



Previsão de vazão na bacia hidrográfica do rio Manuel Alves da Natividade utilizando o modelo de séries temporais SARIMA

Victor Braga Rodrigues Duarte^{a*}, Francisca de Cássia Silva da Silva^a, Igor Viana Souza^a,
Marcos Vinicius Cardoso Silva^a, Hygor Gomes de Almeida Sousa^a, Marcos Giongo^a, Marcelo Ribeiro Viola^b

^a Universidade Federal do Tocantins (UFT), Brasil

^b Universidade Federal de Lavras (UFLA), Brasil

* Autor correspondente (victorbrduarte@uft.edu.br)

INFO

Keywords

hydrological simulation
hydrology
time series
water flow

ABSTRACT

Streamflow forecast in the Manuel Alves da Natividade watershed using SARIMA time series model

Streamflow knowledge and modeling is fundamental for water resources management. Hydrologic models help on water supply planning and projects regarding water budget, hydropower generation, irrigation systems, sustainable use and biodiversity conservation. The study area is Manuel Alves da Natividade (BH_{MAN}) watershed, located in the Tocantins-Araguaia river basin. The BH_{MAN} is one of the main tributaries of the Tocantins river. In this context, the objective was to analyze, model and make predictions of the surface runoff based on the time series model SARIMA, with monthly step. The methodology adopted was that of Box and Jenkins which consists of: identifying the model; estimate their parameters and apply the adjusted model for forecasting. The streamflow series was verified for its trend, seasonality and stationarity. The identification of the component orders of the models was made through the graphical analysis of the correlograms and periodograms. The selection of the best model was performed based on the selection criteria and quality measures. The model chosen was the SARIMA(1,0,1)(1,1,4)₁₂, which presented Nash-Sutcliffe coefficients of 0.61 and 0.81 for the calibration and validation steps, respectively, which are considered good according to the classification applied for conceptual hydrological models. The model proved good performance in predicting hydrography recession, being indicated mainly as a tool for water resources management, irrigation planning and water supply.

RESUMO

O estudo e simulação do escoamento superficial de uma bacia hidrográfica são fundamentais para o gerenciamento dos recursos hídricos. Os modelos hidrológicos são ferramentas que auxiliam o planejamento do abastecimento hídrico, projetos de uso da água, geração de energia hidrelétrica, sistemas de irrigação, utilização sustentável e conservação da biodiversidade de uma bacia hidrográfica. A área em estudo é a bacia hidrográfica do rio Manuel Alves da Natividade, uma importante bacia hidrográfica pertencente a região hidrográfica Tocantins-Araguaia, um dos principais afluentes do rio Tocantins. Neste contexto, objetivou-se analisar, modelar e fazer previsões com base nas séries temporais de vazões mensais da bacia hidrográfica do rio Manuel Alves da Natividade utilizando o modelo SARIMA. A metodologia adotada foi a de Box e Jenkins que consiste em: identificar o modelo; estimar seus parâmetros e aplicar o modelo ajustado para previsão. A série foi verificada quanto a tendência, sazonalidade e estacionariedade e a identificação da ordem dos componentes dos modelos se deu por meio da análise gráfica dos correlogramas e periodograma. A seleção do melhor modelo foi realizada com base nos critérios de seleção e medidas de qualidade. O modelo escolhido foi o SARIMA(1,0,1)(1,1,4)₁₂, que apresentou coeficiente de Nash-Sutcliffe de 0,61 e 0,81 para as etapas de calibração e validação, respectivamente, valores estes classificados como “bom”, segundo a classificação para modelos hidrológicos. O modelo mostrou-se eficaz na previsão de períodos de recessão do escoamento, sendo indicado principalmente como ferramenta na gestão de recursos hídricos para a previsão da oferta hídrica em períodos de estiagem críticos.

Palavras-chaves

simulação hidrológica
hidrologia
séries temporais
escoamento superficial

INTRODUÇÃO

A água é um recurso natural indispensável à vida. Seu consumo aumenta à medida que ocorre o desenvolvimento econômico e urbano nos países, cenário evidenciado pelas previsões publicadas no Relatório Mundial das Nações Unidas sobre o Desenvolvimento dos Recursos Hídricos em 2015, que indicam um aumento de 55% no consumo mundial de água para o ano de 2050 e uma redução hídrica em 40% até o ano de 2030 (UNESCO, 2015). Nas duas últimas décadas houve no Brasil o aumento de aproximadamente 80% da retirada de água dos rios e reservatórios, com previsão de aumento em 24% até o ano 2030 (ANA, 2018).

Nota-se a importância do estudo temporal do escoamento superficial no planejamento do abastecimento hídrico, projetos de uso da água, geração de energia hidrelétrica, sistemas de irrigação, utilização sustentável e conservação da biodiversidade de uma bacia hidrográfica (Viola et al., 2009; Zhang et al., 2011; Zhao et al., 2017). Dentre as ferramentas utilizadas na gestão hídrica das bacias, modelos hidrológicos são representações matemáticas dos fenômenos do ciclo hidrológico, úteis na gestão ambiental com vistas a lidar com problemas ambientais, avaliar impactos futuros e subsidiar planejamentos (Almeida & Serra, 2017).

A previsão do escoamento superficial tem sido o foco de pesquisas hidrológicas com a utilização de modelos hidrológicos conceituais e semi-conceituais desde a década de 1950. Este tipo de pesquisa tem como obstáculo a reduzida disponibilidade de séries históricas hidrológicas, e a obtenção de uma vasta quantidade de parâmetros de solo e vegetação de difícil obtenção, requeridas por modelos tais como o SWAT (Viana et al., 2018), SMAP (Miranda et al., 2017), TOPMODEL (Xue et al., 2018), IPH II (Uliana et al., 2018), dentre outros. Diante disso, muitos são os métodos e modelos estudados com a finalidade de simplificar e melhorar a previsão de vazões. Ganha destaque nessa abordagem os modelos estocásticos, que consideram os processos controlados por leis probabilísticas (Tucci, 1998), com destaque para os modelos de séries temporais.

As séries temporais são capazes de representar a trajetória de um fenômeno durante o tempo e prever seu comportamento futuro. Séries temporais são conjuntos de observações de algum fenômeno ordenado no tempo, cujo objetivo da análise é identificar dependências temporais, conhecer as periodicidades pertinentes, caracterizar seu comportamento e fazer previsões (Bayer & Souza, 2010).

Os modelos de séries temporais mais utilizados são os modelos Autorregressivos (AR), Médias Móveis (MA), e Autorregressivos Integrados de

Médias Móveis (ARIMA). Especificamente os modelos ARIMA, os quais combinam os parâmetros e realizam a interação entre eles, podem incorporar a sazonalidade das séries gerando um modelo SARIMA, destaque em previsões do escoamento superficial (Bayer et al., 2012; Bleidorn et al., 2019).

Santos & Oliveira (2016) ao analisarem a série de vazões mensais do rio Tocantins escolheram o modelo SARIMA(0,0,2)(1,2,2)₁₂ que apresentou previsão da variabilidade da série com coeficiente de Nash-Sutcliffe (C_{NS}) de 0,92. Dhote et al. (2018) objetivando a predição da vazão mensal da bacia hidrográfica do rio Narmada, Índia, após processamento dos dados no Hadoop data cluster, HDFS e MapReduce, encontraram como melhor conformação o SARIMA(1,1,1)(1,1,1)₁₂. Bleidorn et al. (2019) ao estudarem a série temporal de vazões mensais do rio Jucu, localizado no estado do Espírito Santo, concluíram que o melhor modelo foi SARIMA (1,0,0)(5,1,0)₁₂.

Diante da necessidade de gerar informações para subsidiar o gerenciamento dos recursos hídricos na bacia hidrográfica do rio Manuel Alves da Natividade, conceituada como importante região ecológica do sul do estado do Tocantins, objetivou-se neste estudo analisar, modelar e fazer previsões da série temporal de vazões mensais utilizando o modelo SARIMA.

MATERIAL E MÉTODOS

A área em estudo é a bacia hidrográfica do rio Manuel Alves da Natividade (BH_{MAN}), localizada na região hidrográfica Tocantins-Araguaia. O rio Manuel Alves da Natividade é um dos principais afluentes do rio Tocantins por sua margem direita. A BH_{MAN} está localizada no sudeste do estado do Tocantins, entre as coordenadas geográficas 11°09'45" e 12°14'54" de latitude Sul e 46°33'04" e 48°18'40" de longitude Oeste. A BH_{MAN} é limitada pelas bacias do rio das Balsas ao norte, rio Palma ao sul, rio Tocantins a oeste e São Francisco a leste, integrando território de 11 municípios: Rio da Conceição, Porto Alegre do Tocantins, Almas, Dianópolis, Natividade, Chapada da Natividade, São Valério da Natividade, Conceição do Tocantins, Taipas do Tocantins, Pindorama do Tocantins, Santa Rosa do Tocantins. Mediante o modelo digital de elevação TOPODATA, disponibilizado pelo Instituto Nacional de Pesquisas Espaciais (INPE, 2012), com resolução espacial de aproximadamente 30 m, elaborou-se o mapa da bacia (Figura 1), no qual pode-se verificar cotas altimétricas variando de 213 à 954 m. Os recursos hídricos superficiais são utilizados na BH_{MAN} principalmente para irrigação, dessedentação animal e abastecimento urbano.

A série histórica de vazões foi obtida junto ao Sistema de Informações Hidrológicas da Agência Nacional de Águas. Foi selecionado o posto fluviométrico Fazenda Lobeira, localizado no município São Valério da Natividade, cujo código é 22250000, suas coordenadas geográficas são 11°31'58" de latitude Sul e 48°17'18" de longitude Oeste, em uma altitude de 230 m com área de drenagem de aproximadamente 14.500 km².

A BH_{MAN} possui área de drenagem total de 14.917 km². Para a finalidade do presente estudo, a área de drenagem delimitada a montante do posto fluviométrico Fazenda Lobeira é de 14.456,7 km², com perímetro de 868,1 km e 185,2 km de comprimento axial. A precipitação média anual é de 1.382 mm ano⁻¹, o deflúvio médio anual é de 426 mm ano⁻¹ e o relevo é suave-ondulado (Rodrigues et al., 2015).

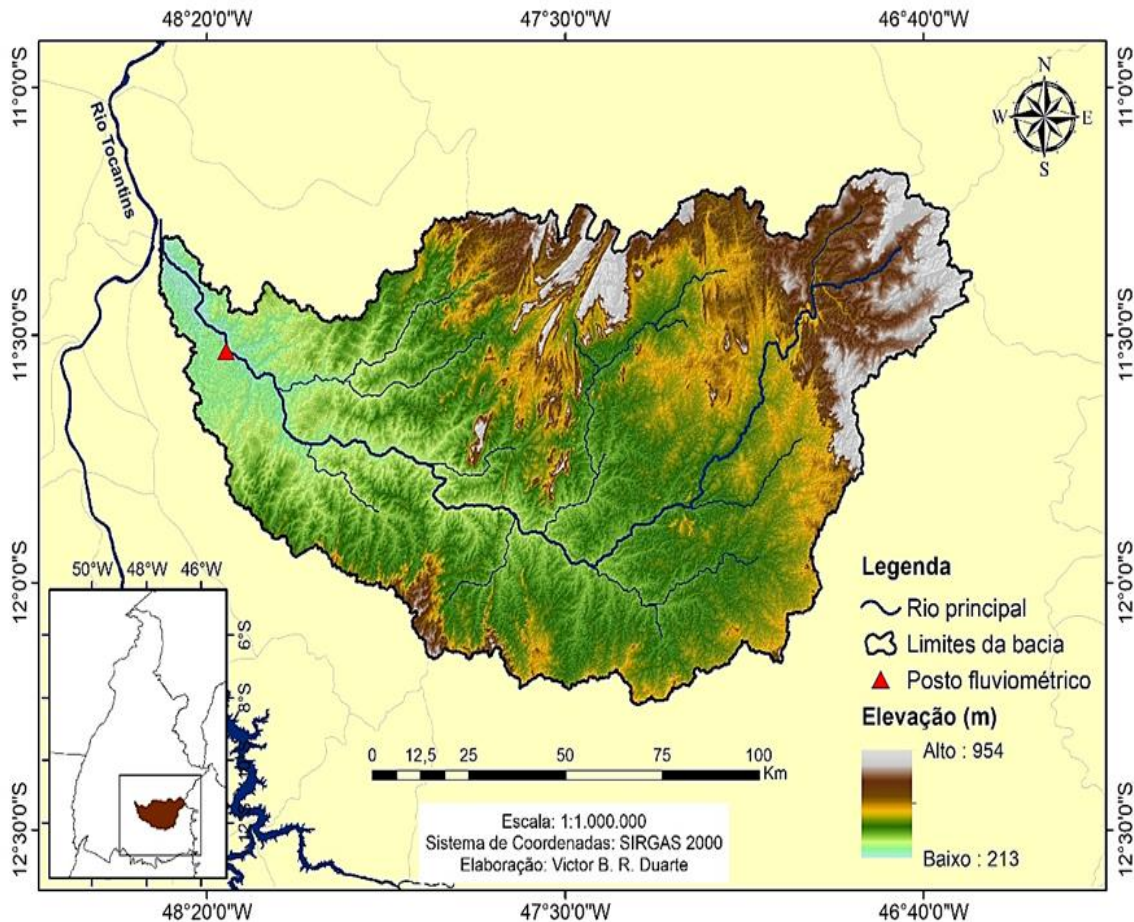


Figura 1 - Localização da bacia hidrográfica do rio Manuel Alves da Natividade

A série histórica de vazão do posto Fazenda Lobeira compreende o período de setembro de 1969 a dezembro de 2018, 49 anos, com apenas 0,5% de seus dados faltantes. A série foi avaliada quanto a qualidade dos dados por meio de inspeção visual e os dados faltantes analisados e imputados por intermédio do Filtro de Kalman, que constitui um conjunto de equações matemáticas que por meio de processos recursivos estimam estados passados, presente e futuro mediante valores observados (Gómez e Maravall, 1994; Reisen et al., 2008; Teodoro, 2019).

Os dados foram divididos em dois subconjuntos para a avaliação da previsão, sendo destinado para

etapa de modelagem (calibração) o período de setembro de 1969 a dezembro de 2017, e o ano de 2018 para avaliação das previsões (validação).

Foi realizada uma análise descritiva da série para elucidação e identificação de tendências e sazonalidades. Modelos para exploração de séries temporais são regidos por leis probabilísticas, visto que uma série temporal é a realização de processos estocásticos (Batista, 2009). Um modelo de séries temporais tradicional pode ser retratado como:

$$Z_t = T_t + S_t + a_t \quad t = 1, 2, 3, \dots, n$$

em que Z_t corresponde a série temporal, T_t a tendência, S_t a sazonalidade e a_t o componente aleatório.

De acordo com Morettin e Toloi (2006), a tendência corresponde ao gradativo aumento ou diminuição das observações ao longo do tempo, a sazonalidade é a flutuação da série em períodos menores e o componente aleatório, ou ruído branco, a oscilação aleatória irregular. A maioria das séries temporais ambientais possuem tendência e sazonalidade, representando uma não-estacionariedade dos dados, ou seja, estes não oscilam ao redor de uma média e variância constantes.

Para verificação de tendência, sazonalidade e estacionariedade a série de vazões mensais foi decomposta para análise visual. Séries estacionárias são necessárias para o ajuste de modelos da classe ARIMA, portanto a série em estudo foi verificada quanto à estacionariedade por meio do teste de Dickey-Fuller Aumentado (Dickey e Fuller, 1981). Segundo Zhang et al. (2011), um procedimento amplamente adotado para tornar as séries estacionárias é a diferenciação sucessiva da série original até que esta se torne estacionária, geralmente sendo suficiente a primeira diferenciação.

Modelos ARIMA admitem que Z_t segue um processo autoregressivo (AR) de ordem p , de médias móveis (MA) de ordem q , integradas (I), ou seja, que sofreu diferenciações para se tornar estacionária de ordem d . O modelo ARIMA(p,d,q) possui como estrutura:

$$\phi_p(B)(1-B)^d Z_t = \theta_q(B)a_t$$

em que $\phi_p(B)$ é o polinômio autorregressivo, $(1-B)^d$ o operador de diferença generalizado, d o número de diferenças para remover a tendência da série, $\theta_q(B)$ o polinômio de médias móveis e a_t o componente aleatório.

Conforme mencionado, processos hidrológicos geralmente possuem sazonalidade, necessitando de modelos que incorporem essa característica temporal. Diante disso modelos da classe ARIMA são ampliados para SARIMA(p,d,q)(P,D,Q)_s, em que (p,d,q) são ordens da dinâmica ordinal e (P,D,Q) são ordens da parte sazonal. Modelos SARIMA apresentam autocorrelação para uma estação de sazonalidade, expressos por:

$$\phi(B)\Phi(B)(1-B)^d(1-B^s)^D Z_t = \theta(B)\Theta(B)a_t$$

os parâmetros Φ_1, \dots, Φ_p e $\Theta_1, \dots, \Theta_p$ correspondem aos autorregressivos e de médias móveis sazonais, respectivamente, e s é o período sazonal da série. A confirmação da sazonalidade da série foi realizada mediante aplicação do teste de Kruskal-Wallis (Morettin & Toloi, 2006) e a identificação do período sazonal foi realizada por meio da análise do periodograma.

Adotando a suposição de que o modelo ARIMA possui base estocástica, podendo ser descrito e caracterizado por seus dados passados, adotou-se a metodologia de Box e Jenkins (Box et al., 2008), que consiste em: identificar o modelo; estimar seus parâmetros; e aplicar o modelo ajustado para previsão.

O processo de identificação tem por objetivo determinar os valores de p , d e q , e P , D e Q , por meio da interpretação dos gráficos de correlação serial e periodograma (Akaike, 1974; 1978). Os correlogramas representam as autocorrelações ao longo do tempo enquanto que o periodograma da frequência, importantes para identificar a periodicidade dos dados, sazonalidades e ciclos da série (Chechi & Bayer, 2012). A função de autocorrelação (FAC) e autocorrelação parcial (FACP) foram observadas para auxiliar na identificação dos modelos.

Para seleção do melhor modelo tomou-se como critério os menores valores das medidas de qualidade de ajuste: critério de informação de Akaike (AIC), critério Bayesiana (BIC) e da raiz do quadrado médio do erro (RQME), expressos conforme as seguintes equações:

$$AIC = -2\log(L) + 2k$$

$$BIC = -2\log(L) + k\log(n)$$

$$RQME = \sqrt{\frac{1}{n} \sum_{t=1}^n (Q_t - \hat{Q}_t)^2}$$

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, Q_t e \hat{Q}_t são, respectivamente, os valores de vazão observados e estimados no instante t .

Os modelos ainda foram analisados quanto a sua eficiência por meio do coeficiente de Nash-Sutcliffe (C_{NS}), segundo a equação a seguir:

$$C_{NS} = 1 - \frac{\sum_{t=1}^n (Q_t - \hat{Q}_t)^2}{\sum_{t=1}^n (Q_t - \bar{Q})^2}$$

De acordo com a classificação proposta por Moriasi et al. (2007), quanto mais alto o C_{NS} melhor o ajuste do modelo, sendo: $C_{NS} = 1$, muito bom; $0,54 < C_{NS} < 0,65$, bom; $0,50 < C_{NS} < 0,54$, satisfatório. Zappa (2002) sugere que os modelos podem ser utilizados para previsão se o C_{NS} for superior à 0,5.

O diagnóstico do modelo que apresentou melhores resultados na etapa de seleção foi realizado mediante análise dos resíduos, que constitui fundamental representação se o modelo capturou adequadamente as informações dos dados. Foram adotadas as suposições que os resíduos não devem ser autocorrelacionados, visto que uma correlação entre eles significa que existem informações que deveriam ter sido incorporadas no modelo; que os resíduos devem ter distribuição normal; e que a média dos resíduos deve ser zero ou próxima de zero para que as previsões não sejam tendenciosas (Wei, 2006). O teste de normalidade dos erros de Shapiro-Wilk foi aplicado (Shapiro & Wilk, 1965) bem como o teste de Ljung-Box (Ljung e Box, 1978) para autocorrelação dos resíduos.

Tendo sido concluída a calibração do modelo prosseguiu-se para a análise das previsões, sendo estas verificadas por meio da comparação entre valores de vazão observados e os previstos pelo modelo, segundo o RQME (Equação 6), C_{NS} (Equação 7) e coeficiente de determinação (R^2) de

acordo com a fórmula a seguir:

$$R^2 = \frac{\sum(Q_t - \bar{Q}_t)^2 - \sum(Q_t - \hat{Q}_t)^2}{\sum(Q_t - \bar{Q}_t)^2}$$

em que Q_t , \bar{Q}_t e \hat{Q}_t são, respectivamente, os valores de vazão observados, médios e estimados no instante t e n é o número total de dados.

O software R (R Development Core Team, 2019) foi utilizado em todas as etapas de desenvolvimento deste trabalho.

RESULTADOS E DISCUSSÃO

A série histórica de vazões mensais da BH_{MAN} é representada na Figura 2. A série possui dados faltantes nos meses de outubro de 1992 e julho de 2017. A imputação dos dados faltantes por meio do Filtro de Kalman apresentou resultados satisfatórios de acordo com análise visual do hidrograma, visto que os valores imputados se ajustaram ao comportamento da série de forma adequada.

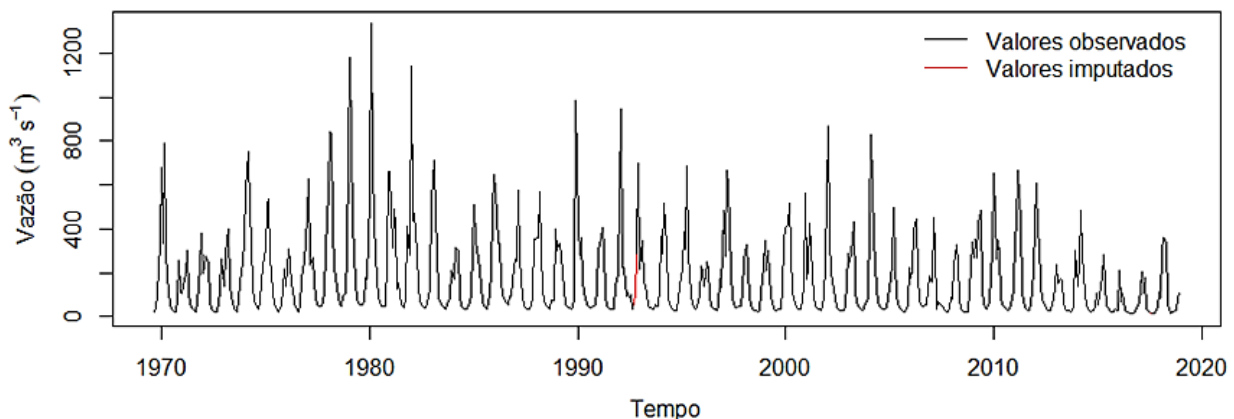


Figura 2 - Série completa de vazões mensais do rio Manuel Alves da Natividade, com destaques para os valores imputados por meio do filtro de Kalman.

Observa-se uma pequena tendência negativa na série, confirmada por meio do teste Cox-Stuart ($p = 0,023$). A partir do teste de Dickey-Fuller Aumentado nota-se que a série é estacionária na média ($p = 0,01$). A vazão média do período foi de $179,91 \text{ m}^3$

s^{-1} , com desvio padrão de $192,56 \text{ m}^3 \text{ s}^{-1}$ e coeficiente de variação de 107,03% (Tabela 1). O alto coeficiente de variação e desvio padrão indicam que a média é pouco representativa, visto a alta variabilidade intra-anual (Figura 2), indicando característica sazonal da série (Bayer et al., 2012).

Tabela 1 – Dados da análise descritiva dos dados

Medidas descritivas	Valor
Mínimo ($\text{m}^3 \text{ s}^{-1}$)	12,91
Máximo ($\text{m}^3 \text{ s}^{-1}$)	1334,51
Média ($\text{m}^3 \text{ s}^{-1}$)	179,91
Mediana ($\text{m}^3 \text{ s}^{-1}$)	97,30
Desvio Padrão ($\text{m}^3 \text{ s}^{-1}$)	192,56
Coeficiente de variação (%)	107,03
Assimetria	2,04
Curtose	5,45

A menor vazão foi registrada no mês de setembro do ano 2017 ($12,91 \text{ m}^3 \text{ s}^{-1}$) e a maior no mês de fevereiro de 1980 ($1334,51 \text{ m}^3 \text{ s}^{-1}$). A distribuição da série possui assimetria positiva de 2,04 e curtose de 5,45, indicando uma distribuição leptocúrtica, ou seja, possui cauda pesada. As vazões possuem

uma sucessão regular de picos e recessões, flutuações sazonais, o que indica uma não estacionariedade da variância. As maiores vazões foram registradas de dezembro a abril e as menores de maio a novembro (Figura 3). Mediante o teste de Kruskal-Wallis confirmou-se a não estacionariedade na variância, sendo caracterizada por sazonalidades.

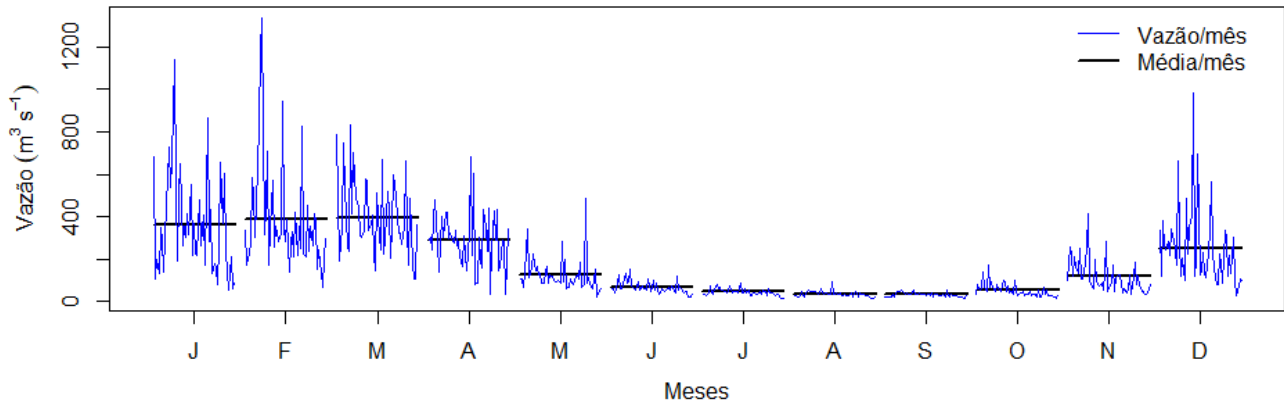


Figura 3 - Sazonalidade da série de vazões médias.

A representação das vazões médias mensais em gráficos boxplot (Figura 4) revela a grande quantidade de valores discrepantes que a série possui, também chamados de outliers, indicando possíveis erros de coleta das vazões ou vazões influenciadas por eventos externos extremos (Valladares Neto et

al., 2017). Observa-se no mês de fevereiro a maior ocorrência de valores atípicos, seguido dos meses de maio, novembro e abril, enquanto que os meses de março e setembro não apresentaram valores discrepantes.

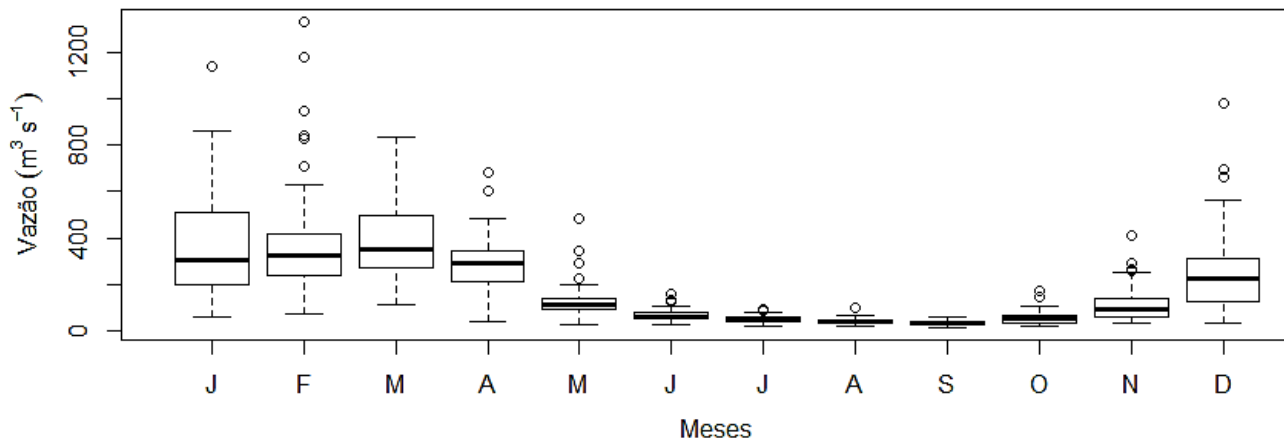


Figura 4 - Boxplot da série de vazões médias.

De acordo com os resultados encontrados por Reisen et al. (2008), após a modelagem de séries temporais sazonais de vazão máxima do rio Jucu, a modelagem de séries com presença de valores atípicos forneceu previsões 46% mais precisas quando comparadas às séries tratadas para retirada dos outliers. Tais resultados evidenciam a importância da presença dos valores discrepantes como

componentes da série, se estes não são decorrentes de falhas ou erros.

Na decomposição das séries em tendência, sazonalidade e componentes aleatórios (Figura 5) nota-se que os valores discrepantes da série não foram classificados como tendência e sazonalidade, restando seus valores aos componentes aleatórios.

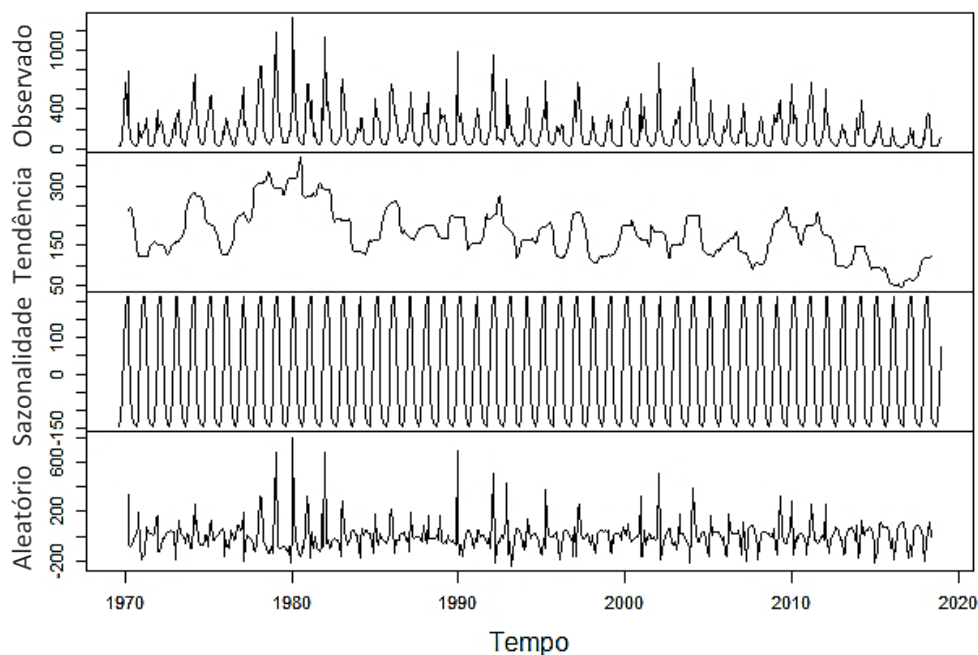


Figura 5 - Série decomposta em tendência, sazonalidade e componentes aleatórios.

A identificação dos modelos foi realizada analisando os gráficos das funções de autocorrelação e autocorrelação parcial (Figura 6), e das funções das séries diferenciadas de ordem 12 (Figura 7). Os correlogramas apontam um possível comportamento

autoregressivo da série com correlação serial de ordens próximas de 1 e 2. Os correlogramas da série diferenciada de ordem 12 mostram características de um modelo estacionário, autorregressivo e de médias móveis de ordem próxima de 1. Ambas as análises demonstrando um período sazonal.

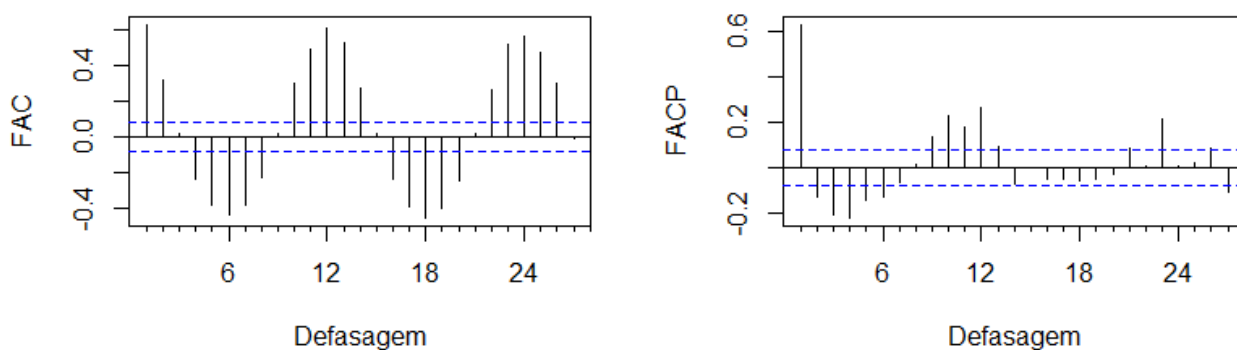


Figura 6 - Função de autocorrelação (FAC) à esquerda, e função de autocorrelação parcial (FACP) da série de vazões médias mensais, à direita.

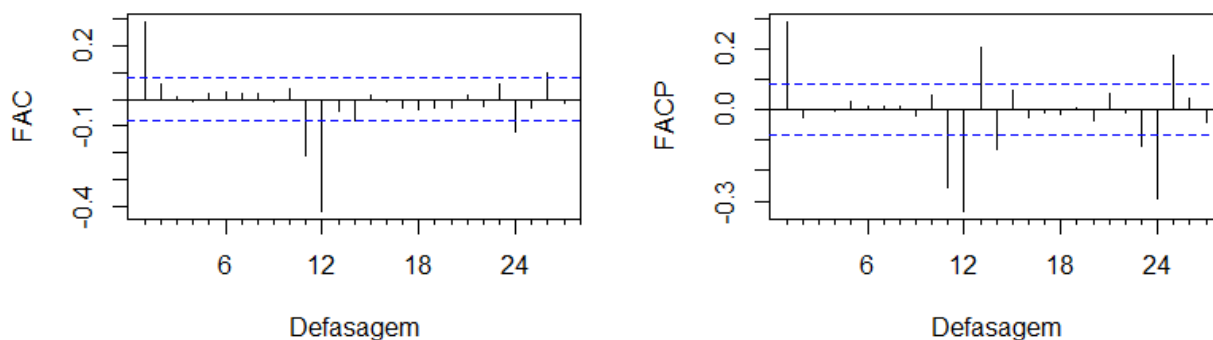


Figura 7 - Função de autocorrelação (FAC), à esquerda, e função de autocorrelação parcial (FACP) da série de vazões mensais diferenciadas de ordem 12, à direita.

O período sazonal foi observado por meio do periodograma (Figura 8), que apresentou a maior ordenada na frequência 0,0833333, implicando em uma periodicidade de $s = 1/0,0833333 = 12$ meses.

A presença de sazonalidade indica que o modelo deve considerar uma diferença sazonal, tornando-se SARIMA.

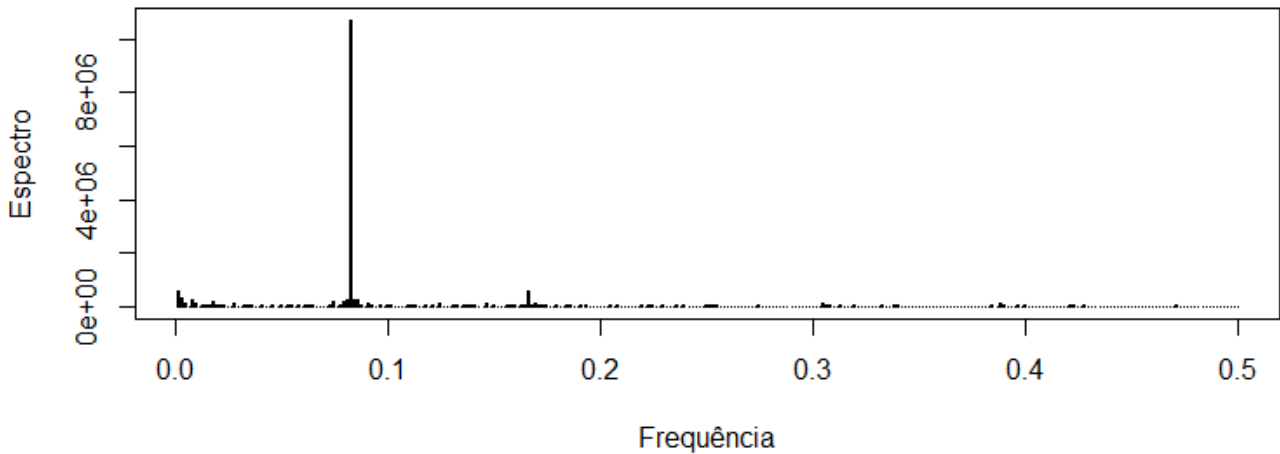


Figura 8 - Periodograma da série de vazões mensais

Os valores autorregressivos e de médias móveis foram verificados e modelos com combinações destes valores foram testados e avaliados quanto

aos critérios de informação e medidas de qualidade (Tabela 2).

Tabela 2 - Modelos SARIMA e seus valores de AIC, BIC e medidas de qualidade

(p,d,q)(P,D,Q) _s	Critérios de informação		Medidas de qualidade	
	AIC	BIC	RQME	C _{NS}
(1,0,1)(1,1,4) ₁₂	7114,57	7149,30	121,21	0,61
(1,0,1)(1,1,3) ₁₂	7122,62	7153,01	122,56	0,60
(1,0,1)(2,1,1) ₁₂	7125,42	7151,47	123,08	0,59
(3,0,2)(2,1,0) ₁₂	7175,95	7210,69	129,64	0,55
(1,0,2)(1,1,1) ₁₂	7125,54	7151,59	123,06	0,59
(1,0,2)(1,1,2) ₁₂	7122,62	7153,01	122,55	0,60
(2,0,1)(2,1,1) ₁₂	7127,55	7157,95	123,11	0,59

Os resultados encontrados na Tabela 2 demonstram que o modelo SARIMA(1,0,1)(1,1,4)₁₂ foi o que melhor se ajustou a série, com o menor valor de AIC (7114,57), BIC (7149,30) e maior C_{NS} (0,61). O resultado do RQME (121,21) foi elevado devido a este parâmetro sofrer grande influência dos picos de vazão e valores discrepantes, não sendo possível comparar seus resultados com os obtidos para outras séries temporais (Caldeira, 2016). Analisando o C_{NS}, observa-se que todos os modelos são considerados aptos para a etapa de previsão (C_{NS} > 0,5) conforme Zappa (2002). O modelo SARIMA(1,0,1)(1,1,4)₁₂ pode ser caracterizado como bom, de acordo com a classificação de

Moriasi et al. (2007).

O gráfico quantil-quantil e correlograma mostraram que os resíduos do modelo SARIMA(1,0,1)(1,1,4)₁₂ não estão distribuídos de forma normal e que não há correlação estatisticamente significativa entre os resíduos (Figura 9). A não normalidade dos resíduos, confirmada por meio do teste de Shapiro-Wilk ($p = 2,2 \times 10^{-16}$), é aceitável visto que a série é estocástica e desta maneira sensível a mudanças mediante estímulos externos e possui valores discrepantes. A não correlação significativa entre os resíduos foi confirmada por meio do teste Ljung-Box ($p = 0,12$).

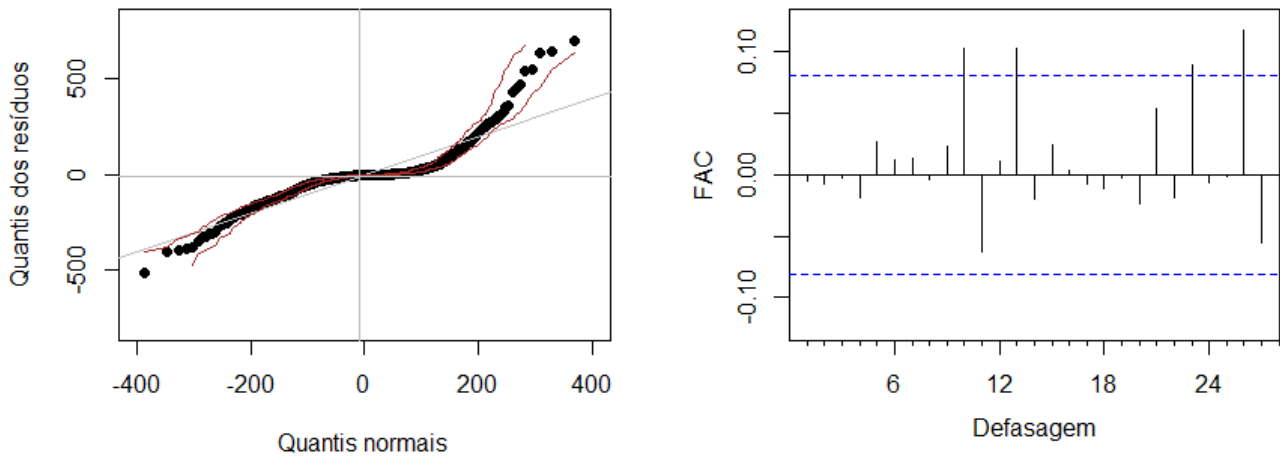


Figura 9 - Gráfico quantil-quantil, à esquerda, e correlograma, à direita, dos resíduos do modelo SARIMA(1,0,1)(1,1,4)₁₂

A média dos resíduos foi de $-9,44 \text{ m}^3 \text{ s}^{-1}$, distante de zero e da mediana, indicando que as previsões resultantes do modelo serão tendenciosas, diante disso foi adicionado o valor médio dos resíduos encontrado aos resultados das previsões, ajustando o viés conforme a metodologia proposta por Hyndman & Athanasopoulos (2018).

Mediante a adequação dos resíduos aos principais pressupostos estabelecidos o modelo

SARIMA(1,0,1)(1,1,4)₁₂ foi o escolhido para previsão das vazões mensais na BH_{MAN}. As estimativas dos parâmetros do modelo, assim como o erro padrão de cada estimativa estão presentes na Tabela 3. O modelo escolhido apresenta um parâmetro autorregressivo estimado ($\hat{\phi}_1$) que mostra que a vazão Q_t é influenciada em 38,87% pela vazão do mês anterior Q_{t-1} .

Tabela 3 - Estimativas e erro padrão do modelo escolhido

Modelo	Parâmetro	Estimativa	Erro padrão
SARIMA(1,0,1)(1,1,4) ₁₂	$\hat{\phi}_1$	0,3887	0,1114
	$\hat{\theta}_1$	-0,0368	0,1198
	$\hat{\Phi}_1$	0,1416	0,2084
	$\hat{\Theta}_1$	-0,9519	0,2066
	$\hat{\Theta}_2$	-0,0118	0,1700
	$\hat{\Theta}_3$	0,1724	0,0616
	$\hat{\Theta}_4$	-0,1613	0,0430

As previsões realizadas para um horizonte de 12 meses à frente resultaram em um RQME de 55,41 e C_{NS} de 0,81. O valor de RQME diminuiu consideravelmente na etapa de validação e o C_{NS} é considerado adequado e bom ($C_{NS} > 0,75$) segundo a classificação de Gotschalk et al. (1999). Os resultados para simulação e previsão do escoamento superficial na bacia são satisfatórios, visto que foi utilizada apenas a correlação serial das vazões médias mensais.

Os resultados das previsões mostram uma vazão média para o ano de 2018 de $134,59 \text{ m}^3 \text{ s}^{-1}$, uma máxima de $289,18 \text{ m}^3 \text{ s}^{-1}$ e mínima de $26,46 \text{ m}^3 \text{ s}^{-1}$, o que equivale a erros da ordem de X%, Y% e Z%, respectivamente. Observa-se desta maneira subestimativa da vazão máxima, denotando que a

aplicação do modelo para previsão de vazões máximas deve ser feita com cautela. De acordo com Caldeira (2016), prever altas vazões é importante para o dimensionamento de projetos hidráulicos, conservação do solo e gestão de cheias.

Comparando os resultados obtidos neste trabalho aos encontrados por outros autores observa-se que de maneira geral os resultados foram condizentes. Bayer et al. (2012) avaliando modelos de séries temporais para a bacia hidrográfica do rio Potiribu escolheram como melhor modelo SARIMA(3,0,0)(2,1,2)₁₂, que apresentou um C_{NS} de 0,68 e 0,81 para as etapas de calibração e validação, respectivamente. Figueiredo & Blanco (2014) simulando vazões e níveis de água médios mensais para o rio Tapajós,

PA, escolheram o modelo SARIMA(1,0,0)(1,1,1)₁₂, que apresentou C_{NS} médio de 0,92 para etapa de calibração e 0,91 para a validação.

O coeficiente de determinação para o período de calibração foi de 0,61 e de 0,82 para validação, significando que 82% da variação das vazões foi prevista pelo modelo na fase de validação. Dhote et al. (2018) realizando a previsão da vazão mensal da bacia hidrográfica do rio Narmada, Índia, após processamento dos dados no Hadoop data cluster, HDFS e MapReduce, encontraram como melhor conformação o SARIMA(1,1,1)(1,1,1)₁₂ que apresentou um R² de 0,95 e 0,99 para as etapas de calibração e validação, respectivamente.

Machekposhti et al. (2018) utilizaram modelos ARIMA para análise da vazão máxima anual e previsão de inundações no rio Karkheh, Irã, e obtiveram como resultado da previsão da vazão máxima anual por meio do modelo ARIMA(4,1,1) um R² de 0,84. Mishra et al. (2018) investigando modelos para previsão do escoamento superficial da bacia hidrográfica do rio Brahmaputra, Índia, obtiveram para os modelos ARIMA um R² médio

de 0,92. Hao et al. (2017) objetivando a previsão do escoamento superficial e de sedimentos do rio Yangtze, China, obtiveram como melhor modelo para o escoamento superficial o modelo SARIMA(1,1,1)(1,1,1)₁₂, com um R² de 0,93.

A análise visual dos dados observados, estimados e previstos pelo modelo (Figura 10) indica um bom desempenho nos períodos de calibração e previsão da vazão mensal. Nota-se que o modelo SARIMA foi eficaz principalmente nos períodos de recessão, com boa aderência às vazões mínimas observadas. De acordo com Pereira et al. (2016), a previsão de vazões mínimas subsidia o dimensionamento de sistemas de bombeamento para irrigação, a previsão de estiagens críticas e a determinação de vazões outorgáveis, sendo esta a principal finalidade identificada para o modelo SARIMA ajustado para a BH_{MAN} no presente estudo. Cabe ressaltar que conforme a Figura 10, nota-se dificuldade em prever as vazões máximas, o que pode ser justificado segundo Viola et al. (2009) pela maior complexidade da gênese do escoamento superficial direto.

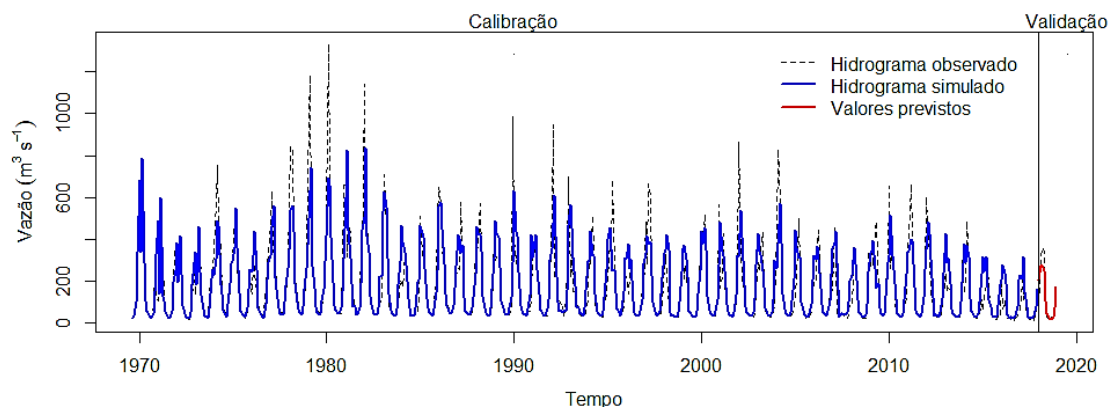


Figura 10 - Valores observados e estimados de vazão média mensal pelo modelo SARIMA(1,0,1)(1,1,4)₁₂

CONCLUSÕES

Neste trabalho foi realizada a análise da série temporal de vazões mensais da bacia hidrográfica do rio Manuel Alves da Natividade e ajuste de um modelo SARIMA para previsão. Mediante a análise da série constatou-se uma tendência de decréscimo das vazões, que a série possui sazonalidade anual e que é estacionária. O modelo escolhido foi o SARIMA(1,0,1)(1,1,4)₁₂, que apresentou bons resultados para as etapas de calibração e validação, com coeficientes de Nash-Sutcliffe iguais à 0,61 e 0,81, respectivamente. O modelo mostrou-se eficaz na previsão de períodos de recessão, sendo indicado principalmente como ferramenta na gestão de recursos hídricos para a previsão de vazões mínimas.

REFERÊNCIAS BIBLIOGRÁFICAS

- AKAIKE, H. A bayesian analysis of the minimum AIC procedure. *Ann Inst Stat Math*, v.30, n.1, p.9–14, 1978.
- AKAIKE, H. A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, [s.l.], v.19, n.6, p.716-723, 1974.
- ALMEIDA, L.; SERRA, J. C.V. Modelos hidrológicos, tipos e aplicações mais utilizadas. *Revista da FAE*, v.20, n.1, p.129-137, 2017.
- ANA. Agência Nacional de Águas. Ministério do Meio Ambiente (Org.). *Conjuntura dos recursos hídricos no Brasil 2018: informe anual*. Brasília: ANA, 72p. 2018.
- BATISTA, A.L.F. Modelos de séries temporais e redes neurais artificiais na previsão de vazão. Ano de obtenção:

2009. 79p. Dissertação (Mestrado em engenharia de sistemas) - Universidade Federal de Lavras, Lavras, Minas Gerais.
- BAYER, D.M.; CASTRO, N.M.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, p.229-239, 2012.
- BAYER, F.M., SOUZA, A.M. Wavelets e modelos tradicionais de previsão: um estudo comparativo. *Revista Brasileira de Biometria*, v.28, n.2, 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 Jucu, ES, utilizando o modelo SARIMA. *IRRIGA*, v.24, n.2, p.320-335, 2019.
- BOX, G.E.; JENKINS, G.M.; REINSEL, G.C.; LJUNG, G.M. *Time series analysis: forecasting and control*. John Wiley & Sons, 2008.
- CALDEIRA, T.L. Aprimoramento computacional do modelo Lavras Simulation of Hydrology (LASH). Ano de obtenção: 2016. 213 p. Dissertação (Mestrado em Recursos Hídricos) - Universidade Federal de Pelotas, Pelotas.
- CHECHI, L.; BAYER, F.M. Modelos univariados de séries temporais para previsão das temperaturas médias mensais de Erechim, RS. *Revista Brasileira de Engenharia Agrícola e Ambiental-Agriambi*, v.16, n.12, 2012.
- DHOTE, V.; MISHRA, S.; SHUKLA, J. P.; PANDEY, S.K. Runoff prediction using big data analytics based on ARIMA model. *Indian Journal of Geo Marine Sciences*, v.47, n.11, p.2163-2170, 2018.
- DICKEY, D.A.; FULLER, W.A. Distribution of the Estimators for Autoregressive Time Series with a Unit Root. *Journal of the American Statistical Association*, v.74, n.366a, p.427-431, 1979.
- FIGUEIREDO, N.M., BLANCO, C.J.C. Simulação de vazão e níveis de água médio mensais para o rio Tapajó usando modelos ARIMA. *Revista Brasileira de Recursos Hídricos*, v.19, n.3, p.111-126, 2014.
- GÓMEZ, V.; MARAVALL, A. Estimation, prediction, and interpolation for nonstationary series with the Kalman filter. *Journal of the American Statistical Association*, v. 89, n.426, p.611-624, 1994.
- GOTSCHALK, L.; BATCHVAROVA, E.; GRYNING, S. E.; LINDROTH, A.; MELAS, D.; MOTOVILOV, Y. U. G.; FREEH, M.; HEIKINHEIMO, M.; SAMUELSSON, P.; GRELE, A.; PERSSON, T. Scale aggregation: Comparison of flux estimates from NOPEX. *Journal of Agricultural and Forest Meteorology*, v.98-99, p.103-120, 1999.
- HAO, C.-F.; QIU, J.; LI, F.-F. Methodology for Analyzing and Predicting the Runoff and Sediment into a Reservoir. *Water*, v.9, n.6, p.440, 2017.
- HYNDMAN, R. J.; ATHANASOPOULOS, G. *Forecasting: principles and practice*, 2nd edition, Melbourne: OTexts. 2018.
- Instituto Nacional de Pesquisas Espaciais - INPE. Disponível em: <<http://www.dsr.inpe.br/topodata/>>. Acesso em: 20 de setembro de 2012.
- LJUNG, G. M.; BOX, G. E. P. On a measure of lack of fit in time series models. *Biometrika*. v.65, n.2, p.297-303, 1978.
- MACHEKPOSHTI, K.H.; WEDGHI, H.; TELVARI, A.; BABAZADEH, H. Flood Predicting in Karkheh River Basin Using Stochastic ARIMA Model. *International Journal of Agricultural and Biosystems Engineering*, v.12, n.3, p.89-96, 2018.
- MIRANDA, N. M.; CATALDI, M.; SILVA, F. N. R. Simulação do regime hidrológico da cabeceira do rio São Francisco a partir da utilização dos modelos SMAP e RegCM. *Anuário do Instituto de Geociências*, v.40, n.3, 2017.
- MISHRA S.; SARAVANAN, C.; DWIVEDI, V.K.; SHUKLA, J.P. Rainfall-runoff modeling using clustering and regression analyses for the river Brahmaputra basin. *Journal Geological Society of India*, v.92, n.3, p.305-312, 2018.
- MORETTIN, P.; TOLOI, C. *Análise de Séries Temporais*. São Paulo: ABE - Projeto Fisher – Blucher, 2006.
- MORIASI, D.N.; ARNOLD, J.G.; LIEW, M.W. VAN; BINGER, R.L.; HARMEL, R.D.; VEITH, T. Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. *Transactions of the ASABE*, v.50, n.3, p.885-900, 2007.
- PEREIRA, D.R.; ULIANA, E.M.; MARTINEZ, M.A.; SILVA, D.D. Desempenho de um modelo hidrológico concentrado e de um semidistribuído na predição de vazões diárias. *Irriga*, v.21, n.2, p.409-409, 2016.
- R Development Core Team. *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria, 2014. URL <http://www.R-project.org>, ISBN 3-900051-07-0.
- REISEN, V.A., MOLINARES, F.A.F., TEIXEIRA, E.C. Modelagem de séries temporais sazonais na presença de outliers: estudo de caso da vazão máxima mensal do rio Jucu, ES, Brasil. *Revista Brasileira de Recursos Hídricos*, v. 13, n. 2, p. 45-53, 2008.
- RODRIGUES, J. A. M.; ANDRADE, A. C. O.; VIOLA, M. R.; MORAIS, M. A. V. Indicadores hidrológicos para a gestão de recursos hídricos na bacia hidrográfica do rio Manuel Alves da Natividade, Tocantins. *Scientia Agraria*, v.16, n.4, 2015.
- SANTOS, P.G.P.; OLIVEIRA, T.F. Modelagem e previsão de vazão afluente média mensal no Rio Tocantins, Usina Hidrelétrica Tucuruí-Pará, Amazônia, Brasil. *Biota Amazônia*, v.6, n.2, p.9-16, 2016.
- SHAPIRO, S.S.; WILK, M.B. An analysis of variance test for normality (complete samples). *Biometrika*, v.52, p.591-611, 1965.
- TEODORO, V.A. Comparação de método de imputação para dados de precipitação diária. Ano de obtenção: 2019. 85 p. Tese (Doutorado) - Escola Superior de Agricultura “Luiz de Queiroz”. Piracicaba.
- TUCCI, C.E.M. *Modelos hidrológicos*. Porto Alegre: Editora da UFRGS, 1998.
- ULIANA, E.M.; PEREIRA, D.R.; SILVA, D.D.; ALMEIDA, F.T.; SOUZA, A.P. Modelagem chuva-vazão em uma bacia tropical utilizando o modelo IPH II. *Geo Uerj*, n.33, 2018.

- UNESCO. United Nations Educational Scientific and Cultural Organization. The united nations world water development report: Managing Water under Uncertainty and Risk, 2015.
- VALLADARES NETO, J.; SANTOS, C.B.; TORRES, E.M.; ESTRELA, C. Boxplot: um recurso gráfico para a análise e interpretação de dados quantitativos. *Revista Odontológica do Brasil Central*, v.26, n.76, p.1-6, 2017.
- VIANA, J.F.S.; MONTENEGRO, S.M.G.L.; SILVA, B.B.; SILVA, R.M.; SOUSA, W.S. Modelagem hidrológica da bacia hidrográfica do rio Pirapama-PE utilizando o modelo SWAT. *Journal of Environmental Analysis and Progress*, v.3, n.1, p.155-172, 2018.
- VIOLA, M.R.; MELLO, C.R.; ACERBI JUNIOR, F.W.; SILVA, A.M. Modelagem hidrológica na bacia hidrográfica do rio Aiuruoca, MG. *Revista Brasileira de Engenharia Agrícola e Ambiental*, v.13, n.5, p.581-590, 2009.
- WEI, W. *Time Series Analysis: Univariate and Multivariate Methods*. New York: Addison Wesley, 2006.
- XUE, L.; YANG, F.; YANG, C.; WEI, G.; LI, W.; HE, X. Hydrological simulation and uncertainty analysis using the improved TOPMODEL in the arid Manas River basin, China. *Scientific reports*, v.8, n.1, p.452, 2018.
- ZAPPA, M. Multiple-response verification of a distributed hydrological model at different spatial scales. Ano de obtenção: 2002. 157p. Tese (Pós-doutorado). Swiss Federal Institute of Technology. Switzerland.
- ZHANG, Q.; WANG, B-D.; HE, B.; PENG, Y.; REN, M-L. Singular spectrum analysis and ARIMA hybrid model for annual runoff forecasting. *Water Resour Manage*, v. 25, n. 11, p. 2683-2703, 2011.
- ZHAO, X.; CHEN, X.; XU, Y.; XI, D.; ZHANG, Y.; ZHENG, X. An EMD-based chaotic least squares support vector machine hybrid model for annual runoff forecasting. *Water*, v.9, n.3, p.153, 2017.