

Análise do comportamento da vazão de água na usina hidrelétrica do Funil/MG para realizar previsões e auxiliar na tomada de decisões

Otávio Augusto S. Delfino^{1†}, Pedro Henrique A. B. Santos¹, Ricardo Vitor R. Santos¹

¹Centro Federal de Educação Tecnológica de Minas Gerais (CEFET-MG).

Resumo: *A instalação de barragens para a construção de usinas hidrelétricas no Brasil iniciou-se no século XIX. Esse é um modelo de geração de energia limpa e segura. Um dos fatores que mais influenciam a operação das usinas hidrelétricas é a vazão de água nos reservatórios. Existe um volume mínimo de água que deve estar represada para conseguir acionar a turbina e gerar energia elétrica. Prever os valores dessa vazão é de extrema importância para o planejamento dos sistemas hidrelétricos. Diante do exposto, este trabalho teve como objetivo compreender o comportamento da vazão de água na usina hidrelétrica do Funil/MG e fazer previsões para que decisões administrativas e de engenharia possam ser tomadas com segurança. Os dados utilizados foram obtidos junto ao Operador Nacional do Sistema Elétrico (ONS). As análises foram realizadas pelo software R e a metodologia utilizada para ajustar o modelo de séries temporais foi a de Box e Jenkins. A partir dos modelos adequados à série, foi utilizado o critério de informação de Akaike e o erro quadrático médio para selecionar o melhor modelo. O modelo ajustado para realizar as previsões foi o SARIMA (2, 0, 1) × (0, 1, 1)₁₂. Observou-se que as estimativas de previsão da vazão média mensal de água seguem um ciclo, entre meses de novembro e março, as vazões são elevadas (período de chuvas), entre meses de abril e junho, intermediárias, já entre meses de julho e outubro, as vazões são baixas (período de seca).*

Palavras-chave: vazão, previsões, séries temporais, SARIMA.

Abstract: *The installation of dams for the construction of hydroelectric plants in Brazil began in the 19th century. It is a model of clean and safe energy generation. One of the factors that most influence the operation of hydroelectric plants is the water flow. There is a minimum volume that must be dammed to start the turbine and generate electricity. Predict the values of that flow is extremely important for planning the hydroelectric systems. Thus, this study aimed to understand the behavior of the water flow in the Funil/MG hydroelectric plants and make predictions so that administrative and engineering decisions can be safely made. The data used in this work were obtained from Operador Nacional do Sistema Elétrico (ONS). The analysis was realized by software R and the methodology used to adjust the model time series was that of Box e Jenkins. From the models suggested and suitable the series have used the criteria of information of Akaike and the mean quadratic error for selecting the best model. The adjustment model for making the predictions was SARIMA (2, 0, 1) X (0, 1, 1)₁₂. It was observed that the estimates of predicts of the average monthly water flow follow a cycle, in-between months of November and March, the flow is elevated (rainy season), in-between months of April and June, intermediaries, in-between months of July and October, the flow are casualties (drought).*

Keywords: flow rate, predictions, time series, SARIMA.

[†]Autor correspondente: otavioaugusto.2503@gmail.com.

Introdução

A instalação de barragens para a construção de usinas hidrelétricas no Brasil iniciou-se a partir do final do século XIX. É um modelo de geração de energia limpa e segura, que demanda investimentos modestos em relação aos sistemas que predominam em outras nações. Apesar de o país apresentar o terceiro maior potencial hidráulico do mundo, importa parte da energia hidrelétrica que consome, isso porque a maior hidrelétrica das Américas, e segunda maior do mundo, a Usina de Itaipu, não é totalmente brasileira. Por se encontrar na divisa do país com o Paraguai, 50% da produção da usina pertence ao país vizinho que, na incapacidade de consumir esse montante, vende o excedente para os brasileiros.

A energia gerada por uma hidrelétrica é através das turbinas, que são movimentadas pela água que passa por elas, transformando energia mecânica em elétrica, e um dos principais problemas enfrentados é a variabilidade natural do regime hidrológico, decorrente da precipitação. Quando o volume de chuvas está dentro do normal para a época e para a região, a vazão de água é maior e as usinas hidrelétricas chegam ao nível previsto de produção de energia, suprimindo a necessidade do país. Quando ocorre escassez de chuva e o reservatório atinge um volume mínimo, a vazão de água tem que ser reduzida. O volume mínimo corresponde a menor quantidade de água que deve estar armazenada num reservatório para que a usina preserve a sua capacidade de geração. Se o volume de água armazenado no reservatório for inferior ao mínimo, a usina não será capaz de gerar energia, ou porque não existe uma queda mínima entre a superfície do reservatório e as turbinas, ou porque a cota da superfície de água do reservatório é inferior à cota de tomada d'água do reservatório. O volume mínimo é também chamado de volume morto do reservatório.

Planejar os sistemas de produção de energia elétrica determina uma política de operação para usinas hidrelétricas e termelétricas que atenda a demanda de energia elétrica de forma econômica e confiável. Em linhas gerais, a meta é utilizar da forma mais eficiente possível a geração de base hidráulica para substituir ao máximo a geração térmica, pois além de ter um custo elevado, causa impactos ambientais. As usinas hidrelétricas possuem características de produção não-lineares e suas vazões são variáveis aleatórias, levando assim a um problema estocástico não-linear. Em razão da complexidade do problema, é comum sua decomposição em uma cadeia de planejamento que inclui as etapas de longo, médio e curto prazos (BATISTA, 2009).

Devido a importância desse planejamento nos sistemas hidrelétricos, justifica-se a aplicação de técnicas estatísticas para o estudo das séries históricas da vazão de água. Nesse contexto, por se tratar de dados observados ao longo do tempo, insere-se o estudo sobre Séries Temporais. Sendo assim, este trabalho tem como objetivo compreender o comportamento da vazão de água na usina hidrelétrica do Funil/MG e fazer previsões para que decisões administrativas e de engenharia possam ser tomadas com segurança.

Material e Métodos

A base de dados utilizada neste trabalho está disponível na página do Operador Nacional do Sistema Elétrico (ONS), e consiste de uma série histórica com 262 observações de vazões médias mensais de água, em metros cúbicos por segundo (m^3/s), coletadas na usina hidrelétrica do Funil/MG, localizada no Rio Grande, na região sul de Minas Gerais, entre os municípios de Perdões e Lavras. Inicialmente, o estudo utilizou 252 observações,

período de jan/1999 a dez/2019, deixando as 10 últimas, meses de janeiro a outubro de 2020, reservadas para comparar com os valores preditos. Posteriormente, estas observações foram incluídas no estudo para realizar as previsões dos meses de nov/2020 a dez/2021.

Uma abordagem bastante utilizada para a construção de modelos paramétricos para séries temporais univariadas é conhecida como metodologia Box e Jenkins (1976), e para utilizá-la, é necessário que a série seja estacionária, isto é, que ela se desenvolva aleatoriamente no tempo em torno de uma média constante. No entanto, a maioria das séries apresenta componentes como tendência e/ou sazonalidade, que são uma forma de não-estacionariedade.

Neste contexto, realizou-se uma análise no gráfico da série original para verificar indícios da existência de componentes de tendência e/ou sazonalidade, e no gráfico da função de autocorrelação para examinar a estacionariedade da mesma.

Antes de se fazer qualquer inferência sobre as componentes da série, é necessário verificar se o modelo é aditivo, ou seja, a componente sazonal independe das outras componentes. Se as amplitudes sazonais variarem com a tendência, o modelo é dito multiplicativo, e nesse caso, a variância não é constante, sendo necessária uma transformação na escala dos dados. A mais comum delas é a transformação logarítmica. Um gráfico comumente utilizado para examinar a necessidade de transformação dos dados traz no eixo das abscissas as médias das observações de subconjuntos da série e no eixo das ordenadas, a amplitude desses subconjuntos. Quando a amplitude independe das médias, os pontos ficam espalhados ao redor de uma reta paralela ao eixo das abscissas. Um teste t para testar a hipótese nula de que o coeficiente de inclinação da reta é zero, indica a necessidade de transformação quando retornar um valor- p inferior a 0,05 (PAIVA et al., 2019).

Para verificar a existência da componente de tendência pode-se utilizar um dos seguintes testes: Cox-Stuart, Dickey Fuller Aumentado, Wald-Wolfowitz ou Mann-Kendall. Neste trabalho aplicou-se o teste de Cox-Stuart (ou teste do Sinal), o qual é baseado no coeficiente de relação de Spearman e no teste F para análise de variância. O procedimento para aplicar o teste é o seguinte: (i) agrupam-se as observações em pares (Z_1, Z_{c+1}) , (Z_2, Z_{c+2}) , (Z_{N-c}, Z_N) , em que $c = \frac{N}{2}$ se N , números de observações, for par e $c = \frac{N+1}{2}$ se N for ímpar; (ii) para cada par (Z_i, Z_{c+i}) , associa-se um sinal. Este sinal é positivo (+) se $Z_i < Z_{c+i}$ e negativo (-) se $Z_i > Z_{c+i}$. Descarta-se a comparação se $Z_i = Z_{c+i}$; (iii) testa-se a hipótese de nulidade. H_0 : Não existe tendência. A estatística do teste a ser utilizada é T , sendo que T é igual ao número de sinais positivos atribuídos no final da comparação entre Z_i e Z_{c+i} ; (iv) calcula-se o valor de t , que é dado por: $t = \frac{1}{2} (n + \omega_{\alpha/2} \sqrt{n})$ em que n é o número de vezes que não houve empate na comparação entre Z_i e Z_{c+i} , e $\omega_{\alpha/2}$ é o valor tabelado ao nível de significância α . Caso $n \leq 20$, o valor tabelado de $\omega_{\alpha/2}$ é dado pela distribuição binomial com $p = 0,5$ e, caso $n > 20$, utiliza-se a distribuição normal; (v) se $T \leq t$ ou $T \geq n - t$, rejeita-se H_0 , isto é, a série possui tendência (BATISTA, 2009).

Para verificar a presença da componente de sazonalidade pode-se aplicar um dos seguintes testes: Kruskal-Wallis, G de Fisher ou Friedman. Neste trabalho foi utilizado o teste de Kruskal-Wallis, sendo este, análogo ao teste F utilizado na ANOVA de um fator, exceto pelo fato de que ele não coloca nenhuma restrição sobre a comparação, enquanto na análise de variância as populações em confronto são independentes e normalmente distribuídas. Os passos para realização deste teste são: (i) estabelecem-se as hipóteses: $H_0 : \tau_1 = \tau_2 = \dots = \tau_k$ e $H_1 : \tau_1, \tau_2, \dots, \tau_k$ não são todas iguais; (ii) ordenam-se de forma crescente de magnitude os valores deste novo conjunto de dados e associa-se a cada valor seu posto correspondente, tendo cada posto o mesmo sinal do valor que este representa;

(iii) calcula-se o valor da estatística H . Em seguida, fixa-se o nível de significância α ; (iv) encontram-se os valores críticos referentes ao nível de significância fixado. Neste caso, calcula-se o valor Q_α de modo que $P[H > Q_\alpha] = \alpha$ (sob H_0); (v) se $H_{obs} > Q_\alpha$ rejeita-se a hipótese nula de que as amostras provêm de populações igualmente distribuídas, isto é, a série possui sazonalidade determinística; (vi) O valor-p é calculado da seguinte forma $P_{valor} = P[\chi_{k-1}^2 \geq H|H_0]$ (MORETTIN; TOLOI, 2006).

A existência de sazonalidade também foi verificada pelo periodograma da análise espectral que consiste na decomposição da série temporal em uma série de Fourier. No eixo das ordenadas ficam localizadas as frequências da série (f_i) e no eixo das abscissas as respectivas intensidades de cada frequência $I_p(f_i)$, definidas como:

$$I_p(f_i) = \frac{2}{\left(\frac{n}{2}\right)} \left(\sum_{t=1}^n \epsilon_t \cos\left[\frac{2\pi i}{\left(\frac{n}{2}\right)}t\right]^2 \sum_{t=1}^n \epsilon_t \sen\left[\frac{2\pi i}{\left(\frac{n}{2}\right)}t\right]^2 \right),$$

em que ϵ_t é o componente estocástico da série temporal associado ao tempo t e n é o número de observações da série. Para determinar a periodicidade sazonal (s) basta verificar que frequência f_i está associado o maior valor de intensidade $I_p(f_i)$ e então dividir 1 por este valor de frequência (BARBOSA et al., 2015).

Para tornar uma série estacionária, um procedimento comum é tomar diferenças sucessivas da série original. Para eliminar a componente de tendência, tome a primeira diferença, definida por $\Delta Z_t = Z_t - Z_{t-1}$ com $t = 1, 2, \dots, n$. Geralmente, duas diferenças são suficientes. A sazonalidade determinística pode ser eliminada por meio de uma diferença realizada na série através do método de médias móveis, levando-se em conta a ordem do período s .

Depois de eliminadas as componentes de tendência e/ou sazonalidade, a série se encontra estacionária, parte-se então para a identificação do modelo. Para sugerir um modelo é necessário analisar as funções de autocorrelação (ACF) e autocorrelação parcial (PACF) da série diferenciada. As correlações significativas da ACF indicam a ordem do modelo de médias móveis (MA) e as da PACF indicam a ordem do modelo autorregressivo (AR). Caso haja correlações significativas nos lags múltiplos de s (período sazonal), deve-se ajustar a parte sazonal do modelo.

Box e Jenkins (1976) propõem o modelo sazonal multiplicativo de período s , denotado por SARIMA $(p, d, q) \times (P, D, Q)_s$ ou ARIMA Sazonal, representado por:

$$\phi(B)\Phi(B^s)\Delta^d\Delta_s^D Z_t = \theta(B)\Theta(B^s)a_t,$$

em que

$\phi(B) = 1 - \phi_1 B - \dots - \phi_p B^p$ é o polinômio autorregressivo de ordem p ;

$\Phi(B^s) = 1 - \Phi_1 B^s - \dots - \Phi_P B^{Ps}$ é o polinômio autorregressivo sazonal de ordem P ;

$\Delta^d = (1 - B)^d$ é o operador diferença e d é o número de diferenças necessárias para retirar a tendência da série;

$\Delta_s^D = (1 - B^s)^D$ é o operador diferença generalizado, quando duas observações estão distantes entre si de s intervalos de tempos que apresentam alguma semelhança, e D é o número de diferenças de lags s necessárias para retirar a sazonalidade da série;

$\theta(B) = 1 - \theta_1 B - \dots - \theta_q B^q$ é o polinômio médias móveis de ordem q ;

$\Theta(B^s) = 1 - \Theta_1 B^s - \dots - \Theta_Q B^{Qs}$ é o polinômio médias móveis sazonal de ordem Q ;

Z_t é a variável resposta do modelo;

a_t é o ruído aleatório no tempo t .

Segundo Morettin e Tolo (2006), a identificação do modelo é a fase mais crítica do ciclo iterativo na metodologia Box e Jenkins (1976), pois essa escolha é baseada, principalmente, nas funções de autocorrelação e autocorrelação parcial estimadas, nesse sentido, é importante salientar que pode ser identificado mais de um modelo para a mesma série. Tendo-se identificado o modelo provisório para a série, o passo seguinte é estimar os parâmetros através do método de máxima verossimilhança.

Após a fase de estimação, o próximo passo é fazer o diagnóstico da adequação do modelo ajustado, isto é, verificar se os resíduos são independentes e normalmente distribuídos, podendo então ser denominados de ruído branco. Para avaliar a presença de autocorrelação serial utilizou-se o teste estatístico Ljung-Box (LJUNG; BOX, 1978). Se o modelo ajustado para um processo SARIMA $(p, d, q) \times (P, D, Q)_s$ é apropriado, então

$$Q(K) = n(n+2) \sum_{j=1}^K \frac{\hat{r}_j^2}{(n-j)} \sim \chi_{K-p-q-P-Q}^2.$$

A hipótese de ruído branco é rejeitada para $Q(K) > \chi_{K-p-q-P-Q}^2$.

Os modelos ajustados em que se obtém um resíduo ruído branco são comparados usando o Critério de Informação de Akaike (AIC). Akaike (1973) sugere escolher como melhor modelo, aquele cujas ordens p e q minimizam o critério de Akaike (AIC) normalizado. Na comparação de diversos modelos, com N fixo, o critério de Akaike pode ser expresso por:

$$AIC(k, l) = N \ln \hat{\sigma}_a^2 + 2(k + l + 2)$$

em que $\hat{\sigma}_a^2$ é o estimador de máxima verossimilhança de σ_a^2 , $0 \leq k \leq p$ e $0 \leq l \leq q$. Onde k e l correspondem ao número de parâmetros que minimizam o critério, independente do modelo ser completo ou incompleto. Entre os vários modelos propostos, será escolhido como o melhor modelo aquele que tiver o menor valor de AIC. Além do AIC, existem outros critérios de informação, tais como: Bayesiano, Hanna-Quinn e Schwarz.

Com o modelo selecionado e ajustado, pode-se utilizá-lo para prever valores futuros. Morettin e Tolo (2006) salientam que a previsão não constitui um fim em si, mas apenas um meio de fornecer informações para uma consequente tomada de decisões, visando a determinar um objetivo. A previsão Z_{t+h} denotada por $\hat{Z}_t(h)$ é a esperança matemática condicional de Z_{t+h} , dado Z_t, Z_{t-1}, \dots , isto é,

$$\hat{Z}_t(h) = [Z_{t+h}] = E[Z_{t+h} | Z_t, Z_{t-1}, \dots].$$

Se o propósito for o de fazer previsões, pode-se utilizar o erro quadrático médio de previsão (EQMP), critério este que fornece as melhores previsões, e é uma boa medida de comparação entre modelos. As estimativas EQMP são dadas pela média dos quadrados das diferenças entre valores observados e previstos. Logo, o EQMP com origem em t é dado por

$$EQMP_t = \frac{1}{n} \sum_{h=1}^n [Z_{t+h} - \hat{Z}_t(h)]^2.$$

O modelo escolhido é o que possuir menor erro quadrático médio.

Utilizou-se o software livre R (R CORE TEAM, 2021) para manipulação e análise estatística dos dados, sendo o *TSA* (CHAN; RIPLEY, 2018) e o *forecast* (HYNDMAN, 2008) os principais pacotes utilizados. As análises foram realizadas considerando um nível de significância $\alpha = 5\%$.

Resultados e Discussões

Na Figura 1 observa-se o gráfico da série temporal sobre as vazões médias mensais de água, em metros cúbicos por segundo (m^3/s), na usina hidrelétrica do Funil/MG, durante o período de janeiro de 1999 a dezembro de 2019.

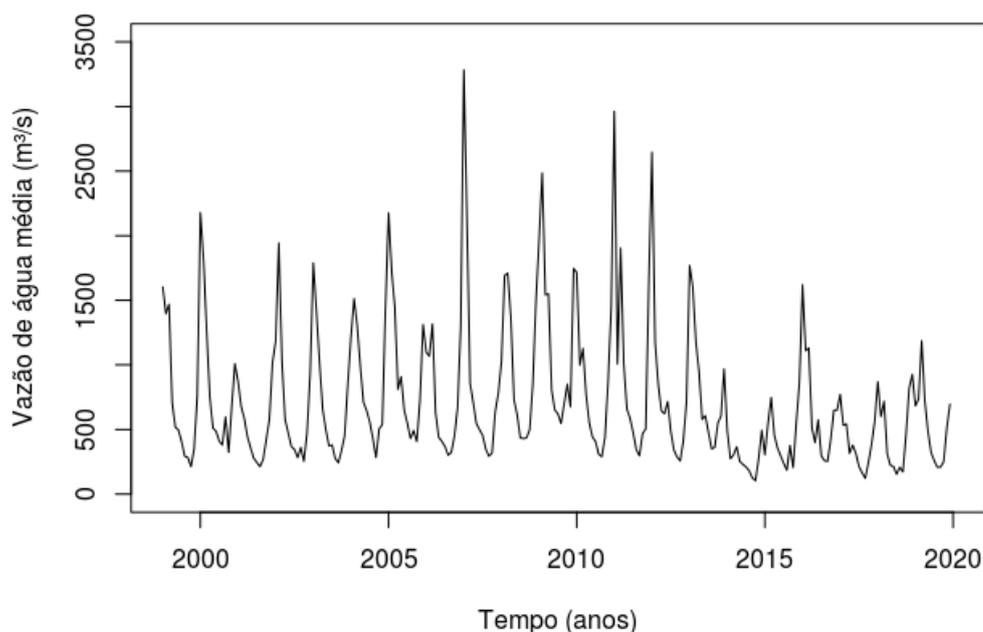


Figura 1: Série temporal das vazões médias mensais de água, em metros cúbicos por segundo (m^3/s), na usina hidrelétrica do Funil/MG, durante o período de janeiro de 1999 a dezembro de 2019.

A partir da inspeção visual preliminar pode-se sugerir que a série apresenta componente sazonal, e aparentemente, não possui tendência, mas é necessário confirmar tais informações através de testes matemático-estatísticos.

Na Figura 2 nota-se que os lags da função de autocorrelação (ACF) não decaem rapidamente para zero, indicando que a série não é estacionária. Percebe-se também que seu comportamento é senoidal, típico de uma série que contém sazonalidade (MORETTIN; TOLOI, 2006).

Para verificar a necessidade de transformação dos dados, aplicou-se o teste t , tendo como resultado um valor-p menor que 0,01, portanto, há evidências, ao nível de 5% de significância, para rejeitar a hipótese nula de que o coeficiente de inclinação da reta é igual a zero e, conseqüentemente, aplicar uma transformação na escala dos dados. Neste caso foi realizada a transformação logarítmica.

Pelo teste de Cox-Stuart, verificou-se a ausência da componente de tendência com um valor-p de 0,2467, não rejeitando, ao nível de 5% de significância, a hipótese nula de não existência de tendência. Sendo assim, não foi necessário realizar uma diferença de ordem 1 na série a fim de remover o efeito de tendência.

Ao aplicar o teste não paramétrico de Kruskal-Wallis o resultado do valor-p foi menor

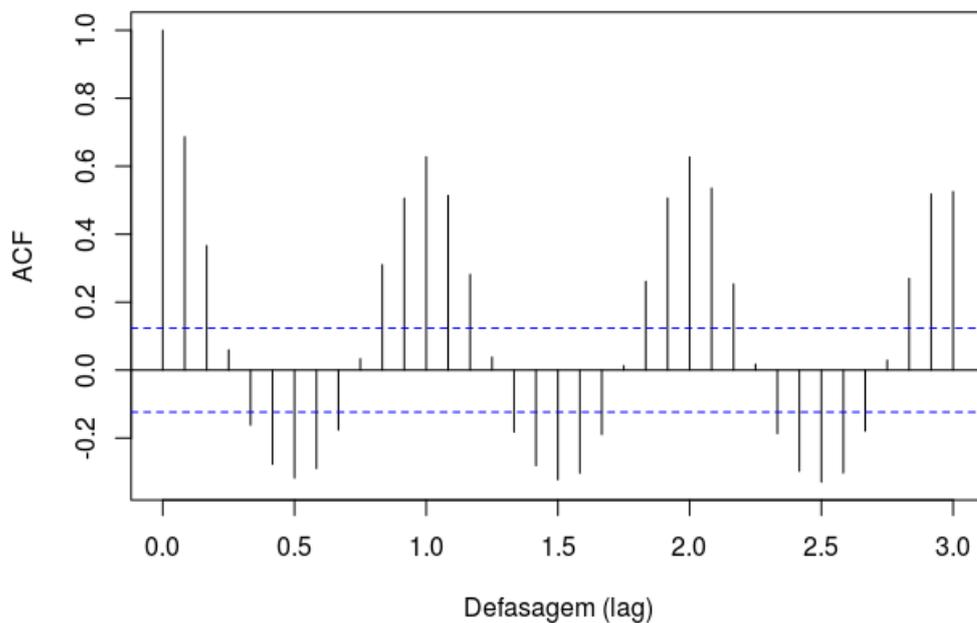


Figura 2: Função de autocorrelação (ACF) da série temporal das vazões médias mensais de água na usina hidrelétrica do Funil/MG.

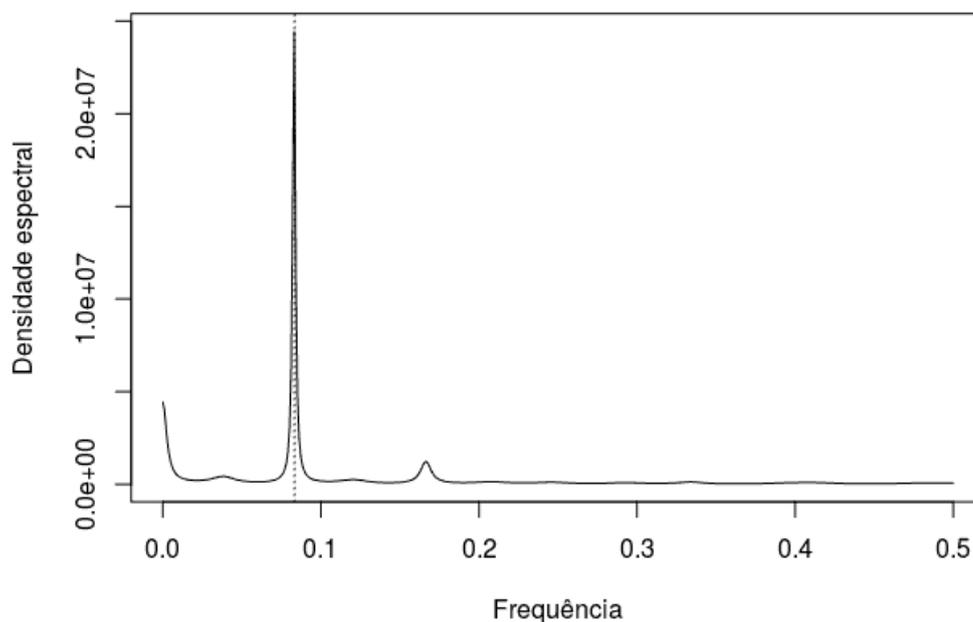


Figura 3: Periodograma da série das vazões médias mensais de água na usina hidrelétrica do Funil/MG.

que 0,01, dessa forma, rejeitou-se a hipótese nula de ausência de sazonalidade determinística, corroborando com a inspeção visual realizada na Figura 2.

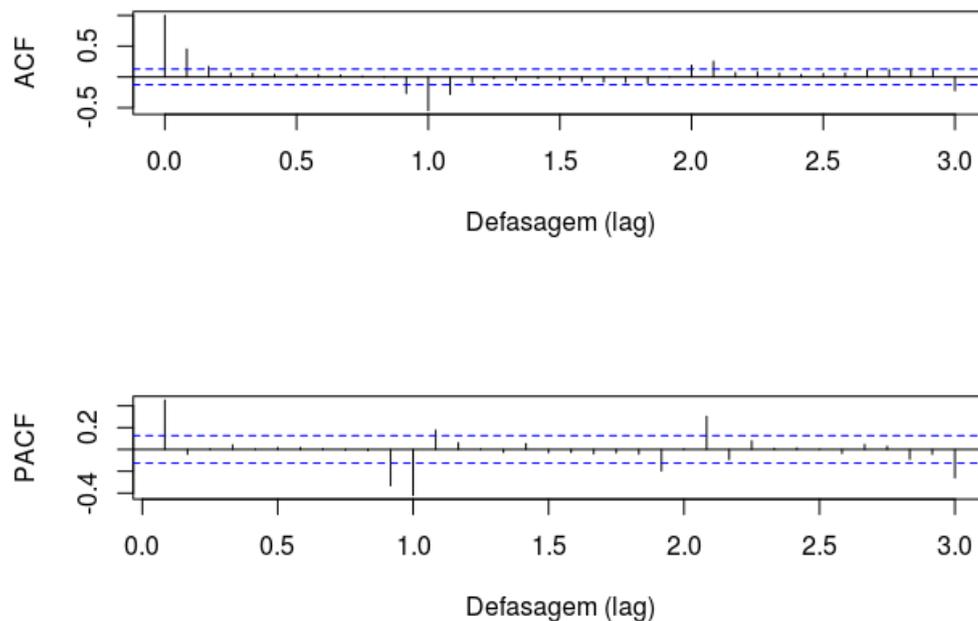


Figura 4: Gráfico das funções de autocorrelação (ACF) e autocorrelação parcial (PACF) após aplicação da diferença sazonal.

Tabela 1: Comparação entre modelos SARIMA através do critério de informação de Akaike (AIC), valor-p para o teste de Ljung-Box e erro quadrático médio (EQM).

Modelo	SARIMA ($s = 12$)	AIC	Ljung-Box (valor-p)	EQM
1	$(2, 0, 1) \times (0, 1, 1)$	113,62	0,492	256177,69
2	$(1, 0, 1) \times (0, 1, 1)$	115,05	0,447	262533,26
3	$(3, 0, 3) \times (0, 1, 1)$	117,12	0,679	267309,68
4	$(1, 0, 1) \times (1, 1, 2)$	118,61	0,467	265420,73
5	$(1, 0, 0) \times (0, 1, 1)$	119,28	0,114	261969,94

Para descobrir o padrão periódico da série e detectar sua frequência de ocorrência, foi confeccionado o gráfico do periodograma (ver Figura 3). Observa-se que a maior densidade espectral (pico) está associada à frequência de 0,08333, o que implica em uma componente sazonal com periodicidade de $s = 12$ meses.

Tabela 2: Estimativas para o modelo SARIMA $(2, 0, 1) \times (0, 1, 1)_{12}$ no período de janeiro de 1999 a dezembro de 2019.

Parâmetro	Coefficiente	Erro padrão	Valor-p
ϕ_1	1,3815	0,2050	0,0000*
ϕ_2	- 0,4130	0,1718	0,0162*
θ_1	- 0,7846	0,1713	0,0000*
Θ_1	- 0,8835	0,0536	0,0000*

*Significativo a 0,05 de probabilidade

Percebe-se na Figura 4 que não existem lags múltiplos de 12 significativos na ACF e PACF depois de aplicada a diferença sazonal, o que segundo Morettin e Tolo (2006)

assinala a ausência de sazonalidade estocástica.

Tabela 3: Valores das vazões médias mensais de água observadas, previstas e o intervalo de confiança a 95%.

Meses/Ano	Observado	Previsto	Intervalo de confiança
janeiro/2020	903	910	[513,1613]
fevereiro/2020	2279	768	[394,1494]
março/2020	1311	807	[398,1636]
abril/2020	651	516	[249,1072]
maio/2020	428	393	[186,828]
junho/2020	360	355	[165,758]
julho/2020	273	276	[127,598]
agosto/2020	242	235	[107,512]
setembro/2020	250	231	[104,508]
outubro/2020	258	298	[134,662]

A partir dos correlogramas da ACF e PACF (Figura 4), foi possível propor alguns modelos para a série estudada. Segundo Barbosa et al. (2015), obter a ordem exata do modelo através destes correlogramas é algo complexo e subjetivo, devido ao grande número de modelos que podem ser ajustados.

Tabela 4: Estimativas para o modelo SARIMA $(2, 0, 1) \times (0, 1, 1)_{12}$ no período de janeiro de 1999 a outubro de 2020.

Parâmetro	Coefficiente	Erro padrão	Valor-p
ϕ_1	1,4130	0,1777	0,0000*
ϕ_2	- 0,4370	0,1513	0,0038*
θ_1	- 0,8212	0,1437	0,0000*
Θ_1	- 0,9215	0,0673	0,0000*

*Significativo a 0,05 de probabilidade

Como a série não apresentou tendência, não houve a necessidade de aplicar diferenças de ordem 1, portanto $d = 0$, mas como uma diferença sazonal de ordem $s = 12$ foi aplicada para eliminar a sazonalidade, tem-se $D = 1$.

Dentre os modelos ajustados, foram selecionados os cinco que apresentaram menor AIC e tiveram valor-p superior a 5% (nível de significância) para o teste de Ljung-Box, o que implica em resíduos independentes e normalmente distribuídos, ou seja, ruído branco (ver Tabela 1). Destes, partiu-se então para o cálculo da previsão e do erro quadrático médio das vazões médias mensais de água para os meses de janeiro a outubro de 2020.

O modelo SARIMA $(2, 0, 1) \times (0, 1, 1)_{12}$ foi escolhido, pois apresentou menor valor para o critério de informação de Akaike (AIC) e para o erro quadrático médio (EQM) (ver Tabela 1). Os coeficientes estimados deste modelo estão na Tabela 2.

As previsões e os intervalos de confiança a 95% para o modelo SARIMA $(2, 0, 1) \times (0, 1, 1)_{12}$, bem como as vazões médias mensais observadas para o período de janeiro a outubro de 2020 estão na Tabela 3. Nota-se que a estimativa para o mês de fevereiro ficou bem distante do valor real, inclusive fora do intervalo de confiança a 95%, fato explicado pela grande quantidade de chuva atípica neste período. Segundo levantamento

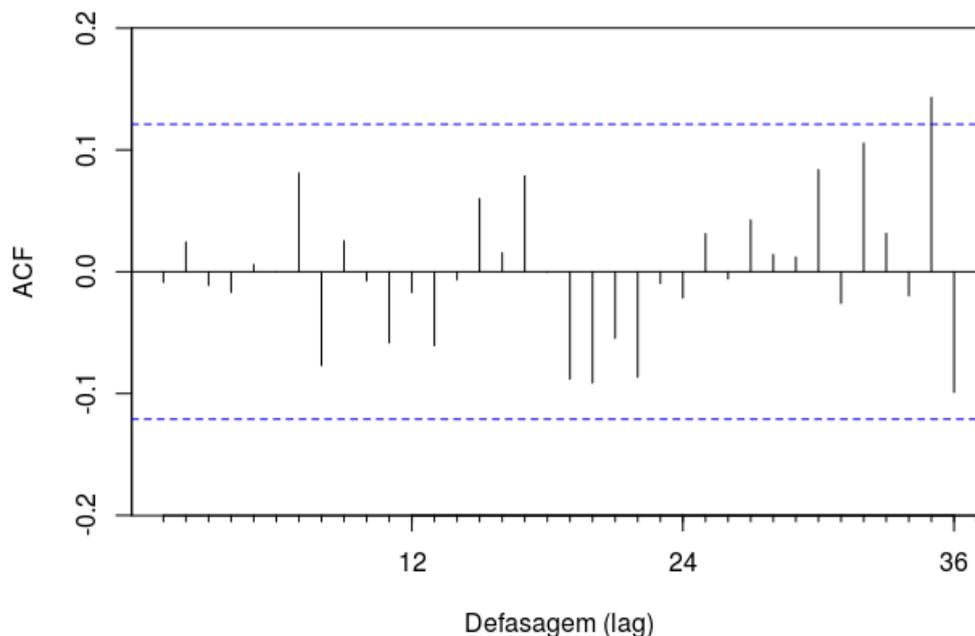


Figura 5: Função de autocorrelação (ACF) dos resíduos da série temporal das vazões médias mensais de água na usina hidrelétrica do Funil/MG, no período de janeiro de 1999 a outubro de 2020.

do Instituto Nacional de Meteorologia (INMET, 2020), choveu 183% a mais do que era esperado para todo o mês no estado.

Realizou-se um novo ajuste de modelo, agora com todos os dados da série histórica de vazões médias mensais de água para a usina hidrelétrica do Funil/MG, abrangendo o período de janeiro de 1999 a outubro de 2020, para, enfim, predizer vazões médias mensais para os meses de novembro e dezembro de 2020, e para o ano de 2021.

Ressalta-se que, assim como para o primeiro ajuste, as etapas de transformar a escala de dados, verificar a existência das componentes de tendência e/ou sazonalidade e removê-las, caso necessário, foram realizadas, contudo, não apresentadas, pois tornaria o texto repetitivo.

Mesmo com a inserção de 10 novas vazões médias mensais de água referentes ao período de janeiro a outubro de 2020 na série de dados, o modelo SARIMA $(2, 0, 1) \times (0, 1, 1)_{12}$ prevaleceu como melhor ajuste.

As estimativas dos coeficientes e respectivos erros padrão para o novo ajuste encontram-se na Tabela 4.

Pode-se verificar a qualidade do ajuste do modelo através do gráfico da ACF dos resíduos (ver Figura 5), onde se observa que mais de 95% dos lags estão contidos no intervalo de confiança e também através do valor-p (0,481) superior a 5% (nível de significância) para o teste de Ljung-Box, indicando uma série de resíduos estacionária e com ruído branco.

Este modelo de previsão proposto pode ser utilizado pela usina hidrelétrica do Funil/MG para realizar previsões a curto prazo, por exemplo, trimestrais, dessa forma, a cada trimestre o modelo pode ser atualizado com dados reais.

A Figura 6 ilustra o ajuste de tal modelo aos dados históricos de vazão média mensal de água na usina hidrelétrica do Funil/MG. Nota-se que os valores ajustados estão bem

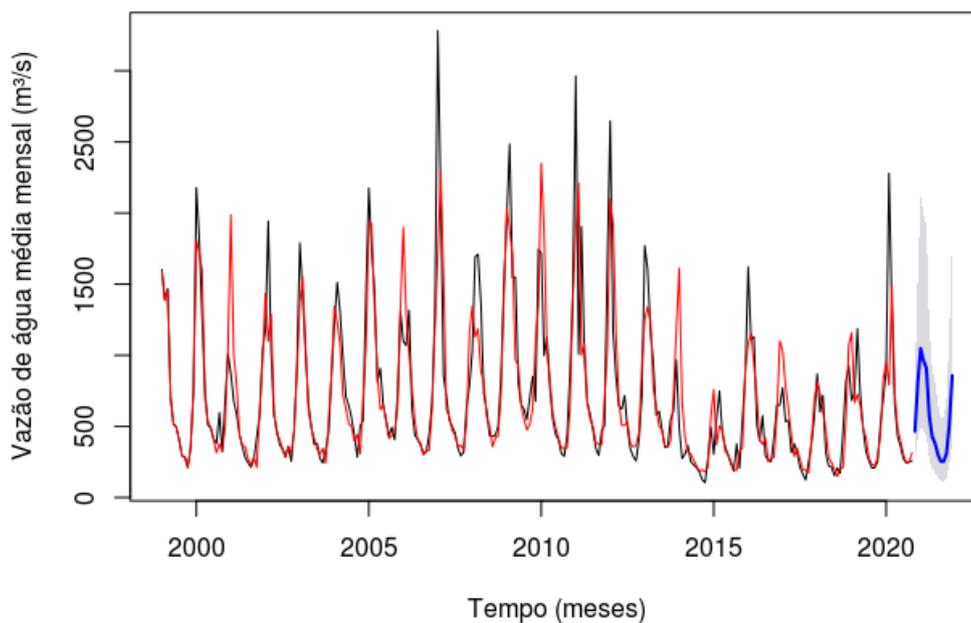


Figura 6: Gráfico da série de vazões médias mensais de água para os valores observados (preto), ajustados (vermelho) e previstos (azul).

Tabela 5: Estimativas das vazões médias mensais de água previstas para o período de novembro de 2020 a dezembro de 2021, assim como os seus respectivos intervalos de confiança a 95%.

Meses/Ano	Limite inferior (2,5%)	Previsto	Limite superior (97,5%)
novembro/2020	261	465	826
dezembro/2020	412	803	1565
janeiro/2021	517	1048	2122
fevereiro/2021	464	961	1989
março/2021	433	910	1910
abril/2021	270	575	1222
maio/2021	199	428	918
junho/2021	176	381	826
julho/2021	138	302	659
agosto/2021	115	254	557
setembro/2021	114	253	559
outubro/2021	139	309	686
novembro/2021	229	514	1156
dezembro/2021	378	856	1940

próximos dos reais, o que significa que o modelo conseguiu captar a dinâmica temporal da série.

Na Tabela 5 são apresentados os valores das vazões médias mensais previstas para o período de novembro de 2020 a dezembro de 2021 na usina hidrelétrica do Funil/MG, assim como os intervalos de confiança a 95%.

Como foi verificada uma periodicidade sazonal anual na série de vazões, observou-

se que as estimativas de previsão média mensal de água respeitam os ciclos de chuvas, portanto, entre meses de novembro e março, as vazões são elevadas (próximo do período de chuvas), entre meses de abril e junho, intermediárias, já entre meses de julho e outubro, as vazões são baixas (próximo do período de seca). Segundo Siqueira (2013), o regime de chuvas no Brasil varia de acordo com as estações do ano, sendo que janeiro é o mês com a maior densidade pluviométrica e agosto com a menor, o que comprova que a sazonalidade da série é resultante deste comportamento.

Conclusões

Nos períodos de baixas vazões, a hidrelétrica pode reduzir a vazão ao longo do dia, pois a demanda energética é menor e utilizá-la de forma acumulada na geração de energia nos horários de pico, compreendidos entre 18 e 21 horas.

Em função de haver uma variabilidade temporal, espacial e quantitativa, tanto na demanda quanto na disponibilidade de água, outra usina hidrelétrica poderia compensar a baixa geração de energia da usina hidrelétrica do Funil no período de seca (julho a outubro).

A estratégia de utilizar o modelo de previsão proposto a curto prazo vai permitir aos gestores da usina terem tempo suficiente para programar a operação e manutenção da vazão de forma adequada.

Outra alternativa seria os gestores da usina hidrelétrica do Funil estudarem a possibilidade de recorrerem à geração de energia através de fontes eólicas ou solares.

A metodologia de Box e Jenkins (1976) proposta mostrou-se adequada para a modelagem de dados de vazão, neste sentido, para futuros trabalhos, sugere-se sua aplicação com a inclusão da análise de intervenção.

Agradecimentos

Agradecemos ao Programa Institucional de Iniciação Científica Voluntária - PICV do CEFET-MG.

Referências Bibliográficas

AKAIKE, H. *Maximum likelihood identification of gaussian autoregressive moving average models*. Biometrika, London, v. 21, p. 243-247, 1973.

BARBOSA, E.C.; SÁFADI, T.; NASCIMENTO, M.; NASCIMENTO, A.C.C.; SILVA, C.H.O.; MANULI, R.C. *Metodologia Box & Jenkins para previsão de temperatura média mensal da cidade de Bauru (SP)*. Revista Brasileira de Biometria, v.33, n.1, p.104-117, 2015.

BATISTA, A.L.F. *Modelos de séries temporais e redes neurais artificiais na previsão de vazão*. Dissertação (Mestrado) - Universidade Federal de Lavras, 2009.

BOX, G.E.P.; JENKINS, G.M. *Time Series Analysis: Forecasting and Control*. San Francisco: HoldenDay, 1970 (Revised edition, 1976).

CHAN, K.; RIPLEY, B. TSA: Time Series Analysis:

<https://CRAN.Rproject.org/package=TSA>, 2018. (R package version 1.2).

HANNAN, E.J.; QUINN, B.G. *The determination of the order of autoregression.*

Journal of the Royal Statistical Society. Series B (Methodological), 41, n.2, p.190-195, 1979.

HYNDMAN, R. J. FORECAST: forecasting Functions for Time Series and Linear

Models: <https://CRAN.Rproject.org/package=forecast>, 2008. (R package version 8.14).

INSTITUTO NACIONAL DE METEOROLOGIA - INMET. Disponível em:

<https://portal.inmet.gov.br/boletinsagro>. Acesso em 15 de outubro de 2020.

LJUNG, G.M.; BOX, G.E.P. *On a measure of lack of fit in time series models.*

Biometrika, v.65, n.2, p.97-303, 1978.

MIRANDA, T. P. *Previsão da precipitação mensal do município de Ouro Branco - MG, por meio de modelos de séries temporais.* Dissertação (Mestrado) - Universidade Federal de São João Del Rei, 2016.

MORETTIN, P.A.; TOLOI, C.M.C. *Análise de Séries Temporais.* 2^a ed. São Paulo: E. Blucher, 2006, 538p.

OPERADOR NACIONAL DO SISTEMA ELÉTRICO - ONS. Disponível em:

<http://www.ons.org.br/Paginas/resultados-da-operacao/historico-da-operacao/>. Acesso em 15 de setembro de 2020.

PAIVA, D.A.; HERVAL, A.C.F.; SÁFADI, T. *Metodologia de Séries Temporais como ferramenta de análise na produção de frangos no Brasil.* Sigmae, v.8, n.2, p.227-237, 2019.

R CORE TEAM. *R: a language and environment for statistical computing.* Vienna: R Foundation for statistical computing. <http://www.Rproject.org>.

SIQUEIRA, H. V. *Máquinas desorganizadas para previsão de séries de vazões.* Tese (Doutorado) - Universidade Estadual de Campinas, 2013.