



Desempenho de modelos SARIMA e SARIMAX para previsão de vazões médias mensais

José Rodrigues da SILVA NETO ¹, Frederico Carlos Martins de MENEZES FILHO ^{*1}

¹ Instituto de Ciências Exatas e Tecnológicas, Universidade Federal de Viçosa, Rio Paranaíba, MG, Brasil.

*E-mail: frederico.menezes@ufv.br

Submetido em: 14/02/2026; Aceito em: 03/08/2026; Publicado em: 13/08/2026.

RESUMO: A modelagem estocástica de séries temporais para a previsão de vazões tem se consolidado como uma alternativa relevante para o planejamento e a gestão de recursos hídricos em bacias hidrográficas não monitoradas. Este estudo teve como objetivo avaliar o desempenho dos modelos SARIMA e SARIMAX, considerando diferentes períodos de treino e de teste, bem como a influência da inclusão de variáveis exógenas, na previsão das vazões médias mensais de uma sub-bacia do rio São Francisco. A metodologia adotada seguiu a abordagem de Box e Jenkins (1976). Foram selecionados os modelos SARIMA (1,0,0)(4,0,0)_[12] e SARIMAX (1,0,0)(2,0,0)_[12], sendo este último incluindo a precipitação como variável exógena. As previsões foram realizadas para quatro horizontes (3, 6, 9 e 12 meses) em 2022. O desempenho foi avaliado por meio das métricas NSE, KGE e PBIAS. Em todos os critérios, o modelo SARIMAX superou o SARIMA. Verificou-se que a escolha dos intervalos de treino e de teste não foi determinante para a melhora do desempenho, sendo os melhores resultados obtidos com o maior período de treino (98% da série). O SARIMAX apresentou NSE = 0,65, KGE = 0,68 e PBIAS = 0,1% no treino, e NSE = 0,45, KGE = 0,5 e PBIAS = 20,9% no teste.

Palavras-chave: modelagem hidrológica; modelos estocásticos; séries temporais.

Performance of SARIMA and SARIMAX models for monthly mean streamflow forecasting

ABSTRACT: Stochastic time series modeling for streamflow forecasting has become an important approach for water resources planning and management in ungauged river basins. This study aimed to evaluate the performance of SARIMA and SARIMAX models under different training and testing periods, as well as the influence of including exogenous variables, in forecasting monthly mean streamflow in a sub-basin of the São Francisco River. The methodology followed the Box and Jenkins (1976) approach. The selected models were SARIMA (1,0,0)(4,0,0)_[12] and SARIMAX (1,0,0)(2,0,0)_[12], the latter including precipitation as an exogenous variable. Forecasts were generated for four forecasting horizons (3, 6, 9, and 12 months) during 2022. Model performance was evaluated using the Nash–Sutcliffe Efficiency (NSE), Kling–Gupta Efficiency (KGE), and Percent Bias (PBIAS) metrics. The SARIMAX model outperformed the SARIMA model across all evaluation criteria. The results indicated that the choice of training and testing intervals was not a determining factor in improving model performance, with the best results obtained using the longest training period (98% of the time series). During the training period, the SARIMAX model achieved an NSE of 0.65, KGE of 0.68, and PBIAS of 0.1%, whereas during the testing period, it yielded an NSE of 0.45, KGE of 0.50, and PBIAS of 20.9%.

Keywords: hydrological modeling; stochastic models; time series.

1. INTRODUÇÃO

O aprimoramento da gestão dos recursos hídricos e das estratégias de previsão torna-se imprescindível diante da escassez hídrica, agravada pelo crescimento econômico e pelo aumento da população. Para contemplar a acessibilidade aos recursos hídricos faz-se necessário uma gestão eficiente, assim os setores como: agricultura, abastecimento de água, geração de energia e transporte hidroviário serão plenamente satisfeitos (SHAH et al., 2023).

Os avanços tecnológicos e matemáticos, impulsionados pela acessibilidade a linguagens de programação e a bancos de dados, têm se mostrado uma estratégia importante para a

tomada de decisão. Portanto, é necessária a utilização de métodos para prever a vazão, como a modelagem hidrológica, que abrange modelos físicos, estocásticos e de inteligência artificial. Os modelos físicos são complexos, pois, além da série de vazões, é indispensável uma grande quantidade de dados e o estabelecimento de parâmetros referentes à pedologia, ao uso e ocupação do solo e à cobertura vegetal da bacia hidrográfica. Já os modelos estocásticos e de aprendizado de máquina exigem menor volume de dados de entrada do que os modelos físicos. Destacam-se, dentre eles, respectivamente, os modelos da família ARIMA, utilizados mundialmente para a previsão de

vazão (BARRIENTOS-TORRES et al., 2023), e as redes neurais artificiais (SUN et al., 2019; KASSEM et al., 2020).

Estudos utilizando modelo SARIMA foram conduzidos por Bayer et al. (2012) para a previsão de vazões mensais do Rio Potiribu (RS), considerando a relação entre períodos de treino e de teste de 93% e 7%, ou 93/7. Bleidorn et al. (2019), com aplicação no Rio Jucu, no Espírito Santo, consideraram a relação entre os períodos de treino e de teste de 98/2.

Costa et al. (2023) avaliaram a aplicação do modelo SARIMAX para a previsão de vazões, incluindo a precipitação como variável exógena, considerando a relação entre os períodos de treino e de teste de 89/11. O melhor modelo obtido foi o SARIMAX(2,0,0)(2,1,0)_[12], com horizonte de 9 meses e NSE próximo de 0,81.

Consoante Hyndman; Athanasopoulos (2021), a dimensão do conjunto de teste deve ser de cerca de 20% da amostra total, embora este valor dependa do tamanho da amostra e de quanto antes se deseja prever. Em geral, na modelagem de dados, as porcentagens utilizadas para o conjunto de teste são 70% e 80%, e a determinação da proporção ideal depende de fatores como o tamanho do conjunto de dados, sua complexidade e do modelo (SIVAKUMAR et al., 2024).

Como descrito, não há padronização quanto à divisão da série temporal em períodos de treino e de teste para modelos estocásticos, que são geralmente definidos em função da extensão da série temporal disponível ou de períodos sem falhas, por exemplo. Além da limitação regional à modelagem estocástica, em que cada bacia terá um modelo diferente, o que dificulta a generalização desses modelos. Portanto, tal padronização traria maior rigor metodológico e confiabilidade, o que melhoraria a aplicação dos modelos hidrológicos na gestão dos recursos hídricos.

Dessa forma, o presente artigo teve como objetivos: (i) comparar o desempenho preditivo dos modelos SARIMA e SARIMAX, avaliando a influência da inclusão da precipitação como variável exógena; (ii) analisar o efeito de diferentes relações treino-teste sobre a qualidade das previsões em horizontes de 3,6,9 e 12 meses; e (iii) discutir as implicações hidrológicas dos resultados, destacando a capacidade dos modelos em capturar a sazonalidade e os eventos extremos da série histórica de vazões médias mensais do Rio Indaiá, afluente do São Francisco.

2. MATERIAL E MÉTODOS

2.1. Área de estudo

A área de estudo, com aproximadamente 875,31 km² está inserida na Bacia Hidrográfica do São Francisco (Figura 2). Tal sub-bacia pertence à circunscrição hidrográfica (CH) do Entorno da Represa de Três Marias (SF4), com seção de controle na estação fluviométrica Barra do Funchal, sub-bacia do Rio Indaiá, em Minas Gerais.

2.2. Obtenção dos dados

Os dados utilizados para a modelagem da série histórica de vazões foram obtidos da Agência Nacional de Águas e Saneamento Básico (ANA) por meio de sua plataforma Hidroweb.

As estações hidrometeorológicas de Barra do Funchal incluem uma estação fluviométrica (código 40930000), em operação desde 1939, a 679 m de altitude, e uma estação pluviométrica (código 1945002), ativa desde 1941, a 720 m de altitude. Ambas localizadas nas coordenadas 19,39° S e 45,88° W (ANA, 2024).

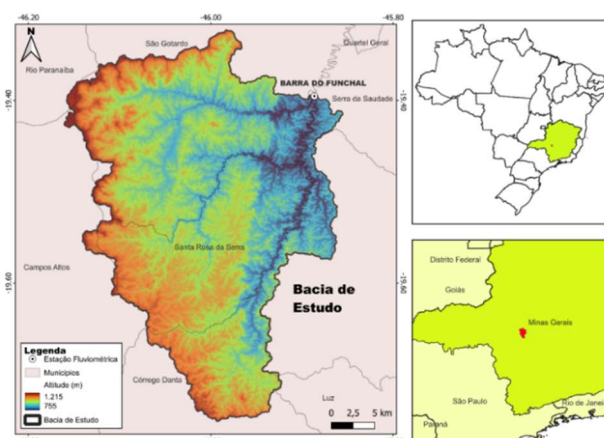


Figura 1. Localização da sub-bacia hidrográfica do rio Indaiá em Minas Gerais.

Figure 1. Location of the Indaiá River sub-basin in Minas Gerais State, Brazil.

De posse dos dados de vazão e de precipitação, realizou-se o pré-processamento por meio da identificação de erros tais como discordâncias de valores ou ausências para a devida correção, se necessária. Os valores ausentes foram preenchidos com a média histórica do mês e selecionado o período de maior abrangência e de menor percentual de falhas. A utilização da precipitação se deu pela facilidade de acesso aos dados, disponíveis no próprio site da ANA.

Desta maneira, definiu-se a série temporal de estudo, compreendendo o período de janeiro de 1970 a dezembro de 2022, totalizando 636 observações. A estratégia de validação temporal adotada neste estudo compreendeu duas etapas distintas. Na primeira etapa, para cada configuração de treino e teste apresentada na Tabela 1, os modelos SARIMA/SARIMAX foram estimados uma única vez, utilizando exclusivamente os dados de treino, mantendo seus parâmetros fixos durante todo o processo de previsão.

Tabela 1. Informações sobre a divisão dos períodos de treino e de teste.

Table 1. Details of the data partitioning into training and test periods.

Intervalo de treino	Intervalo de teste	Treino	Teste
1970 – 2021	2022	98%	2%
1970 – 2017	2018 – 2022	90%	10%
1970 – 2011	2012 – 2022	80%	20%
1970 – 2006	2007 – 2022	70%	30%
1970 – 2001	2002 – 2022	60%	40%
1970 – 1996	1997 – 2022	50%	50%

Porcentagens aproximadas.

Approximate percentages.

Em seguida, geraram-se previsões para todo o intervalo de teste e calcularam-se os erros correspondentes, sem utilizar horizontes específicos. Na segunda etapa, com base no desempenho geral obtido (avaliando as métricas descritas em 2.4), selecionaram-se os melhores modelos e realizou-se uma análise complementar, focada exclusivamente no ano de 2022, na qual foram aplicados horizontes de 3, 6, 9 e 12 meses para avaliar o desempenho do modelo em diferentes defasagens nesse ano. O ano de 2022 não participou do período de treino na seleção dos modelos, sendo utilizado apenas na avaliação final, o que garantiu a inexistência de vazamento de informação (data leakage) entre os conjuntos de treino e de teste. Não foram empregadas janelas

deslizantes (rolling windows) nem reestimação do modelo ao longo do período de teste. Destaca-se também que a variável precipitação, utilizada como variável exógena nos modelos SARIMAX, foi agregada mensalmente às vazões, sem a aplicação de defasagens temporais.

2.3. Modelos de previsão SARIMA e SARIMAX

Os modelos Autorregressivos Integrados de Médias Móveis (ARIMA), introduzidos por Box e Jenkins na década de 1970, são modelos estatísticos lineares amplamente utilizados na análise de séries temporais, com o objetivo de capturar e representar as características de processos estocásticos (BAYER; SOUZA, 2010).

Em virtude da sazonalidade presente nos processos hidrológicos, os modelos SARIMA (Sazonal Autorregressivo Integrado de Médias Móveis) constituem alternativas eficientes para a captura de padrões periódicos inerentes às variáveis ambientais (MENEZES FILHO, 2020; COSTA et al. 2023; CAMINHA et al., 2025).

Os modelos SARIMA (p, d, q) (P, D, Q) têm sua estrutura representada pela Equação 1 e caracterizam-se pelas ordens de autorregressão (p, P), diferenciação (d, D) e média móvel (q, Q).

$$\phi(B) \Phi(B^S) (1 - B^S)^D (1 - B)^d Z_t = \theta(B) \Theta(B^S) \varepsilon_t \quad (01)$$

em que: $\phi(B)$ = operador autorregressivo (AR) não sazonal de ordem p; $\Phi(B^S)$ = operador autorregressivo (AR) sazonal de ordem P; $(1 - B^S)^D$ = operador de diferenciação sazonal (S = sazonalidade, D = ordem); $(1 - B)^d$ = operador de diferenciação não sazonal (d = ordem); Z_t = série temporal observada; $\theta(B)$ = operador de média móvel (MA) não sazonal de ordem q; $\Theta(B^S)$ = operador de média móvel (MA) sazonal de ordem Q; ε_t = termo de erro de ruído branco; B = operador de defasagem (backshift), por exemplo, $B y_t = y_{t-1}$.

O modelo SARIMAX, versão mais avançada do modelo ARIMA (ALHARBI; CSALA, 2022), é uma abordagem estatística cujo principal propósito é prever valores futuros de uma série temporal usando relações lineares entre os valores observados previamente dos dados sequenciais, informações secundárias fornecidas pelas variáveis exógenas e termos de erro (FAZLA et al., 2023). Do mesmo modo que o modelo SARIMA descrito pela Equação 1, o modelo SARIMAX é frequentemente expresso pela ordem de seus parâmetros (p,d,q) (P, D, Q), diferindo pela inclusão de um termo exógeno ou variável regressora, representado pelo termo de soma no lado direito, que modela a relação entre a sequência de dados observados e o vetor de informações explanatórias, como descrito por (FAZLA et al., 2023). O modelo SARIMAX é representado pela Equação 2.

$$\phi(B) \Phi(B^S) (1 - B^S)^D (1 - B)^d Z_t = \theta(B) \Theta(B^S) \varepsilon_t + \sum_{i=1}^m \beta_{ist,i} \quad (02)$$

em que: $\sum_{i=1}^m \beta_{ist,i}$ = Componente de regressão (variáveis exógenas), em que β_i são os coeficientes associados a m regressores externos $s_{t,i}$.

O processo de modelagem estocástica, inicialmente apresentado por Box; Jenkins (1976), pode ser descrito por meio de quatro etapas: (i) identificação do modelo; (ii) estimativa dos parâmetros; (iii) verificação ou avaliação do modelo selecionado; e (iv) previsão (MACHIWAL; JHA, 2012).

Destaca-se que, em etapa anterior à identificação do modelo, verificou-se a estacionariedade das séries de vazões e de precipitação – esta última como variável exógena – por meio do teste de Dickey-Fuller (SAID; DICKEY, 1984). A identificação do modelo e a estimativa de seus parâmetros exigem a extração de padrões intrínsecos à série temporal (COSTA et al., 2023). A obtenção desses padrões dá-se pela decomposição da série temporal em componentes (tendência, sazonalidade e resíduo) e pelas funções de autocorrelação (FAC) e de autocorrelação parcial (PACF) (HYNDMAN; ATHANASOPOULOS, 2021).

Para a seleção dos modelos, além dos critérios mencionados, empregam-se também o critério de informação de Akaike (AIC) (AKAIKE, 1974) e o critério de informação bayesiano de Schwarz (BIC) (SCHWARZ, 1978), ambos amplamente consolidados na literatura (MACHIWAL; JHA, 2012). O modelo selecionado é aquele que minimiza simultaneamente esses dois critérios. As equações que os definem são apresentadas nas Equações 3 e 4.

$$AIC = 2(k) - 2\ln(L) \quad (03)$$

em que: L é a estimativa de máxima verossimilhança e k é o número de parâmetros do modelo.

$$BIC = (k)\ln(N) - 2\ln(L) \quad (04)$$

em que: L é a estimativa de máxima verossimilhança; k é o número de parâmetros do modelo e N o número de observações da série.

Conforme Costa et al. (2023), o AIC e o BIC atendem a propósitos distintos. O AIC seleciona o modelo que melhor representa o processo gerador dos dados, independentemente do tamanho da amostra; o BIC, por sua vez, busca o modelo mais parcimonioso, aplicando uma penalidade mais severa ao número de parâmetros em função do tamanho da amostra.

No presente estudo, a seleção dos modelos mais adequados para a previsão de vazões médias mensais foi realizada por meio do comando `auto.arima()` do pacote `forecast` (HYNDMAN; KHANDAKAR, 2008). Este comando automatiza a seleção do modelo combinando testes de raiz unitária, a minimização do AIC e da estimativa de máxima verossimilhança (MLE), conforme descrito por Hyndman e Athanasopoulos (2021). Destaca-se que o `auto.arima()` é amplamente utilizado por reduzir o tempo necessário para determinar as ordens ótimas dos modelos ARIMA, minimizando o AIC e o BIC (HYNDMAN; KHANDAKAR, 2008; LEE; CHO, 2022).

Após a definição dos melhores modelos e a estimativa de seus parâmetros, verifica-se, além do ajuste adequado aos dados, se os resíduos apresentam comportamento aleatório e não correlacionado, seguindo uma distribuição normal com média zero (COSTA et al., 2023). Segundo Bayer e Souza (2010), o teste de Ljung-Box é adequado para esse fim, pois identifica a presença de autocorrelação nos erros estimados por meio da autocorrelação dos resíduos.

Todas as análises de séries temporais envolvendo os modelos ARIMA foram realizadas no software livre R (R CORE TEAM, 2017), com o auxílio de pacotes específicos, dentre os quais o `forecast`.

2.4. Métricas de avaliação

No presente estudo, a análise de desempenho dos modelos foi realizada com as métricas: coeficiente de Nash-Sutcliffe (NSE), viés percentual (PBIAS), coeficiente de Kling-Gupta (KGE) e raiz do erro quadrático médio (RMSE).

O NSE tem sido amplamente utilizado na avaliação de modelos hidrológicos (Equação 5). No entanto, apresenta limitações quando aplicado a dados com forte sazonalidade ou na comparação entre bacias hidrográficas distintas (KNOBEN et al., 2019). O valor de NSE varia de 1 a $-\infty$, sendo que modelos de referência, como a média dos dados observados, resultam em NSE = 0.

$$NSE = 1 - \left[\frac{\sum_{i=1}^n (Q_{i,obs} - Q_{i,sim})^2}{\sum_{i=1}^n (Q_{i,obs} - \bar{Q}_{obs})^2} \right] \quad (05)$$

em que: n = número de observações; $Q_{i,obs}$ = vazão observada no mês (m^3/s); $Q_{i,sim}$ = vazão simulada (m^3/s); $\bar{Q}_{obs}$ = vazão média observada no período (m^3/s).

O PBIAS expressa a tendência média das vazões simuladas em relação às observadas e é calculado pela Equação 6. O valor ótimo é 0,0; valores positivos indicam superestimativa, enquanto valores negativos indicam subestimativa (GUPTA et al., 1999).

$$PBIAS = \left(\frac{\sum_{i=1}^n Q_{i,obs} - \sum_{i=1}^n Q_{i,sim}}{\sum_{i=1}^n Q_{i,obs}} \right) \cdot 100 \quad (06)$$

O KGE foi desenvolvido para avaliar modelos hidrológicos integrando correlação, viés e variabilidade entre os dados observados e simulados. O valor do KGE varia entre 1 e $-\infty$, semelhante ao NSE, mas, para modelos de referência, o valor esperado é -0,41 (DUC; SAWADA, 2023). Adotou-se como referência para o KGE valores entre 0 e 1 (desempenho aceitável a bom) e inferiores a -0,41 (desempenho ruim). O KGE é calculado pela Equação 7.

$$KGE = 1 - \sqrt{(r - 1)^2 + (\beta - 1)^2 + (\alpha - 1)^2} \quad (07)$$

em que: r = correlação de Pearson entre $Q_{i,sim}$ e $Q_{i,obs}$; β = razão entre $\bar{Q}_{sim}$ e $\bar{Q}_{obs}$ e α = razão entre $Q_{i,sim}$ e $Q_{i,obs}$.

O RMSE, calculado pela Equação 8, avalia a magnitude das diferenças entre valores estimados e observados a partir do quadrado dos desvios (BARBOSA et al., 2024). O melhor modelo apresenta RMSE igual a zero; valores mais elevados indicam pior desempenho. A Tabela 2 apresenta os valores de referência adotados para as métricas de desempenho neste estudo.

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (Q_{i,sim} - Q_{i,obs})^2}{n}} \quad (08)$$

Tabela 2. Valores de referência utilizados nas métricas de desempenho.
Table 2. Reference values used for the performance metrics.

Métricas	NSE	PBIAS (%)	RMSE (m^3/s)
Muito bom	(0,80-1,00]	$PBIAS < \pm 5$	$\leq 0,31DP$
Bom	(0,70-0,80]	$\pm 5 \leq PBIAS < \pm 10$	$\leq 0,45DP$
Satisfatório	(0,50-0,70]	$\pm 10 \leq PBIAS < \pm 15$	$\leq 0,83DP$
Insatisfatório	$(-\infty, 0,50]$	$PBIAS \geq \pm 15$	$> 0,83DP$
Referências	[1]	[1]	[2]

Legenda: DP – desvio padrão; [1] – Moriasi et al. (2015); [2] – Dazzi et al. 2021.
Legend: SD – standard deviation; [1] – Moriasi et al. (2015); [2] – Dazzi et al. 2021.

3. RESULTADOS

3.1. Análise estatística dos dados de vazão e precipitação

A Tabela 3 apresenta as medidas descritivas da vazão média mensal e da precipitação média mensal, observadas na estação de Barra do Funchal, sub-bacia do São Francisco, no período de janeiro de 1970 a dezembro de 2022.

Tabela 3. Medidas descritivas da vazão e precipitação de 1970 a 2022.
Table 3. Descriptive statistics of streamflow and precipitation from 1970 to 2022.

Medidas Descritivas	Vazão	Precipitação (mm)
Mediana	9,02	85,35
Média	17,85	126,88
Máximo	140,61	668,80
Mínimo	0,59	0,00
Desvio Padrão	21,00	130,41
Coeficiente de Variação (%)	117,65	102,78
Assimetria	2,47	1,25

A vazão mínima observada foi de 0,59 m^3/s , registrada em outubro de 2017 durante um período de seca que afetou esse afluente (SILVA et al., 2023), enquanto a vazão máxima ocorreu em fevereiro de 1992. Como o desvio padrão superou a média das vazões mensais, obteve-se um

coeficiente de variação elevado, o que indica forte sazonalidade (Figura 2) e torna a média aritmética pouco representativa para este estudo.

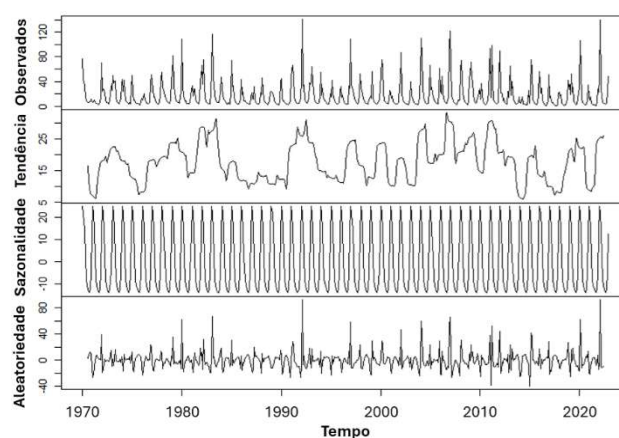


Figura 2. Decomposição da série temporal de vazão em componentes de sazonalidade, tendência e resíduos.

Figure 2. Time series decomposition of streamflow into seasonality, trend and remainder components.

Quanto às medidas descritivas da precipitação mensal, o valor máximo de 668,80 mm foi registrado em fevereiro de

2002. Nesse caso, assim como verificado para as vazões, a média também se mostrou pouco representativa.

3.2. Decomposição das séries temporais

As Figuras 2 e 3 apresentam a decomposição aditiva das séries de vazões e precipitações, em componentes de valores observados, tendência, sazonalidade e resíduos.

A estacionariedade da série de vazões quanto à média e à variância (Figura 2) foi confirmada pelo teste de Dickey-Fuller (SAID e DICKEY, 1984), com rejeição da hipótese nula de não estacionariedade (p -valor $< 0,01$). O mesmo foi verificado para a série de precipitação.

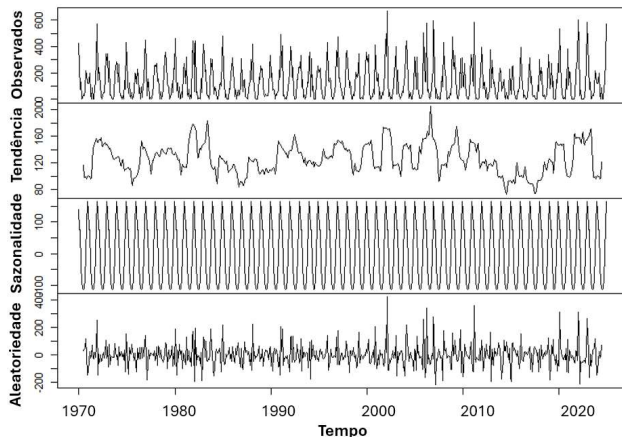


Figura 3. Decomposição da série temporal de precipitação em componentes de sazonalidade, tendência e resíduos. Fonte: Autores.
Figure 3. Time series decomposition of precipitation into seasonality, trend and remainder components. Source: Authors.

3.3. Identificação e seleção dos modelos

A Figura 4 apresenta os correlogramas das funções de autocorrelação amostral (FAC) e de autocorrelação parcial amostral (FACP) para os dados da série temporal de vazões médias mensais.

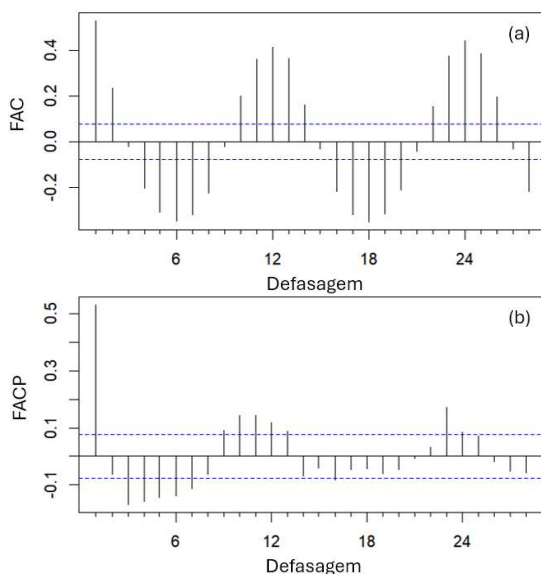


Figura 4. Funções de autocorrelação amostral (FAC) (a) e de autocorrelação parcial amostral (FACP) da série em nível (b).
Figure 4. Sample autocorrelation function (ACF) (a) and sample partial autocorrelation function (PACF) (b) of the series at level.

Verifica-se, por meio dos correlogramas, autocorrelação significativa em defasagens múltiplas de 12, indicando a

presença de sazonalidade. Assim, a identificação da componente sazonal torna-se evidente, incluindo também o comportamento senoidal observado no correlograma da FAC, o que sugere um processo autorregressivo, estacionário e sazonal.

A seleção dos modelos mais adequados foi realizada por meio do comando “auto.arima()”, tendo sido escolhidos o SARIMA (1,0,0)(4,0,0)_[12] e o SARIMAX (1,0,0)(2,0,0)_[12].

Para as diferentes configurações temporais de treino e teste, obtiveram-se valores distintos dos critérios AIC e BIC, cujas menores magnitudes mostraram-se associadas tanto à menor complexidade dos modelos quanto ao emprego de intervalos temporais mais curtos (Tabela 4).

A ausência de correlação serial residual foi verificada por meio do teste de Ljung-Box, que confirmou a condição de ruído branco em todos os modelos (p -valores $> 0,05$), atestando, portanto, a independência dos resíduos.

Tabela 4. Informações sobre os modelos selecionados e os valores de AIC e BIC.
Table 4. Information on selected models and AIC and BIC values.

Modelo	Sigla	Treino/ Teste	AIC	BIC
SARIMA (1,0,0)(4,0,0) _[12]	S6	98/2	5203,982	5235,036
SARIMAX (1,0,0)(2,0,0) _[12]	SX6	98/2	4899,549	4926,165
SARIMA (1,0,0)(4,0,0) _[12]	S5	90/10	4805,575	4836,068
SARIMAX (1,0,0)(2,0,0) _[12]	SX5	90/10	4537,454	4563,591
SARIMAX (1,0,0)(2,0,0) _[12]	SX4	80/20	3972,918	3998,254
SARIMAX (1,0,0)(2,0,0) _[12]	SX3	70/30	3465,769	3490,344
SARIMAX (1,0,0)(2,0,0) _[12]	SX2	60/40	2985,521	3009,225
SARIMAX (1,0,0)(2,0,0) _[12]	SX1	50/50	2519,291	2541,975

3.4. Desempenho dos modelos para diferentes relações entre treino e teste

3.4.1. SARIMA

A avaliação do desempenho dos modelos SARIMA foi conduzida com base nas métricas NSE, KGE, PBIAS e RMSE, além da inspeção visual das previsões (Figura 5). Foram ajustados seis modelos, um para cada relação treino-teste, denominados S1 a S6. Como se observa na Figura 5, o modelo S1 apresenta uma tendência de convergência das previsões para valores próximos de 20 m³/s. Tal comportamento está associado, muito provavelmente, a deficiências no ajuste ou à não inclusão de variáveis exógenas, fatores que limitam a capacidade do modelo em representar tanto os picos quanto os períodos de menor vazão.

No período de treino, o desempenho dos modelos variou entre insatisfatório e muito bom, dependendo da métrica considerada. O NSE apresentou valores entre 0,42 (modelo S3) e 0,45 (modelo S4), ambos classificados como insatisfatórios. O KGE, por sua vez, oscilou de 0,48 (modelo S3) a 0,51 (modelo S4), sendo classificado como bom em ambos os casos. O PBIAS permaneceu entre 0,3% e 0,7% em todos os modelos, o que lhes conferiu a avaliação de muito bom. Com relação ao RMSE (Tabela 5), os valores variaram de 14,20 m³/s (relação 50/50) a 15,48 m³/s (relação 80/20), sendo classificados como satisfatórios.

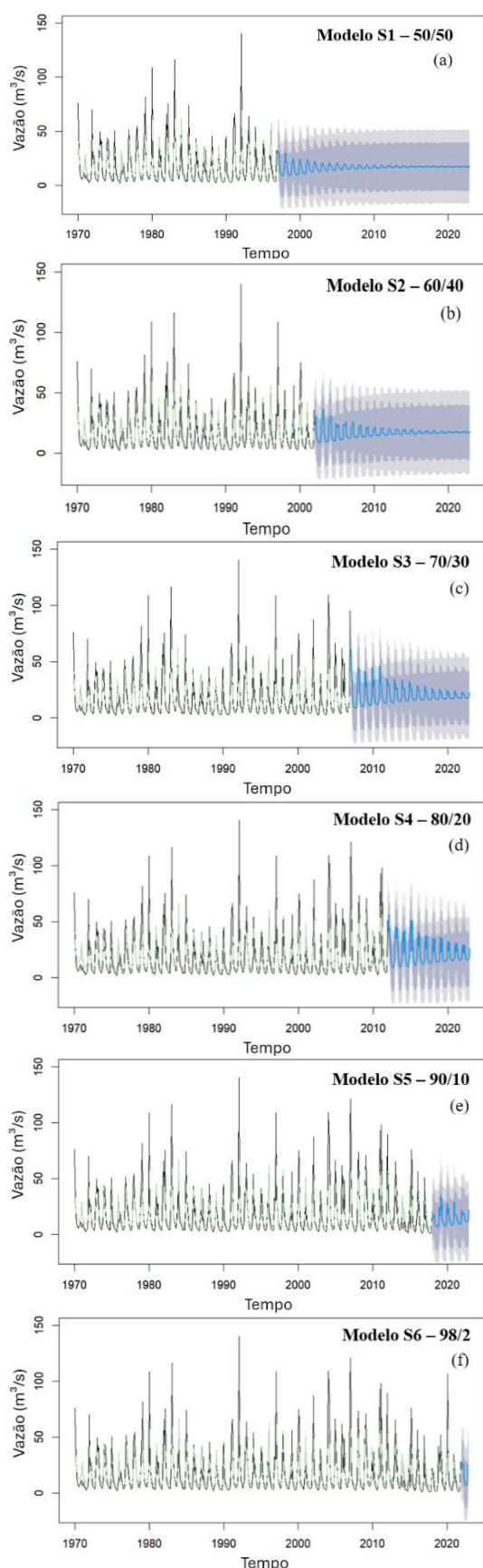


Figura 5. Previsões para os modelos SARIMA com relação treino/teste (a) 50/50, (b) 60/40, (c) 70/30, (d) 80/20, (e) 90/10 e (f) 98/2.

Figure 5. Forecasts for the SARIMA models with training/testing ratios of (a) 50/50, (b) 60/40, (c) 70/30, (d) 80/20, (e) 90/10, and (f) 98/2.

Tabela 5. Dados do RMSE dos modelos nos períodos de treino e de teste.

Table 5. RMSE values for the models during the training and testing periods.

Teste/Treino	SARIMA		SARIMAX	
	Treino	Teste	Treino	Teste
98/2	15,39	34,04	12,13	27,56
90/10	15,39	22,08	12,28	15,99
80/20	15,48	18,94	12,29	14,19
70/30	15,11	19,35	11,80	15,80
60/40	14,39	21,87	11,60	15,71
50/50	14,20	21,86	11,56	15,34

Os valores em negrito correspondem aos menores valores de menores RMSE.

Values in bold indicate the lowest RMSE.

No período de teste, o desempenho foi substancialmente inferior ao observado durante o treinamento. O NSE variou de 0,09 (modelo S1) a 0,31 (modelo S4), sendo classificados como insatisfatórios. O KGE apresentou valores entre -0,04 (modelo S1) e 0,24 (modelo S3), com enquadramento que variou de satisfatório a bom. O PBIAS apresentou grande amplitude: o melhor valor foi obtido pelo modelo S1 (9,7%, classificação boa), enquanto o pior foi registrado pelo modelo S6 (42,5%, classificação insatisfatória). Em relação ao RMSE (Tabela 5), todos os modelos superaram o limiar de 0,83 DP, o que, de acordo com a Tabela 2, os enquadra na categoria insatisfatória.

Embora os resultados das métricas sugerissem o modelo S4 como o mais promissor para previsões, a análise da Figura 5 revela que suas projeções tendem à média em horizontes mais longos, sem captar adequadamente a sazonalidade da série histórica. Portanto, para a previsão de 2022, optou-se por selecionar apenas os modelos S5 e S6, que empregaram os maiores períodos de treinamento.

3.4.2. SARIMAX

A avaliação do desempenho dos modelos SARIMAX foi conduzida com base nas métricas NSE, KGE, PBIAS e RMSE, além da inspeção visual das previsões (Figura 6). Foram ajustados seis modelos, um para cada relação treino-teste, denominados SX1 a SX6.

Na Figura 6, apresentam-se as previsões obtidas com os modelos SARIMAX, que incluem a precipitação como variável exógena, para as diferentes relações treino-teste. Evidencia-se uma melhoria substancial nas previsões, com os modelos capturando a sazonalidade, reproduzindo oscilações entre picos e vales, e sem convergência à média no período de teste. Os dados simulados para o período de treinamento mostraram-se mais próximos dos valores observados.

No período de treino, os valores de NSE variaram de 0,63 (modelo SX2) a 0,65 (modelo SX4), ambos classificados como satisfatórios. O PBIAS situou-se entre 0% e 0,2% em todos os modelos, o que lhes confere uma avaliação muito boa. Quanto ao KGE, os maiores valores foram de 0,68 (modelo SX3) e 0,66 (modelo SX1), ambos classificados como bons.

No período de teste, o NSE variou de 0,45 (modelo SX6) a 0,57 (modelo SX4), com classificações de insatisfatório e satisfatório, respectivamente. O PBIAS apresentou ampla variação: o melhor resultado foi obtido pelo modelo SX4 (3,2%, muito bom), enquanto o pior foi obtido pelo modelo SX6 (20,9%, insatisfatório). Para o KGE, os valores oscilaram entre 0,50 (modelo SX6) e 0,59 (modelo SX4), ambos classificados como bons.

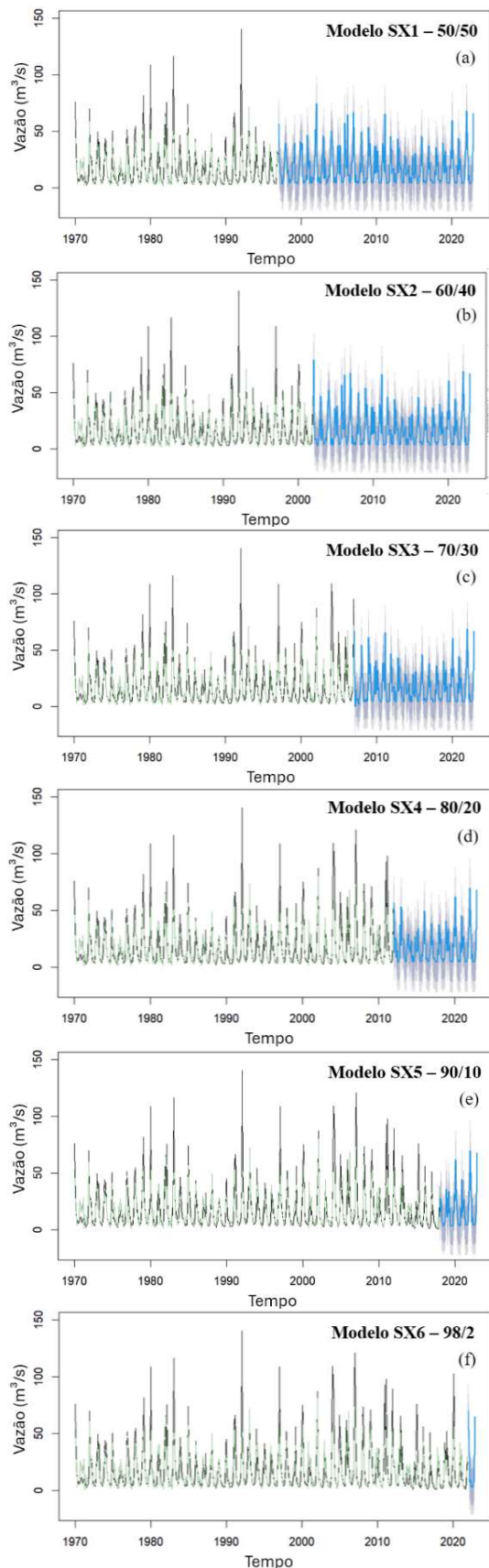


Figura 6. Previsões para os modelos SARIMAX com relação treino/teste (a) 50/50, (b) 60/40, (c) 70/30, (d) 80/20, (e) 90/10 e (f) 98/2.

Figure 6. Forecasts for the SARIMAX models with training/testing ratios of (a) 50/50, (b) 60/40, (c) 70/30, (d) 80/20, (e) 90/10, and (f) 98/2.

Com relação ao RMSE (Tabela 5), os modelos SARIMAX apresentaram valores no período de treino entre 11,56 m³/s e 12,29 m³/s, com melhor desempenho na relação 50/50 e pior na relação 80/20. Segundo os critérios estabelecidos na Tabela 2, ambos os extremos foram enquadrados como satisfatórios. No período de teste, o melhor resultado coube à relação 80/20 (14,19 m³/s, satisfatório), enquanto o pior ocorreu na relação 98/2 (27,56 m³/s), associado ao pico atípico de fevereiro de 2022, o que lhe conferiu classificação insatisfatória.

Com base nos resultados obtidos, todos os modelos SARIMAX foram selecionados para a avaliação das previsões de 2022.

3.5. Comparação SARIMA vs SARIMAX

A Figura 7 apresenta as previsões para o horizonte de 12 meses, referentes ao ano de 2022, considerando os melhores modelos propostos. Observa-se na referida figura, a similaridade entre as previsões dos modelos SARIMAX (SX), com curvas praticamente coincidentes para os valores previstos. Dentre eles, destaca-se a simulação do modelo SX6, que mais se aproximou das vazões observadas. Já para os modelos SARIMA, verificaram-se resultados ruins na previsão, com redução dos valores médios mensais máximos e aumento dos valores médios mensais mínimos. A Figura 8 apresenta os valores das métricas dos modelos SARIMA e SARIMAX para quatro horizontes de previsão.

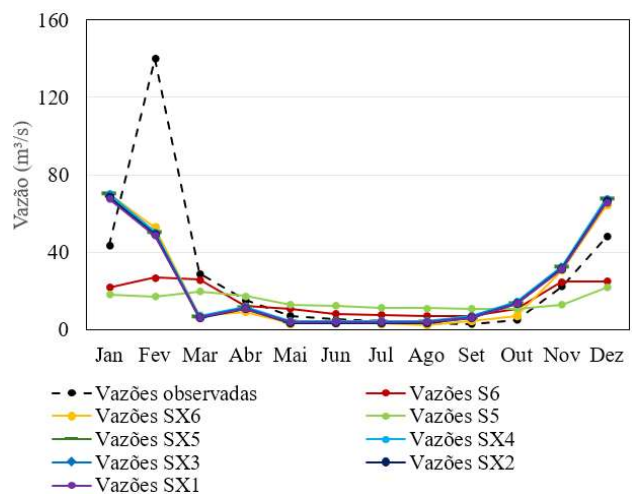


Figura 7. Gráfico das vazões médias mensais observadas e previstas. Figure 7. Observed and predicted monthly average streamflow values.

4. DISCUSSÃO

A presente seção analisa o desempenho dos modelos SARIMA e SARIMAX para diferentes relações treino-teste, discutindo os resultados conforme a literatura. A discussão é organizada por métrica de avaliação, seguida de uma análise integrada dos resultados.

4.1. Coeficiente de Eficiência Nash-Sutcliffe (NSE)

Os modelos SARIMAX apresentaram valores de NSE superiores aos dos modelos SARIMA, indicando um ajuste mais adequado à série temporal. Com base na classificação proposta por Moriasi et al. (2015), no período de treino, apenas os modelos SARIMAX obtiveram coeficientes satisfatórios, enquanto, no SARIMA, apenas coeficientes insatisfatórios.

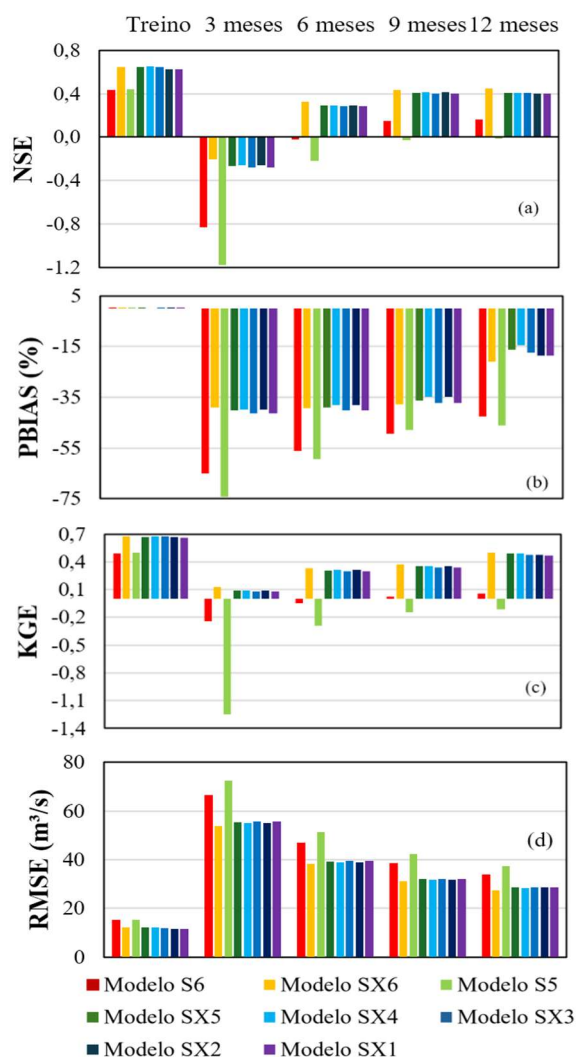


Figura 8. Métricas de desempenho para NSE (a), PBIAS (b), KGE (c) e RMSE (d).

Figure 8. Performance metrics for NSE (a), PBIAS (b), KGE (c), and RMSE (d).

O melhor desempenho no teste foi obtido pelo modelo SARIMA 6 (SX6) no horizonte de 12 meses, além de melhora em todos os modelos SARIMAX nos horizontes de 9 e 12 meses. Esse comportamento se deve ao valor extremo da vazão de fevereiro, o que explica a queda no desempenho nos períodos de 3 e 6 meses.

Em um estudo de reconstrução das vazões naturais de 330 bacias hidrográficas da China, Miao et al. (2022) obtiveram $NSE > 0,70$ em aproximadamente 83% das bacias nas simulações mensais, demonstrando que valores elevados da métrica são alcançáveis em diferentes contextos hidroclimáticos. De modo geral, o desempenho dos modelos SARIMAX foi bastante homogêneo entre si (Figura 6), sugerindo que a variação na relação treino-teste não interfere significativamente na qualidade das previsões nos horizontes analisados.

4.2. Coeficiente de Viés Percentual (PBIAS)

No período de treino, todos os modelos foram classificados como satisfatórios quanto ao PBIAS, com valores próximos de 0%, de acordo com os critérios de Moriasi et al. (2015). No período de teste, o modelo SX4 destacou-se pelos coeficientes mais baixos e foi considerado satisfatório apenas na janela de 12 meses.

Costa et al. (2023), ao avaliarem modelos SARIMA e SARIMAX para previsão mensal de vazões na Bacia do Rio Paranaíba (MG), relataram que o SARIMA apresenta tendência de subestimar picos em média 6,5% e superestimar vales em aproximadamente 76,0%, evidenciando limitações na reprodução de extremos. Em contrapartida, o modelo SARIMAX, com precipitação como variável exógena, superestimou picos em cerca de 11,6% e vales em cerca de 5,0%, demonstrando melhor ajuste para valores extremos.

Em todos os períodos analisados, os modelos SARIMAX superaram os SARIMA, indicando que a inclusão da variável exógena contribui para reduzir o viés sistemático nas previsões.

4.3. Coeficiente de Eficiência Kling-Gupta (KGE)

A métrica KGE corroborou a superioridade dos modelos SARIMAX, com um comportamento análogo ao observado para o NSE no período de treino. Contudo, no horizonte de 3 meses, os valores de NSE e KGE mostraram-se bastante discrepantes, o que se deve à natureza distinta dessas métricas: enquanto o NSE é mais sensível a picos extremos, o KGE pondera de forma balanceada a correlação, o viés e a variabilidade.

Dhanesh et al. (2020), ao avaliarem diferentes produtos de precipitação para simulação mensal de vazões com o modelo SWAT em múltiplas bacias hidrográficas, relataram que os valores de KGE variaram entre 0,62 e 0,82, independentemente das características climáticas, ressaltando a importância da métrica para a avaliação de modelos hidrológicos em diferentes contextos.

Em todos os horizontes, os modelos SARIMAX apresentaram valores de KGE superiores a 0, indicando desempenho satisfatório. O modelo SX6 foi o único a superar o limiar de 0,5 no período de teste de 12 meses.

4.4. Raiz do Erro Quadrático Médio (RMSE)

Todos os modelos SARIMAX apresentaram RMSE inferior ao dos modelos SARIMA, confirmando a maior precisão da modelagem com variável exógena. No período de teste, o RMSE dos modelos SARIMAX reduziu-se à medida que a janela temporal aumentava, um comportamento atribuído ao pico atípico registrado em fevereiro de 2022.

Khodakhah et al. (2022), ao compararem modelos SARIMA com abordagens de *machine learning* para previsão mensal de vazões em duas bacias hidrográficas no Irã, relataram que o SARIMA obteve um RMSE de 1,138 m³/s no período de teste, superando os modelos LSSVM, ANFIS e GMDH, o que evidencia a eficácia de modelos lineares estocásticos para a previsão de vazões em regiões com sazonalidade bem definida.

Os valores de RMSE entre os modelos SARIMAX foram relativamente semelhantes. No horizonte de 12 meses, o SX6 registrou o melhor desempenho (27,56 m³/s), enquanto o SX1 apresentou o pior (28,67 m³/s).

4.5. Comparação com a literatura e implicações

Os resultados indicam que os modelos SARIMA apresentaram evidências de falha de ajuste, uma vez que as métricas NSE, KGE e PBIAS exibiram valores satisfatórios no período de treino e uma acentuada degradação no período de teste, com alguns valores tornando-se negativos. Esse comportamento é característico de modelos que capturam

ruído específico da amostra de treinamento, sem generalizar adequadamente a novos dados.

Barbosa et al. (2024), em estudo com série mensal de vazão de 84 meses em bacia de relevo e em sistemas atmosféricos heterogêneos, constataram que o modelo SARIMA seria suficiente para as previsões. Em contraste, para a Bacia do Indaiá, os resultados indicam que a inclusão de uma variável exógena (precipitação) foi necessária, corroborando os resultados de Bayer et al. (2012), Bleidorn et al. (2019) e Costa et al. (2023), que utilizaram relações treino-teste de 93/7, 98/2 e 89/11, respectivamente.

Conforme descrito na metodologia, a avaliação do desempenho dos modelos baseou-se nos critérios recomendados por Moriasi et al. (2015), que estabelecem faixas de desempenho para as métricas NSE, PBIAS e RMSE aplicadas à modelagem hidrológica em bacias hidrográficas. Para simulação de vazões em escala de bacia, esses autores recomendam como critérios satisfatórios $NSE > 0,50$, $PBIAS < \pm 5\%$ e $R^2 > 0,60$. Nesse contexto, Harmel et al. (2014) destacam que os critérios de desempenho devem ser interpretados de acordo com a finalidade da modelagem, podendo ser mais flexíveis em aplicações de planejamento e de gestão de recursos hídricos do que em estudos de caráter regulatório ou operacional. Para aplicações voltadas ao planejamento, valores de NSE próximos de 0,45 e de PBIAS em torno de $\pm 20\%$ podem ser considerados aceitáveis, desde que o modelo represente adequadamente a dinâmica hidrológica da bacia.

Em relação ao coeficiente de Kling-Gupta (KGE), adotou-se a interpretação proposta por Knoben et al. (2019), segundo a qual valores superiores a -0,41 indicam desempenho superior ao de um modelo de referência (benchmark), enquanto valores iguais ou superiores a 0,50 são geralmente considerados indicativos de bom desempenho.

Dessa forma, os resultados obtidos neste estudo ($NSE \approx 0,45-0,60$, $KGE \geq 0,50$ e PBIAS em torno de $\pm 20\%$ no período de teste) encontram-se dentro das faixas recomendadas na literatura para aplicações de planejamento e gestão de recursos hídricos.

O modelo SX6 (98/2) apresentou desempenho consistentemente superior em diferentes horizontes de previsão, embora ainda tenha subestimado o pico de fevereiro de 2022 (Figura 6). Esse comportamento pode ser atribuído a três fatores: (i) o treinamento com período histórico mais extenso permitiu estimativa mais ampla dos parâmetros autorregressivos e de médias móveis, fundamentais para a captura da memória da série; (ii) a maior diversidade de regimes hidrológicos incluídos no treino ampliou a capacidade de generalização do modelo; e (iii) a relação 98/2, ao utilizar dados imediatamente anteriores ao período de previsão, preservou padrões de correlação temporal de curto prazo, conferindo vantagem preditiva em relação aos demais modelos.

Esse resultado reforça a importância da precipitação como variável exógena para melhorar a acurácia das previsões em bacias com elevada variabilidade hidroclimática, embora a limitação na representação do pico extremo indique que modelos lineares, mesmo com variáveis exógenas, podem não ser suficientes para capturar totalmente eventos de grande magnitude. A relação 98/2, adotada neste trabalho para previsão de um ano, alinha-se à padronização observada na literatura para séries temporais mensais.

5. CONCLUSÕES

O estudo avaliou o desempenho preditivo dos modelos SARIMA e SARIMAX, a influência da inclusão de variáveis exógenas e o efeito de diferentes relações de treino-teste na Bacia do Indaiá, MG. Os resultados demonstram que períodos de treino mais extensos e mais próximos ao horizonte de previsão favorecem ajustes mais consistentes.

Os modelos SARIMAX superaram o SARIMA em todas as métricas avaliadas, confirmando que a inclusão da precipitação como variável exógena melhora significativamente a capacidade preditiva, especialmente diante de eventos extremos, como o pico de precipitação em fevereiro de 2022. Para os modelos SARIMAX, as diferentes relações treino-teste não afetaram substancialmente os resultados, indicando estabilidade estrutural e superioridade em relação aos modelos SARIMA.

Embora a relação 98/2 tenha apresentado o melhor desempenho no horizonte de um ano, essa constatação deve ser interpretada como uma recomendação preliminar, restrita às condições da bacia hidrográfica e ao período analisado, e não deve ser generalizada como regra universal para outras localidades ou escalas temporais.

Destaca-se que os resultados e conclusões deste estudo estão circunscritos ao contexto em que foram gerados: uma única bacia hidrográfica, variáveis com resolução temporal mensal, série histórica restrita a um período específico e consideração de apenas uma variável exógena (precipitação). Ademais, os modelos empregados são de natureza linear, o que pode restringir a capacidade de capturar relações não lineares e eventos extremos mais complexos.

Diante dessas limitações, recomenda-se que estudos futuros avaliem a incorporação de outras variáveis exógenas (p.ex., temperatura), aplicação de modelos de *machine learning* e a investigação de diferentes horizontes de previsão e estratégias de validação, com vistas a ampliar a generalização e a aplicabilidade dos modelos em diferentes contextos hidrológicos.

6. REFERÊNCIAS

- ALHARBI, F. R.; CSALA, D. A. Seasonal autoregressive integrated moving average with exogenous factors (SARIMAX) forecasting model-based time series approach. *Inventions*, v. 7, n. 4, e94, 2022. <https://doi.org/10.3390/inventions7040094>
- AKAIKE, H. A new look at the statistical model identification. In: PARZEN, E.; TAMABE, K.; KITAGAWA, G. (Eds.). *Springer Series in Statistics*. New York: Springer. 1974. p. 215-222. https://doi.org/10.1007/978-1-4612-1694-0_16
- ANA_Agência Nacional de Águas e Saneamento Básico. **Hidroweb: sistema de informações hidrológicas**. Brasília, DF, 2024. Disponível em: <https://www.snirh.gov.br/hidroweb/>. Acesso em: 15 out. 2024.
- BARBOSA, M. H.; CARVALHO, M. de M.; SÁFADI, T.; OLIVEIRA, L. F. C. de. Modelagem estocástica de séries temporais pluviométricas e fluviométrica da bacia hidrográfica do Rio Preto, Sudeste do Brasil. *Revista Brasileira de Meteorologia*, v. 39, e39240051, 2024. <https://doi.org/10.1590/0102-77863910051>
- BARRIENTOS-TORRES, D.; MARTÍNEZ-RÍOS, E. A.; NAVARRO-TUCH, S. A.; PABLOS-HACH, J. L.; BUSTAMANTE-BELLO, R. Water flow modeling and

- forecast in a water branch of Mexico City through ARIMA and transfer function models for anomaly detection. **Water**, v. 15, n. 15, e2792, 2023. <https://doi.org/10.3390/w15152792>
- BAYER, D. M.; CASTRO, N. M. dos R.; BAYER, F. M. Modelagem e previsão de vazões médias mensais do rio Potiribu utilizando modelos de séries temporais. **Revista Brasileira de Recursos Hídricos**, v. 17, n. 2, p. 229-239, 2012. <https://doi.org/10.21168/rbrh.v17n2.p229-239>
- BAYER, F. M.; SOUZA, A. M. Forecasting with Wavelets and Traditional Models: A Comparative. **Revista Brasileira de Biometria**, v. 28, p. 40-61, 2010.
- BLEIDORN, M. T.; PINTO, W. P.; BRAUM, E. S.; LIMA, G. B.; MONTEBELLER, C. A. Modelagem e previsão de vazões médias mensais do rio Jacu, ES, utilizando o modelo SARIMA. **IRRIGA**, v. 24, n. 2, p. 320-335, 2019. <https://doi.org/10.15809/irriga.2019v24n2p320-335>
- BOX, G. E. P.; JENKINS, G. M. **Time series analysis, forecasting and control**. San Francisco: Holden-Day, 1976. 575p.
- CAMINHA, A. R.; GUIMARÃES, L. R.; CORRÊA, F. V.; SOUZA, G. R. de; OLIVEIRA, L. F. C. de; SÁFADI, T. (2025). Previsão de precipitação e vazão a partir da modelagem de séries temporais. **Revista Brasileira de Climatologia**, v. 36, n. 21, p. 264-290, 2025. <https://doi.org/10.55761/abclima.v36i21.18885>
- COSTA, G. E. de M. e; MENEZES FILHO, F. C. M. de; CANALES, F. A.; FAVA, M. C.; BRANDÃO, A. R. A.; PAES, R. P. Assessment of time series models for mean discharge modeling and forecasting in a sub-basin of the Paranaíba River, Brazil. **Hydrology**, v. 10, n. 11, e208, 2023. <https://doi.org/10.3390/hydrology10110208>
- DAZZI, S.; VACONDIO, R.; MIGNOSA, P. Flood stage forecasting using machine-learning methods: a case study on the Parma River (Italy). **Water**, v. 13, n. 12, e1612, 2021. <https://doi.org/10.3390/w13121612>
- DHANESH, Y.; BINDHU, V. M.; SENENT-APARICIO, J.; BRIGHENTI, T. M.; AYANA, E.; SMITHA, P. S.; FEI, C.; SRINIVASAN, R. A comparative evaluation of the performance of CHIRPS and CFSR data for different climate zones using the SWAT model. **Remote Sensing**, v. 12, e3088, 2020. <https://doi.org/10.3390/rs12183088>
- DUC, L.; SAWADA, Y. A signal-processing-based interpretation of the Nash–Sutcliffe efficiency. **Hydrology and Earth System Sciences**, v. 27, n. 9, p. 1827-1839, 2023. <https://doi.org/10.5194/hess-27-1827-2023>
- FAZLA, A.; AYDIN, M.E.; KOZAT, S.S. Joint optimization of linear and nonlinear models for sequential regression. **Digital Signal Processing**, v. 132, e103802, 2023. <https://doi.org/10.1016/j.dsp.2022.103802>
- GUPTA, H. V.; SOROOSHIAN, S.; YAPO, P. O. Status of automatic calibration for hydrologic models: comparison with multilevel expert calibration. **Journal of Hydrologic Engineering**, v. 4, n. 2, p. 135-143, 1999. [https://doi.org/10.1061/\(ASCE\)1084-0699\(1999\)4:2\(135\)](https://doi.org/10.1061/(ASCE)1084-0699(1999)4:2(135))
- HARMEL, R. D.; SMITH, P. K.; MIGLIACCIO, K. W.; CHAUBEI, I.; DOUGLAS-MANKIN, K. R.; BENHAM, B.; SHUKLA, S.; MUÑOZ-CARPENA, R.; ROBSON, B. J. Evaluation, interpreting, and communicating performance of hydrologic/water quality models considering intended use: A review and recommendations. **Environmental Modelling & Software**, v. 57, p. 40-51, 2014. <https://doi.org/10.1016/j.envsoft.2014.02.013>
- HYNDMAN, R.J.; KHANDAKAR, Y. Automatic Time Series Forecasting: The Forecast Package for R. **Journal of Statistical Software**, v. 27, p. 1-22, 2008. <https://doi.org/10.18637/jss.v027.i03>
- HYNDMAN, R. J.; ATHANASOPOULOS, G. **Forecasting: principles and practice**. 3. ed. Melbourne: OTexts, 2021. 442p.
- KASSEM, A. A.; RAHEEM, A. M.; KHIDIR, K. M. Daily streamflow prediction for Khazir River Basin Using ARIMA and ANN Models. **Zanco Journal of Pure and Applied Sciences**, v. 32, p. 30-39, 2020. <https://doi.org/10.21271/zipas.32.3.4>
- KHODAKHAH, H.; AGHELPOUR, P.; HAMED, Z. Comparing linear and non-linear data-driven approaches in monthly river flow prediction, based on the models SARIMA, LSSVM, ANFIS, and GMDH. **Environmental Science and Pollution Research**, v. 29, p. 21935-21954, 2022. <https://doi.org/10.1007/s11356-021-17443-0>
- KNOBEN, W. J. M.; FREER, J. E.; WOODS, R. A. Inherent benchmark or not? Comparing Nash–Sutcliffe and Kling–Gupta efficiency scores. **Hydrology and Earth System Sciences**, v. 23, n. 10, p. 4323-4331, 2019. <https://doi.org/10.5194/hess-23-4323-2019>
- MACHIWAL, D.; JHA, M. K. **Hydrologic time series analysis: theory and practice**. Dordrecht: Springer, 2012. 316p.
- MENEZES FILHO, F. C. M. Modelagem e previsão de temperaturas médias mensais de Rio Paranaíba/MG utilizando modelos de séries temporais. **Revista Ibero Americana de Ciências Ambientais**, v. 11, n. 6, p. 251-261, 2020. <http://doi.org/10.6008/CBPC2179-6858.2020.006.0021>
- MIAO, C.; GOU, J.; FU, B.; TANG, Q.; DUAN, Q.; CHEN, Z.; LEI, H.; CHEN, J.; GUO, J.; BORTHWICK, A. G. L.; DING, W.; DUAN, X.; LI, Y.; KONG, D.; GUO, X.; WU, J. High-quality reconstruction of China's natural streamflow. **Science Bulletin**, v. 67, p. 547-556, 2022. <https://doi.org/10.1016/j.scib.2021.09.022>
- MORIASI, D. N.; GITAU, M. W.; PAI, N.; DAGGUPATI, P. Hydrologic and water quality models: performance measures and evaluation criteria. **Transactions of the ASABE**, v. 58, n. 6, p. 1763-1785, 2015. <https://doi.org/10.13031/trans.58.10715>
- R CORE TEAM. **R: a language and environment for statistical computing**. Vienna: R Foundation for Statistical Computing, 2017.
- SAID, S. E.; DICKEY, D. A. Testing for unit roots in autoregressive-moving average models of unknown order. **Biometrika**, v. 71, n. 3, p. 599-607, 1984. <https://doi.org/10.2307/2336570>
- SHAH, H. S. M.; YASSIN, M. A.; ABBA, S. I.; LAWAL, D. U.; HUSSEIN AL-QADAMI, E. H.; TEO, F. Y.; MUSTAFFA, Z.; ALJUNDI, I. H. Flood risk and vulnerability from a changing climate perspective: an overview focusing on flash floods and associated hazards in Jeddah. **Water**, v. 15, n. 20, e3641, 2023. <https://doi.org/10.3390/w15203641>
- SIVAKUMAR, M.; PARTHASARATHY, S.; PADMAPRIYA, T. Trade-off between training and

testing ratio in machine learning for medical image processing. **PeerJ Computer Science**, v. 10, e2245, 2024. <https://doi.org/10.7717/peerj-cs.2245>

SUN, Y.; NIU, J.; SIVAKUMAR, B. A. Comparative Study of Models for Short-Term Streamflow Forecasting with Emphasis on Wavelet-Based Approach. **Stochastic Environmental Research and Risk Assessment**, v. 33, p. 1875-1891, 2019. <https://doi.org/10.1007/s00477-019-01734-7>

Agradecimentos: Os autores agradecem ao Instituto de Ciências Exatas e Tecnológicas e ao Curso de Engenharia Civil da UFV-CRP pelo apoio institucional ao desenvolvimento desta pesquisa.

Contribuições dos autores: J.R.da.S.N.: concepção, coleta de dados, análise de dados, redação (rascunho original), discussão dos resultados. F.C.M.de.M.F.: concepção, validação, supervisão, metodologia, discussão dos resultados, redação (revisão e edição). Todos os autores leram e concordaram com a versão publicada do manuscrito.

Financiamento: Esta pesquisa não recebeu financiamento de agências públicas, comerciais ou sem fins lucrativos.

Disponibilidade de dados: Os dados desta pesquisa poderão ser obtidos por e-mail, mediante solicitação ao autor correspondente.

Conflito de interesses: Os autores declaram não haver conflitos de interesses.



Copyright: © 2026 by the authors. This article is an Open-Access article distributed under the terms and conditions of the Creative Commons **Attribution-NonCommercial (CC BY-NC)** license (<https://creativecommons.org/licenses/by/4.0/>).