5)$$

Em que:  $n$  é o número de observações,  $Q_{i,obs}$  é a vazão observada no mês,  $Q_{i,sim}$  é a vazão simulada e  $\bar{Q}_{obs}$  é a vazão média observada no período, todas as vazões em  $m^3.s^{-1}$ , de modo que NSE seja um parâmetro adimensional.

$$KGE = 1 - \sqrt{(r-1)^2 + \left( \frac{\sigma_{sim}}{\sigma_{obs}} - 1 \right)^2 + \left( \frac{\mu_{sim}}{\mu_{obs}} - 1 \right)^2} \quad (6)$$

Em que:  $r$  é a correlação linear entre observações e simulações,  $\sigma_{obs}$  é o desvio padrão nas observações,  $\sigma_{sim}$  o desvio padrão nas simulações,  $\mu_{sim}$  a média da simulação e  $\mu_{obs}$  a média da observação.

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (Q_{i,obs} - Q_{i,sim})^2} \quad (7)$$

Nesta equação utilizou-se a mesma notação que a Equação 5, o que muda é a disposição e operação das variáveis.

Para o NSE o valor pode variar de 1 a  $-\infty$ , sendo resultado abaixo de 0 ruim e 1 o melhor resultado possível. De modo semelhante, a métrica KGE se encontra entre 1 e  $-\infty$ . Porém, valores menores que -0,41 são classificados como ruins e valores próximos à unidade, corresponderiam a de melhor desempenho. Consoante Knoben et al. (2019), ao fazer a relação entre o NSE e o KGE, valores satisfatórios para as métricas seriam  $NSE \geq 0,5$  e  $KGE \geq 0,3$ . Em relação ao RMSE, quanto menor valor, melhor será o resultado.

## RESULTADOS

## Descrição dos dados

Com finalidade de compreender o comportamento dos dados utilizados, apresenta-se a Tabela 2, contendo as medidas descritivas da vazão e precipitação média mensal, e temperatura mensal observadas na estação Tapirapuã e Salto do Céu, sub-bacia do Rio Sepotuba, compreendendo o período de janeiro de 2008 a dezembro de 2020.

A vazão mínima observada ocorreu em setembro de 2019 e a máxima vazão registrada aconteceu em março de 2011. A vazão apresentou uma média coerente, o que pode ser constatado pelo coeficiente de variação menor que 50%. Para a precipitação, o valor máximo ocorreu no mês de dezembro de 2009, com as mínimas ocorrendo de maio a setembro, principalmente em junho e julho. A média das precipitações foi pouco significativa, principalmente por ter um coeficiente de variação próximo a 95%. Já para a temperatura, a mínima ocorreu em julho de 2012 e a máxima em setembro de 2019, com uma média muito representativa, já que o coeficiente de variação foi inferior a 10%.

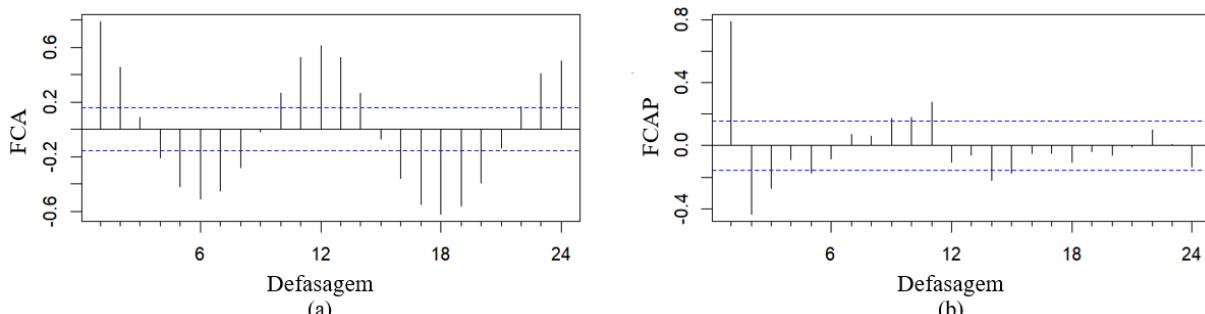
Tabela 2 – Medidas descritivas da Vazão, Precipitação e Temperatura.

Medidas Descritivas	Vazão ( $m^3.s^{-1}$ )	Precipitação (mm)	Temperatura (°C)
Mediana	121,94	115,22	25,10
Média	142,78	146,77	24,56
Mínima	69,58	0,00	20,00
Máximo	346,81	598,60	27,90
Desvio Padrão	58,54	137,82	1,66
Coeficiente de Variação	41,01	93,90	6,78
Assimetria	1,09	0,85	-0,72

Fonte: Autores (2025).

Na Figura 3 apresentam-se as funções de autocorrelação amostral (FAC) e de autocorrelação parcial amostral (FACP) da série temporal de vazões médias sem diferenciação, ou seja, dos dados reais. Percebe-se a presença de autocorrelação com significância diferente de zero em defasagens múltiplas de 12, confirmando a sazonalidade. Logo, a presença do componente sazonal, bem como o comportamento senoidal e infinito no correlograma da ACF, denota um processo do tipo autorregressivo, não estacionário e sazonal.

Figura 3 – Funções de Autocorrelação.



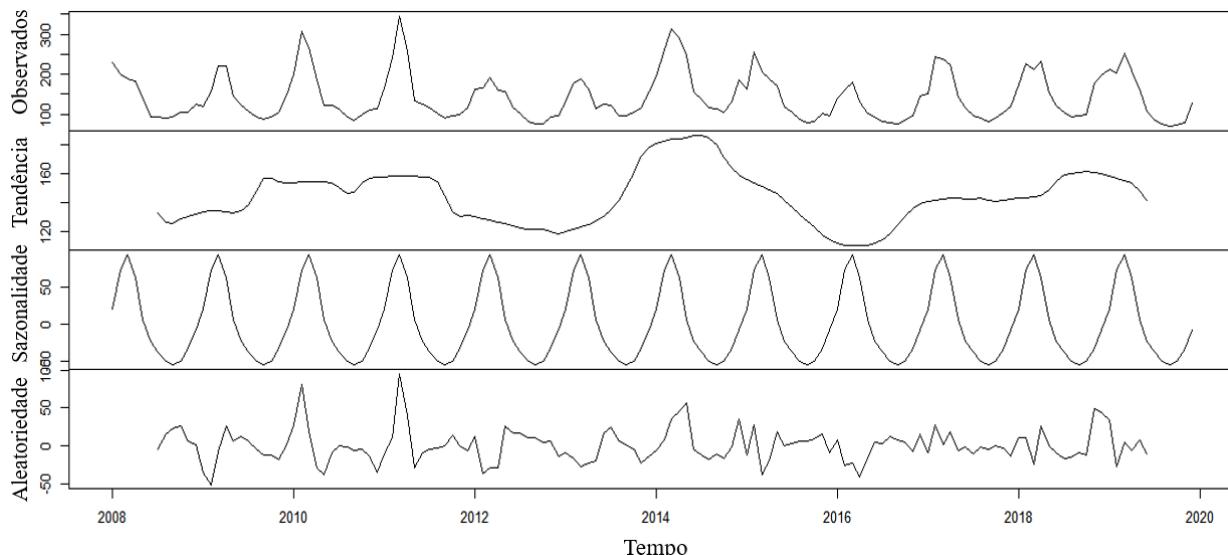
Legenda: a – Função de autocorrelação amostral (FAC); b – Função autocorrelação parcial amostral (FACP).

Fonte: Autores (2025).

A Figura 4 apresenta a série de vazões médias mensais decomposta em componentes de valores observados, tendência, sazonalidade e aleatoriedade (resíduos). Percebe-se na componente de sazonalidade que as séries variam periodicamente em torno de um valor médio, confirmando a

presença dela nas séries analisadas. Na componente tendência verificou-se uma repetição de períodos de estiagens a cada quatro anos, com intensificação no período de 2015 a 2017.

Figura 4 – Decomposição da série de vazão.



Fonte: Autores (2025)

### Avaliação dos modelos desenvolvidos

A tabela 3 apresenta um resumo dos modelos obtidos, considerando apenas a vazão no modelo SARIMA e as variáveis regressoras precipitação e temperatura nos modelos SARIMAX.

Tabela 3 – Informações dos Modelos, Valores do AIC, Testes Estatísticos

Nome	Variável Regressora	Modelo	AIC	Shapiro-Wilk (p-value)	Ljung-Box (p-value)
SARIMA	-	SARIMA(1,0,0)(3,1,0) <sub>[12]</sub>	1154,12	6,39E-07	0,935
SARIMAX 1	Precipitação	SARIMAX(0,0,1)(2,1,2) <sub>[12]</sub>	1160,33	1,59E-06	0,004
SARIMAX 2	Temperatura	SARIMAX(2,0,0)(2,1,1) <sub>[12]</sub>	1152,78	1,14E-06	0,940
SARIMAX 3	Precipitação + Temperatura	SARIMAX(0,0,1)(2,1,2) <sub>[12]</sub>	1162,32	1,46E-06	0,004

Como pode ser observado, os melhores modelos foram: SARIMA de ordens (1,0,0)(3,1,0), para SARIMAX 1 e SARIMAX 3 de ordem (0,0,1)(2,1,2), e para o SARIMAX 2 ordem (2,0,0)(2,1,1).

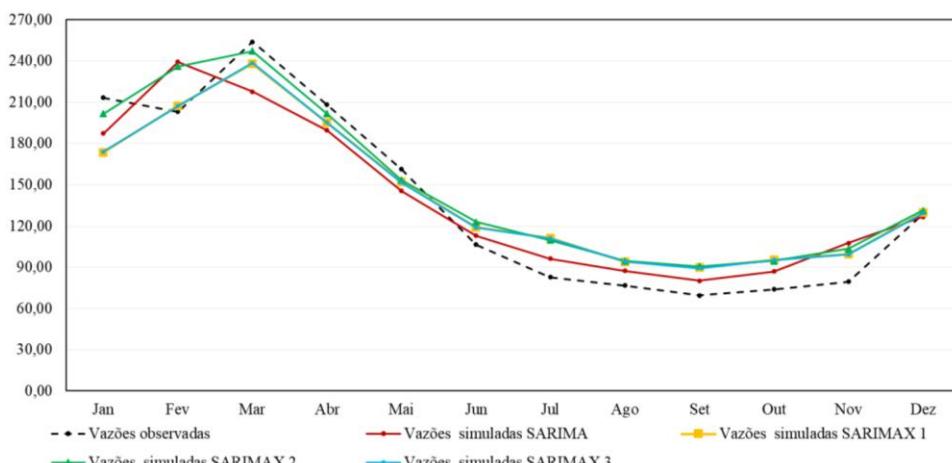
Os valores do critério de AIC tendem a diminuir com a simplicidade do modelo, o que demonstra que o modelo SARIMAX 2 é ligeiramente melhor que os modelos para AIC. Vale ressaltar que os critérios de AIC foram utilizados apenas para a identificação do modelo pelo `auto.arima()`.

Embora os resíduos dos modelos não tenham seguido uma distribuição normal segundo o teste de Shapiro-Wilk ( $p < 0,05$ ), os testes de Ljung-Box ( $p < 0,05$ ) indicaram ausência de autocorrelação para os modelos SARIMAX 1 e 2, sugerindo que os resíduos se comportam como ruído branco.

### Avaliação de desempenho

A Figura 5 apresenta as previsões para o ano de 2019, considerando os melhores modelos propostos.

Figura 5 – Gráfico das vazões médias mensais observadas e previstas.



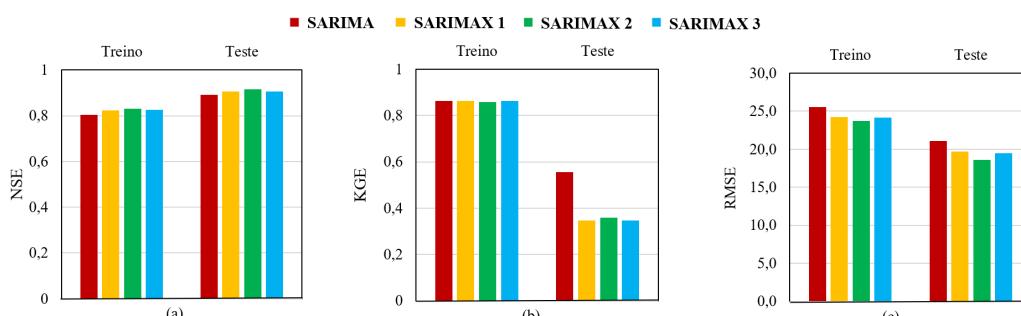
Fonte: Autores (2025).

Na Figura 5, pode-se perceber a similaridade entre as previsões dos modelos SARIMAX 1 e SARIMAX 3, já que agregaram ambos a precipitação como variável exógena, tendo essa mais influência na predição na predição. Para os modelos SARIMA e SARIMAX 2, houve superestimativa de vazões para o mês de fevereiro, assim como no período de junho a novembro, período de vazões mínimas observadas. Nos meses de janeiro e março, todos os modelos subestimaram as vazões de pico, apenas prevendo com pouco erro o mês de dezembro.

Diante disso, é importante salientar que tais previsões são satisfatórias nos períodos de treino e teste, já que estão em consonância aos verificados na literatura. Como pode-se perceber, os modelos SARIMA e SARIMAX tendem a subestimar vazões superiores, ou, ainda subestimar vazões mínimas, principalmente em períodos de estiagem, ou seja, eventos extremos são imprevisíveis para estes modelos (Bayer *et al.*, 2012; Costa *et al.*, 2023 e Caminha *et al.*, 2025).

A Figura 6 apresenta a comparação entre as métricas de desempenho para os períodos de treino e teste considerando os modelos desenvolvidos.

Figura 6 – Gráficos das métricas de avaliação dos modelos.



Legenda: a – Coeficiente de Nash-Sutcliffe (NSE); b – Coeficiente de Kling-Gupta (KGE); c – Raiz do erro quadrático médio (RMSE)

Como se observa, todos os modelos no período de treino apresentaram resultados muito semelhantes como pode ser percebido com o KGE de 0,86. Mesmo assim, é possível ver a ligeira superioridade do modelo SARIMAX 2, que obteve NSE de 0,83 e RMSE de  $23,74 \text{ m}^3 \cdot \text{s}^{-1}$ , no período de teste. Já para o período de teste os modelos demonstraram boa adequação dos dados, por não haver

overfitting (sobreajuste). O que comprova isso é o aumento dos resultados das métricas para todos os modelos do NSE e RMSE, com destaque para o SARIMAX 2 com as melhores previsões (NSE = 0,92; RMSE de  $18,58 \text{ m}^3 \cdot \text{s}^{-1}$ ).

Como o KGE é uma medida que possui características e parâmetros diferentes das outras métricas, houve uma queda na avaliação desse parâmetro para os modelos. O melhor resultado foi para o modelo SARIMA com o valor de 0,56. Enquanto para o modelo SARIMAX 2 foi de 0,36, sendo um valor ainda válido, como apresentado por Knoben (2019), mesmo sendo um valor menor que o desempenho do SARIMA.

Portanto, o modelo que apresentou o melhor desempenho foi o SARIMAX 2, sendo aquele que possuiu como variável exógena a temperatura. Todavia, isso não inviabiliza a validade dos outros modelos, já que para todos os testes, as métricas resultaram em valores próximos ao modelo SARIMAX 2, de modo que o emprego de variáveis regressoras tendem a aprimorar as previsões realizadas pelos modelos da família ARIMA.

## CONCLUSÃO

Este trabalho teve por objetivo o desenvolvimento e análise de modelos estocásticos da família ARIMA para a previsão de vazões médias mensais de uma bacia do Rio Sepotuba, na cidade Tangará da Serra, Mato Grosso.

O modelo de melhor desempenho, SARIMAX 2, considerou o emprego da temperatura como variável regressora. Tal resultado vai de encontro a literatura, em relação à melhoria do desempenho de modelos SARIMA quando variáveis regressoras são implementadas.

Do mesmo modo, todos os modelos obtidos tiveram excelente desempenho verificado nas métricas obtidas, indicando a superioridade dos modelos SARIMAX em relação ao SARIMA para a bacia estudada.

## REFERÊNCIAS

- ALMEIDA, L; SERRA, J.C.V. (2017). “*Modelos hidrológicos, tipos e aplicações mais utilizadas.*”. Revista da FAE, 20 (1), pp. 129-137.
- BAYER, D. M.; CASTRO, N.; BAYER, F. M. (2012). “*Modelagem e Previsão de Vazões Médias Mensais do Rio Potiribu Utilizando Modelos de Séries Temporais*”. Revista Brasileira de Biometria, v. 17, pp. 229–239.
- BAYER, F.M.; SOUZA, A.M. (2010). “*Wavelets e modelos tradicionais de previsão: um estudo comparativo*”. Revista Brasileira de Biometria, v.28, n. 2, pp. 40-61.
- BOX, G.E.P.; JENKINS, G.M. (1976). *Time Series Analysis, Forecasting and Control*; Holden-Day San Francisco, CA, USA, 553 p.
- CAMINHA, A.R.; GUIMARÃES, L.R.; CORRÊA, F. V.; SOUZA, G.R.; SÁFADI, T. (2025). “*Previsão de precipitação e vazão a partir da modelagem de séries temporais*”. Revista Brasileira de Climatologia. 36 (21), pp. 264-290.
- DEUS, R.; DEUS, S.S.; NEVES, R. J.J. (2020). “*Uso do código SWAT na previsão de vazão dos rios da bacia hidrográfica Tocantins-Araguaia.*”. Novos Cadernos NAEA. 23. n.2, pp. 179-200.
- FAZLA, A.; AYDIN, M.E.; KOZAT, S.S. (2023). “*Joint Optimization of Linear and Nonlinear Models for Sequential Regression.*” Digit. Signal Process. 132, 103802

GAMERO, A. R.; NEVES, S.M.A.S.; GALVANIN, E.A.S.; GONZÁLES, A.Z.D.; LEANDRO, G.R.S. (2023). “*Vulnerabilidade natural e ambiental da paisagem da Bacia Hidrográfica do Rio Sepotuba, Mato Grosso - Brasil*”. Revista Caminhos de Geografia. 24 (91), pp 258-271.

HYNDMAN, R.J.; ATHANASOPOULOS, G. (2021). *Forecasting: Principles and Practice*; OTexts: Melbourne, VIC, Australia, 442 p.

JUNIOR, A.D.M.; SILVEIRA, C.S.; COSTA, J.M.F.; GONÇALVES, S.T.N. (2023). “*Previsões de vazões afluentes com modelos de machine learning aplicados ao reservatório de três Marias*”. in Anais do XXV Simpósio Brasileiro de Recursos Hídricos. Aracajú, Nov. 2023, 1. 10p. (ver)

KNOBEN, W.J.M.; FREER, J.E.; WOODS, R.A. (2019) “*Inherent benchmark or not? Comparing Nash–Sutcliffe and Kling–Gupta efficiency scores*”. Hydrology and Earth System Sciences, v. 23, n. 10, p. 4323-4331.

LEANDRO, G.R.S.; ROCHA, P.C. (2019). “*Expansão agropecuária e degradação ambiental na bacia hidrográfica do rio Sepotuba - Alto Paraguai, Mato Grosso- Brasil*”. Revista Sociedade e Natureza 31. 22 p.

SILVA, G.J.O.; NEVES, S.M.A.S.; RAMOS, A.W.P.; GALVANIN, E.A.S. (2021). “*Dinâmica da cobertura vegetal e uso da terra da bacia hidrográfica do rio Sepotuba-MT, Brasil*”. Caminhos de Geografia, Uberlândia, v. 22, n. 80, pp. 81–96.

UNESCO - ORGANIZAÇÃO DAS NAÇÕES UNIDAS PARA A EDUCAÇÃO, CIÊNCIA E CULTURA. “*Relatório Mundial das Nações Unidas sobre o Desenvolvimento dos Recursos Hídricos 2024: Água para a prosperidade e a paz; fatos, dados e exemplos de ações*””. (2024). Perúgia, Itália. 19p.