

## Simulação de Vazões e Níveis de Água Médios Mensais para o Rio Tapajós Usando Modelos ARIMA

Nelio Moura Figueiredo<sup>1</sup>, Claudio José Cavalcante Blanco<sup>2</sup>  
nelio@ufpa.br; blanco@ufpa.br

Recebido: 28/05/13 - revisado: 25/11/13 - aceito: 24/03/14

### RESUMO

Este trabalho apresenta uma análise de previsões de vazão e níveis de água médios mensais, com antecedência de 24 meses, para a bacia do rio Tapajós, PA, utilizando modelos estocásticos do tipo ARIMA. A identificação do modelo foi feita através da análise do coeficiente e função de autocorrelação (ACF) e do coeficiente e função de autocorrelação parcial (PACF). Os critérios de verossimilhança apresentaram menores valores para o modelo ARIMA  $(1,0,0)(1,1,1)_{12}$ . Os critérios de desempenho usados foram o coeficiente de Nash e Sutcliffe  $R^2$  e o RMSE (Root Mean Square Error). O modelo ARIMA  $(1,0,0)(1,1,1)_{12}$  apresentou melhor desempenho para as séries de vazão e de níveis de água, nas fases de calibração e validação. Na calibração, o modelo apresentou  $R^2$  médio de 0,94 para níveis de água e de 0,92 para vazões e RMSE médio de 0,08 para níveis de água e de 0,15 para vazões. Na validação, o  $R^2$  médio foi de 0,95 para níveis de água e de 0,91 para vazões e RMSE médio de 0,07 para níveis de água e de 0,15 para vazões. Foi observado, em função da maior linearidade do comportamento temporal dos níveis de água, que esses foram melhor simulados que as vazões. Resultado que pode ser explicado pela concepção linear dos modelos ARIMA. No contexto de uso futuro da água da bacia do Tapajós, o modelo considerado, é uma ferramenta interessante, podendo ser usada para analisar o regime hidrológico da bacia, através de previsões de níveis de água, diante das UHE a serem instaladas e, conseqüentemente, dos possíveis impactos ambientais e conflitos pelo uso da água.

**Palavras-chave:** Modelagem hidrológica, modelo estocástico ARIMA, previsão de níveis de água e previsão de vazão, Amazônia.

### INTRODUÇÃO

A região amazônica será palco preferencial para a expansão da geração hidrelétrica nas próximas décadas. Do potencial hidrelétrico nacional, 41% encontram-se na região amazônica (MME/EPE, 2007). A construção do complexo de usinas na bacia do rio Tapajós, entre os Estados do Amazonas e do Pará, prevê a construção de cinco usinas hidrelétricas – São Luiz de Tapajós, Jatobá, Cachoeira dos Patos, Jamanxim e Cachoeira do Caí. A Usina Hidrelétrica (UHE) de São Luiz do Tapajós será a de maior importância, com previsão de geração de 6.133 MW, terá potência inferior apenas às de Itaipu, Belo Monte e Tucuruí. Nesse contexto, o estudo do regime hidrológico da bacia do Tapajós, assume, cada vez mais, importância estratégica na

previsão de vazões e de níveis de água; na percepção de alterações do regime hídrico em função de geração em períodos de vazão reduzida a jusante de barragens; na prospecção de condições de navegabilidade e na harmonização de conflitos ambientais e socioeconômicos gerados pela formação dos reservatórios das UHEs. Em síntese, estudos dessa natureza são importantes, tendo em vista a minimização de conflitos inerentes ao uso múltiplo de água.

Nesse caso, simulações hidrológicas estocásticas de séries temporais de vazão e de níveis de água são comumente baseadas em modelos matemáticos. O modelo pode ser escolhido por sua característica hidrológica e pela disponibilidade e propriedade dos dados estatísticos. Segundo Nourani et al. (2013) modelos auto regressivos de média móvel integrada ARIMA (*Auto Regressive Integrated Moving Average*), com sazonalidade agregada, são amplamente utilizados para a previsão de séries temporais hidrológicas. Lohani et al. (2012) utilizaram modelos auto regressivos juntamente com sistemas de inferência fuzzy na previsão de vazões mensais. Wu e Chau (2010) empregaram modelos ARIMA e redes neu-

<sup>1</sup> Faculdade de Engenharia Naval, ITEC, UFPA

<sup>2</sup> Faculdade de Engenharia Sanitária e Ambiental, ITEC, UFPA

rais na previsão de vazões em diferentes bacias hidrográficas na China. Birinci e Akay (2010) utilizaram modelos do tipo ARIMA para a previsão de precipitações, que, por sua vez, são insumos (inputs) em modelos de redes neurais artificiais para previsão de vazões diárias. Koutroumanidis et al. (2009) desenvolveram modelo de previsão de vazão para a bacia do rio Nestos na Bulgária com proposta híbrida baseada em modelos ARIMA e em procedimentos de detecção e genética para identificar pontos de mudança estrutural nas séries temporais. Srinivas e Srinivasan (2006) ressaltaram que nas últimas quatro décadas, desde o trabalho pioneiro de Fiering (1964), uma série de estudos têm abordado a aplicação de modelos paramétricos para simulação estocástica de variáveis hidrológicas.

Acerca do desempenho de modelos estocásticos, Perrin et al. (2001) pontuam que trabalhos como os de Ye et al. (1997) e Gan et al. (1997) já destacavam que a robustez de um modelo não é proporcional à complexidade do mesmo, ou seja, modelos simples podem produzir tão bons resultados quanto os mais complexos.

No Brasil, Bayer et al. (2012) apresentaram um modelo de previsão de vazões para a bacia do rio Potiribu, afluente do rio Ijuí, RS, obtendo resultados satisfatórios para uma antecedência de seis meses. Lucas et al. (2009) utilizaram dois modelos, um determinístico e outro estocástico do tipo ARIMA, para simular a vazão média mensal das sub-bacias da região hidrográfica do Xingu no Estado do Pará. Mine e Tucci (1999) utilizaram um modelo ARIMA, conjuntamente com um modelo determinístico chuva-vazão denominado IPH-II, na previsão de vazões afluentes à estação Usina Foz de Areia, no rio Iguçu/PR.

Assim, o objetivo deste artigo foi realizar a modelagem e a previsão, com antecedência de até vinte e quatro meses, de vazões e de níveis de água médios mensais. Nesse caso, a modelagem dos níveis de água foi tomada sem referências altimétricas ao nível médio dos mares, pois, infelizmente, a maioria das réguas limimétricas instaladas em nosso país não possuem essa referência, não sendo diferente para o rio Tapajós. A análise desenvolvida considerou modelos estocásticos sazonais do tipo ARIMA.

## DESCRIÇÃO DA ÁREA DE ESTUDO

O modelo ARIMA foi aplicado à bacia do rio Tapajós, que está inserida na bacia hidrográfica do rio Amazonas que se constitui na maior bacia hidrográfica do globo terrestre com uma área total

de 6.110.000 km<sup>2</sup>, desde suas nascentes nos Andes Peruanos até sua foz no oceano Atlântico. Esta bacia continental estende-se sobre vários países da América do Sul: Brasil (63%), Peru (17%), Bolívia (11%), Colômbia (5,8%), Equador (2,2%), Venezuela (0,7%) e Guiana (0,2%) (ANA, 2002).

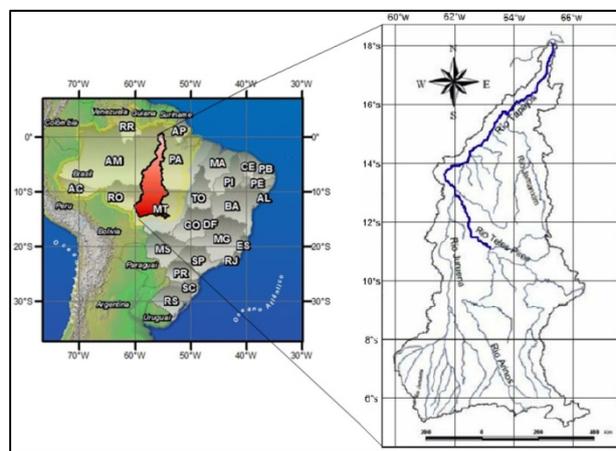


Figura 1 – Localização da bacia do rio Tapajós

A bacia do rio Tapajós está situada nos estados do Pará, Mato Grosso e uma pequena porção do Amazonas, entre latitudes 2° e 15° Sul e 53° e 61° Oeste, possui uma área de drenagem de 493.200 km<sup>2</sup>, com configuração alongada no sentido Sul-Norte. Seus rios formadores são o Juruena e o Teles Pires, com áreas de drenagem de 191.100 km<sup>2</sup> e 139.000 km<sup>2</sup>, respectivamente (Figura 1). A altitude varia de cerca de 900m, nas cabeceiras da bacia, até cerca de 50 metros acima do nível do mar, junto à foz no rio Amazonas. A localização da bacia do rio Tapajós abrange o Oeste do estado do Pará, próximo à divisa com o estado do Amazonas. O rio Tapajós, com extensão de 851 km, é o mais importante afluente da margem direita do rio Amazonas. Nasce na confluência dos rios Teles Pires e Juruena, junto à divisa dos estados do Pará, Amazonas e Mato Grosso, compreendendo a porção média da bacia Amazônica. Em termos de clima, a bacia é dominada pela Zona de Convergência Intertropical (ZCIT). Apresenta baixa pressão, sendo uma característica própria de regiões próximas à linha do Equador. A precipitação média da bacia está situada entre 1.800 e 2.300 mm ao ano (ANA, 2002). A Figura 1 mostra a localização da área de estudo.

Os principais tipos de solo que ocorrem na bacia do rio Tapajós são: solos litólicos, areias quartzosas, latos solos, podzólicos, solos de várzea, solos concrecionários e cambissolos. Na ocupação e uso

do solo, ocorre a predominância de 75% de cobertura de floresta, 10% de pastagem, 7% de solo com pouca cobertura, 10% de cultivo agrícola e cerca de 1% está coberta pelos cursos de água (COLLISCHONN, 2006).

**DADOS UTILIZADOS**

A base de dados utilizada no modelo foi obtida junto ao banco de dados da Agência Nacional de Águas (ANA), constituindo-se de séries históricas de níveis de água médios diários e de vazões médias diárias. Esses dados encontram-se disponibilizados no endereço eletrônico <http://hidroweb.ana.gov.br>.

A modelagem das vazões médias mensais foi desenvolvida em quatro estações fluviométricas; e a dos níveis de água médios mensais em sete estações fluviométricas localizadas na bacia do rio Tapajós. A identificação e a localização das estações utilizadas na modelagem das vazões são apresentadas na Figura 2; e a dos níveis de água na Figura 3. Em ambas, foram utilizadas séries históricas com 33 anos de observação, no período de 1976 a 2008. A Tabela 1 resume os dados e informações relativas às estações utilizadas na modelagem das vazões e a Tabela 2 dos níveis de água.

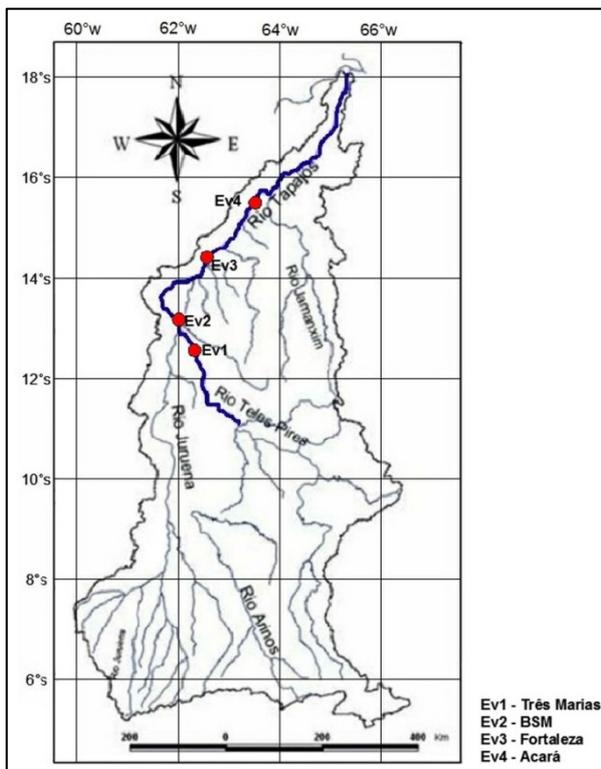


Figura 2 – Localização das estações de vazão

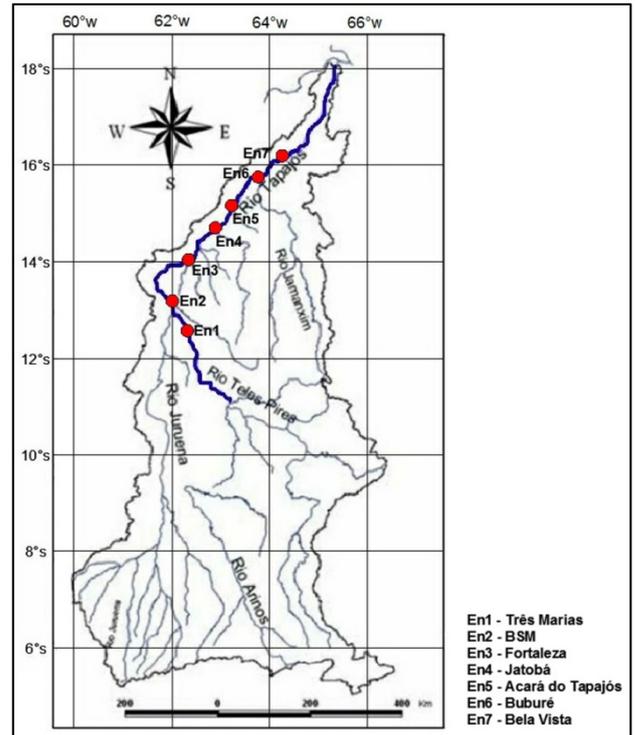


Figura 3 – Localização das estações de níveis de água

Tabela 1 - Informações das estações utilizadas na modelagem das vazões

Código ANA	Nome	Área (Km <sup>2</sup> )	Rio
17420000	Três Marias	138000	Teles Pires
17430000	BSM	333000	Tapajós
17500000	Fortaleza	363000	Tapajós
17650002	Acará	390000	Tapajós

Tabela 2 - Informações das estações utilizadas na modelagem dos níveis de água.

Código ANA	Nome	Área (Km <sup>2</sup> )	Rio
17420000	Três Marias	138000	Teles Pires
17430000	BSM	333000	Tapajós
17500000	Fortaleza	363000	Tapajós
17650000	Jatobá	387000	Tapajós
17650002	Acará	390000	Tapajós
17710000	Buburé	450000	Tapajós
17720000	Bela Vista	453000	Tapajós

O preenchimento de falhas nas observações foi feito com a utilização do método da Ponderação Regional com base em Regressão Linear (Equação 1) descrito por Bertoni e Tucci (2007) e Alves et al. (2006). Segundo Bertoni e Tucci (2007), Alves et al. (2006) e Pruski et al. (2004), trata-se de um método que estabelece regressões lineares entre as estações com os dados a serem preenchidos  $Y_m$  e cada um dos postos vizinhos  $Y_i$ .

$$Y_m = \frac{\sum_{i=1}^n Y_i r_{mi}}{\sum_{i=1}^n r_{mi}} \quad (1)$$

onde,  $Y_m$  é a observação da estação base  $m$  a ter a falha na série preenchida;  $n$  é o número de estações;  $Y_i$  é a observação da estação  $i$  e  $r_{mi}$  é o coeficiente de correlação linear entre os postos  $m$  e  $i$ .

A consistência dos dados fluviométricos foi feita com a utilização do método de regressão múltipla (Equação 2).

$$Y_{est} = \beta_1 + \beta_2 Y_m + \beta_3 Y_j \quad (2)$$

onde,  $Y_{est}$  é a observação a ser estimada na estação em análise;  $Y_m$  é a observação do posto a montante do posto em análise;  $Y_j$  é a observação do posto a jusante do posto em análise e  $\beta_1$ ,  $\beta_2$  e  $\beta_3$  são coeficientes de regressão múltipla.

A eficiência da correlação e a confiabilidade da regressão foram verificadas através da confrontação e da análise dos erros entre os valores estimados e os observados.

## METODOLOGIA

Box e Jenkins (1976), conceituando série temporal como um conjunto de observações de uma variável dispostas sequencialmente no tempo, presuppõe que exista um processo estocástico gerador da série e que cada possível realização aleatória da variável esteja associada a uma probabilidade de ocorrência da observação. A metodologia está assentada na suposição de que a série foi gerada por um processo estocástico que pode ser descrito e caracterizado, com base no comportamento passado da variável.

O processo de construção e de ajuste do modelo proposto por Box e Jenkins está alicerçado em um ciclo iterativo. A Figura 4 apresenta o fluxograma da metodologia de Box e Jenkins, evidenciando as etapas inerentes ao processo de aplicação e utilização do modelo (BOX e JENKINS, 1976).

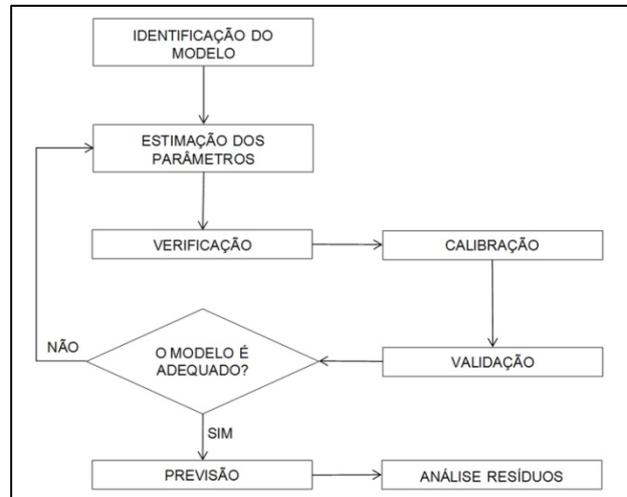


Figura 4 – Fluxograma da metodologia de Box e Jenkins

A etapa de identificação do modelo foi realizada com a observância dos gráficos da função de autocorrelação ACF e da função de autocorrelação parcial PACF das séries de vazão e de níveis de água.

O coeficiente de autocorrelação ( $\rho_k$ ), definido pela Equação 3, tem a função de medir como estão relacionadas as observações de uma série temporal afastadas  $k$  períodos entre si. Com o cálculo dos coeficientes de autocorrelação  $\rho_k$  para diversas defasagens, obtém-se uma série que permite construir a função de autocorrelação (*autocorrelation function – ACF*).

$$\rho_k = \frac{\sum_{t=k+1}^n (Y_t - \bar{Y})(Y_{t-k} - \bar{Y})}{\sum_{t=1}^n (Y_t - \bar{Y})^2} \quad (3)$$

onde,  $\bar{Y}$  é a média das observações, e  $k$  é o número de defasagens das observações.

O coeficiente de autocorrelação parcial ( $\phi_{kk}$ ), definido pela Equação 4, mede o nível de associação entre observações após a remoção dos efeitos das auto correlações. Com o cálculo dos coeficientes de autocorrelação parcial  $\phi_{kk}$  para diversas defasagens, a função de autocorrelação parcial (*partial autocorrelation function – PACF*) é obtida.

$$\phi_{kk} = \left( \rho_k - \sum_{j=1}^{k-1} \phi_{k-1j} \rho_{k-j} \right) \cdot \left( 1 - \sum_{j=1}^{k-1} \phi_{k-1j} \rho_j \right)^{-1} \quad (4)$$

A seleção do modelo foi feita com a utilização dos critérios de AIC (*Akaike Information Criterion*), de AICC (*Akaike Information Criterion Corrected*) e de BIC (*Bayesian Information Criteria*). Esses critérios baseiam-se na metodologia da máxima verossimilhança e são expressos pelas Equações 5, 6 e 7, respectivamente.

$$AIC = -2 \log \hat{L} + 2k \quad (5)$$

$$AICC = -2 \log \hat{L} + \frac{2(p+q)T}{T-p-q-1} \quad (6)$$

$$BIC = -2 \log \hat{L} + (p+q) \log T \quad (7)$$

onde T é o número de observações,  $k = p + q + 1$ , se o modelo tiver um termo constante e  $k = p + q$  para os demais modelos.  $\hat{L}$  é a verossimilhança maximizada (SIMAS et al.(2010), OSPINA et al. (2006) e VASCONCELLOS e CRIBARI-NETO (2005)). O melhor modelo é aquele que apresentar o menor AIC, AICC e/ou o menor BIC.

A calibração e a validação do modelo foram desenvolvidas com a utilização de função objetivo que adotou como coeficientes de desempenho o coeficiente de determinação  $R^2$  ou de Nash e Sutcliffe e o RMSE (*Root Mean Square Error*- raiz do erro relativo quadrático médio), descritos pelas Equações 8 e 9, respectivamente.

$$R^2 = 1 - \frac{\sum_{t=1}^n (Y_t - \hat{Y}_t)^2}{\sum_{t=1}^n (Y_t - \bar{Y})^2} \quad (8)$$

$$RMSE = \sqrt{\frac{\sum_{t=1}^n \left( \frac{Y_t - \hat{Y}_t}{Y_t} \right)^2}{n}} \quad (9)$$

onde n é o número de observações;  $Y_t$  é a observação no instante t;  $\hat{Y}_t$  é a observação simulada; e  $\bar{Y}$  é a média das n observações.

O processo não estacionário homogêneo, descrito pela transformação da série em estacionária e, posteriormente, pela utilização de um processo misto auto regressivo e de média móvel, é chamado de processo auto regressivo integrado de média móvel de ordem (p,d,q) ou ARIMA (p,d,q).

Box e Jenkins (1976) sugerem a aplicação de um modelo ARIMA sazonal para descrever a série possuidora de correlação serial nos períodos sazonalmente defasados, conforme Equação 10. Esse modelo é representado por ARIMA (p,d,q) (P,D,Q)<sub>s</sub>, onde p, d e q são parâmetros não sazonais e P, D e Q são parâmetros sazonais auto regressivos, de integração e de média móvel, respectivamente.

$$\phi(B) \cdot \Phi(B^s) \cdot (1 - B^s)^D \cdot (1 - B)^d \cdot Y_t = \theta(B) \cdot \Theta(B^s) \cdot a_t \quad (10)$$

onde:  $\phi$  é o coeficiente auto regressivo não sazonal;  $\Phi$  é o coeficiente auto regressivo sazonal;  $\theta$  é o coeficiente média móvel não sazonal e  $\Theta$  é o coeficiente média móvel sazonal. Os demais termos presentes na Equação 10 são operadores assim definidos:

- B é o operador translação, sendo  $B^m Z_t = Z_{t-m}$ ;
- $\phi(B) = (1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p)$  é o operador de translação auto regressivo não sazonal de ordem “p”;
- $\Phi(B^s) = (1 - \Phi_1 B^s - \Phi_2 B^{2s} - \dots - \Phi_p B^{ps})$  é o operador de translação auto regressivo sazonal de ordem “P”;
- $\theta(B) = (1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q)$  é o operador de translação média móvel não sazonal de ordem “q”; e
- $\Theta(B^s) = (1 - \Theta_1 B^s - \Theta_2 B^{2s} - \dots - \Theta_Q B^{Qs})$  é o operador de translação média sazonal de ordem “Q”.

Os coeficientes auto regressivos e médias móveis, sazonais e não sazonais, foram obtidos através de métodos de estimação, com a busca interativa da minimização do erro quadrático dos resíduos. Na presença dos termos de médias móveis, há não linearidade, logo, usou-se um procedimento iterativo de estimação não linear para obter os valores das estimativas dos parâmetros do modelo (DELURGIO, 1998).

O procedimento iterativo de estimação não linear dos parâmetros permitiu o teste da significância dos estimadores por meio de estatística  $t$ . A função que busca a otimização dos coeficientes foi linearizada através de expansão em séries de Taylor em torno de um conjunto inicial de parâmetros (LEVENBACH e CLEARY, 1984).

A previsão de observações com modelos de séries temporais é um procedimento que visa simular observações futuras com modelos calibrados, validados e ajustados a valores passados e presentes. A previsão do modelo foi feita com a utilização da equação característica geral de um ARIMA (p,d,q) (P,D,Q)<sub>s</sub> sazonal (Equação 10).

Assim, para o modelo ARIMA (1,0,0) (1,1,1)<sub>12</sub>, a Equação 10, com a substituição e a operacionalização dos valores dos parâmetros p=1, d=0, q=0 e P=1, D=1, Q=1, pode ser escrita em sua forma linear, conforme a Equação 11.

$$\hat{Y}_t = \Phi_1 Y_{t-12} + Y_{t-12} - \Phi_1 Y_{t-24} + \phi_1 Y_{t-1} - \phi_1 \Phi_1 Y_{t-13} - \phi_1 Y_{t-13} + \phi_1 \Phi_1 Y_{t-25} - \theta_1 a_{t-12} \quad (11)$$

onde:

$\hat{Y}_t = Y_t - a_t$  é a previsão mensal de vazão ou de nível de água no instante  $t$ ;

$Y_t$  é a observação mensal de vazão ou de nível de água no instante  $t$ ; e

$a_t = \hat{Y}_t - Y_t$  é o resíduo da previsão no instante  $t$ .

A operacionalização da metodologia sugerida por Box e Jenkins foi feita com a utilização de algoritmos implementados no software R da *R Foundation for Statistical Computing*, versão 2.14.1, que é uma linguagem e ambiente de computação estatística e de construção de gráficos, sendo um projeto GNU (*General Public License da Free Software Foundation*).

## RESULTADOS E DISCUSSÃO

Os gráficos de boxplot das séries de vazão consistidas mostram a ocorrência de poucos *outliers*, evidenciando um considerável desempenho dos processos utilizados no preenchimento das falhas e na consistência dos dados dessas séries. Os *outliers* encontram-se, em sua grande maioria, localizados acima dos valores máximos de cada diagrama de Boxplot mensal, evidenciando, para as análises das vazões e de níveis de água de estiagem, que as poucas discrepâncias ocorridas são desprezíveis.

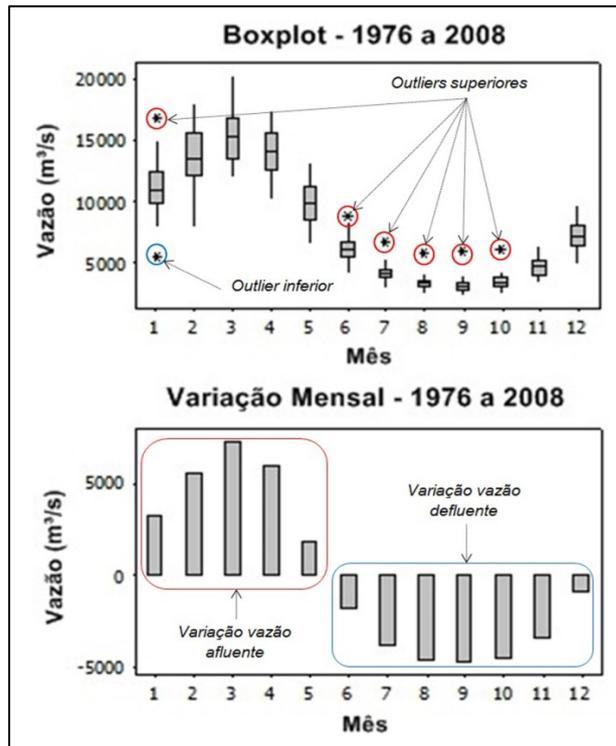


Figura 5 – Comportamento da série de vazões mensais na estação Barra do São Manuel

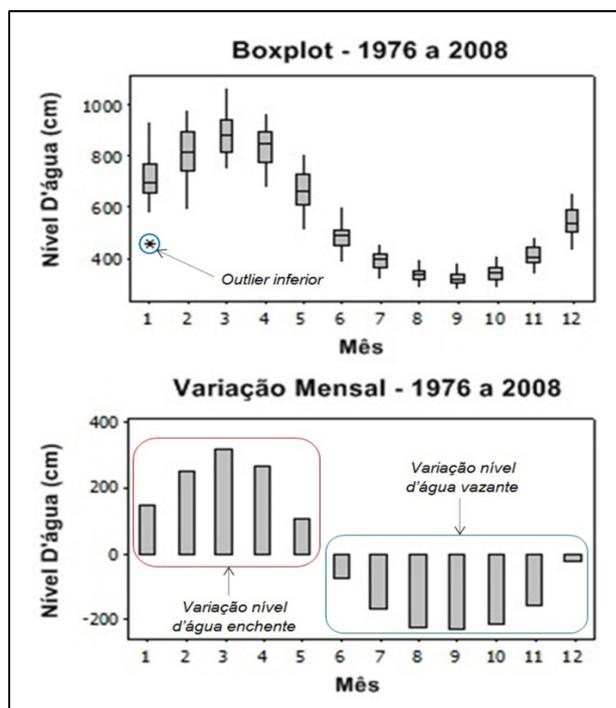


Figura 6 – Comportamento da série de níveis de água mensais na estação Barra do São Manuel

Após o preenchimento de falhas e da consistência dos dados, o comportamento das séries de vazão e de níveis de água, no período de 1976 a 2008, foi analisado através de gráficos de Boxplot e de variação mensal. Como exemplo, a Figura 5 mostra o comportamento da série de vazão para a estação Barra do São Manuel; e a Figura 6 o da série de níveis de água para a mesma estação.

O comportamento da amplitude de variação da componente sazonal das séries de vazão e de níveis de água é uniforme. O período de enchente e de vazão afluente ocorre entre os meses de janeiro a maio e o de vazante e de vazão defluente entre junho e dezembro. A estiagem dá-se entre agosto e outubro.

Observe-se que nos diagramas de boxplot das séries de vazão e de níveis de água uma maior amplitude interquartil entre os meses de janeiro a maio. No período defluente as séries apresentam menores variâncias e menores amplitudes interquartis. Na estiagem as séries apresentaram uma baixa variabilidade temporal de suas componentes sazonais, o que refletiu em um melhor desempenho dos modelos para esses períodos.

**Tabela 3 - Critérios de AIC, AICc e BIC para as séries de vazão na calibração.**

Estação	Modelo ARIMA – Vazão Calibração	
	(0,0,0)	(1,0,0)
	(1,1,1)	(1,1,1)
	AIC	AIC
Três Marias	5493,20	5336,59
Fortaleza	6220,27	5941,45
BSM	6291,27	5713,65
Acará	6287,78	5956,72
	AICc	AICc
Três Marias	5493,27	5336,77
Fortaleza	6220,34	5941,63
BSM	6291,34	5713,83
Acará	6287,85	5956,90
	BIC	BIC
Três Marias	5504,65	5355,66
Fortaleza	6231,72	5960,52
BSM	6302,93	5732,72
Acará	6299,23	5975,79

A identificação dos modelos ARIMA baseou-se na análise gráfica das funções de autocorrelação ACF e de autocorrelação parciais PACF geradas com base nos coeficientes de autocorrelação  $\rho_k$  (Equação 3) e de autocorrelação parcial  $\phi_{kk}$  (Equação 4). O

comportamento dessas funções auxiliou, também, na verificação da estacionariedade e na proposição dos modelos.

**Calibração**

Na fase de calibração, a estimação dos modelos compreendeu na aferição dos critérios das Equações 5, 6 e 7. Dentre as diversas combinações de parâmetros atribuídas ao modelo ARIMA (p,d,q) (P,D,Q)<sub>s</sub>, as que apresentaram menores valores para esses critérios foram: ARIMA (0,0,0) (1,1,1)<sub>12</sub> e ARIMA (1,0,0) (1,1,1)<sub>12</sub>. Outras combinações de parâmetros apresentaram valores próximos, mas o princípio da parcimônia prevaleceu. As Tabelas 3 e 4 mostram os valores dos critérios AIC, AICc e BIC aferidos para as séries de vazão e de níveis de água.

**Tabela 4 - Critérios de AIC, AICc e BIC para as séries de níveis de água na calibração.**

Estação	Modelo ARIMA – Níveis de água Calibração	
	(0,0,0)	(1,0,0)
	(1,1,1)	(1,1,1)
	AIC	AIC
Três Marias	3821,04	3647,32
Jatobá	3700,79	3640,47
Fortaleza	3532,10	3354,59
Buburé	3466,26	3269,62
BSM	3767,57	3588,15
Bela Vista	3954,53	3751,36
Acará	3697,60	3512,64
	AICc	AICc
Três Marias	3821,11	3647,50
Jatobá	3700,86	3640,65
Fortaleza	3532,18	3354,77
Buburé	3466,33	3269,80
BSM	3767,64	3588,33
Bela Vista	3954,60	3751,54
Acará	3697,67	3512,82
	BIC	BIC
Três Marias	3832,49	3666,39
Jatobá	3712,24	3659,54
Fortaleza	3543,56	3373,66
Buburé	3477,71	3286,34
BSM	3779,02	3607,22
Bela Vista	3965,98	3769,03
Acará	3709,05	3530,52

Os valores dos critérios de estimação da calibração (Tabelas 3 e 4), evidenciaram que o modelo

ARIMA (1,0,0)(1,1,1)<sub>12</sub> foi o que apresentou melhor resultado.

O desempenho dos modelos na calibração foi aferido através do coeficiente de Nash e Sutcliffe R<sup>2</sup> (Equação 8) e da raiz do erro relativo quadrático médio RMSE (Equação 9). As Tabelas 5 e 6 mostram, para as séries de vazão e de níveis de água, os valores do R<sup>2</sup> e do RMSE, respectivamente, para cada modelo analisado.

**Tabela 5 - R<sup>2</sup> e RMSE para as séries de vazão na calibração.**

Estação	Modelo ARIMA – Vazão Calibração	
	(0,0,0) (1,1,1) R <sup>2</sup>	(1,0,0) (1,1,1) R <sup>2</sup>
Três Marias	0,914	0,918
Fortaleza	0,904	0,907
BSM	0,930	0,944
Acará	0,894	0,914
	RMSE	
Três Marias	0,167	0,154
Fortaleza	0,186	0,160
BSM	0,146	0,127
Acará	0,215	0,167

**Tabela 6 - R<sup>2</sup> e RMSE para as séries de níveis de água na calibração**

Estação	Modelo ARIMA – Níveis de água Calibração	
	(0,0,0) (1,1,1) R <sup>2</sup>	(1,0,0) (1,1,1) R <sup>2</sup>
Três Marias	0,941	0,946
Jatobá	0,929	0,932
Fortaleza	0,922	0,923
Buburé	0,917	0,923
BSM	0,944	0,951
Bela Vista	0,909	0,917
Acará	0,921	0,928
	RMSE	
Três Marias	0,076	0,070
Jatobá	0,065	0,063
Fortaleza	0,078	0,075
Buburé	0,062	0,059
BSM	0,092	0,084
Bela Vista	0,155	0,137
Acará	0,090	0,084

Tendo por base os valores do R<sup>2</sup> e do RMSE da função objetivo de calibração do modelo aplicado às vazões e níveis de água, constatou-se que o modelo ARIMA (1,0,0)(1,1,1)<sub>12</sub>, apresentou melhor desempenho.

**Validação**

A estimação do modelo na fase de validação deu-se com base nos mesmos critérios da calibração.

**Tabela 7 - R<sup>2</sup> e RMSE para as séries de vazão na validação.**

Estação	Modelo ARIMA – Vazão Validação	
	(1,0,0) (1,1,1) R <sup>2</sup>	
Três Marias	0,896	
Fortaleza	0,902	
BSM	0,925	
Acará	0,894	
	RMSE	
Três Marias	0,157	
Fortaleza	0,141	
BSM	0,130	
Acará	0,153	

**Tabela 8 - R<sup>2</sup> e RMSE para as séries de níveis de água na validação.**

Estação	Modelo ARIMA – Níveis de água Validação	
	(1,0,0) (1,1,1) R <sup>2</sup>	
Três Marias	0,945	
Jatobá	0,952	
Fortaleza	0,933	
Buburé	0,957	
BSM	0,941	
Bela Vista	0,952	
Acará	0,940	
	RMSE	
Três Marias	0,068	
Jatobá	0,052	
Fortaleza	0,065	
Buburé	0,044	
BSM	0,083	
Bela Vista	0,099	
Acará	0,078	

Na validação, o desempenho dos modelos foi aferido igualmente aos da calibração, com a adoção do coeficiente de Nash e Sutcliffe R<sup>2</sup> e da raiz do

erro relativo quadrático médio RMSE. As Tabelas 7 e 8 apresentam os valores dos coeficientes para cada modelo analisado.

Os valores do R<sup>2</sup> e do RMSE da função objetivo de validação do modelo, constantes das Tabelas 7 e 8, mostraram que o modelo ARIMA (1,0,0) (1,1,1)<sub>12</sub>, apresentou melhor desempenho nessa fase.

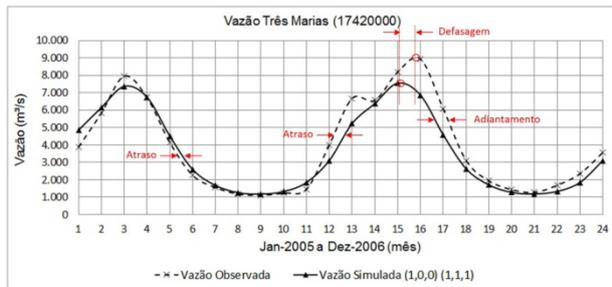
A Tabela 9 resume, através dos R<sup>2</sup> e dos RMSE médios, o desempenho do modelo para as séries de vazão e de níveis de água nas fases de calibração e de validação.

**Tabela 9 - R<sup>2</sup> e RMSE médios das séries de vazão e de níveis de água.**

Fase	Vazão		Níveis de água	
	R <sup>2</sup>	RMSE	R <sup>2</sup>	RMSE
Calibração	0,92	0,15	0,93	0,08
Validação	0,90	0,16	0,95	0,06

**Aferição de Desempenho**

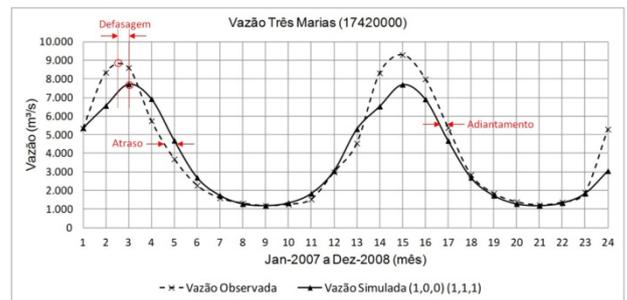
A aferição do desempenho dos modelos nas fases de calibração e de validação deu-se mediante confrontação entre os valores simulados e os dados observados. As séries de dados relativas ao período de 1976 a 2004 formaram a memória do modelo. A calibração foi feita para o período de 2005 a 2006 e a validação para o de 2007 a 2008.



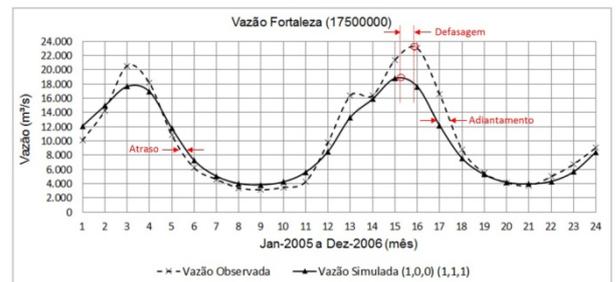
**Figura 7 – Hidrogramas das vazões observadas e simuladas na estação Três Marias - Calibração**

As Figuras 7 a 14 apresentam os hidrogramas das vazões simuladas e observadas para as fases de calibração e de validação em cada estação. Na estigem, o modelo apresentou desempenho satisfatório na simulação das vazões nos dois períodos sazonais, tanto na fase de calibração quanto na de validação, não apresentando diferenças significativas

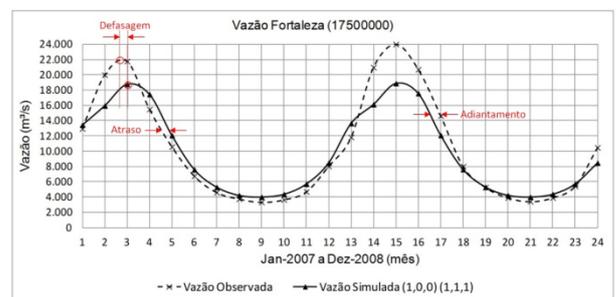
de comportamento entre essas fases. Na defluência das vazões, do primeiro período sazonal, tanto na calibração quanto na validação, observou-se um atraso das vazões observadas em relação às simuladas, tendo maior significância na validação. Na calibração, em todas as estações, os hidrogramas evidenciaram, entre as lags 13 e 15, que o modelo simulou de forma uniforme à inflexão da afluência das vazões junto do ponto correspondente a lag 14, mantendo constante a deflexão ascendente das vazões.



**Figura 8 - Hidrogramas das vazões observadas e simuladas na estação Três Marias - Validação**



**Figura 9 - Hidrogramas das vazões observadas e simuladas na estação Fortaleza - Calibração**



**Figura 10 - Hidrogramas das vazões observadas e simuladas na estação Fortaleza - Validação**

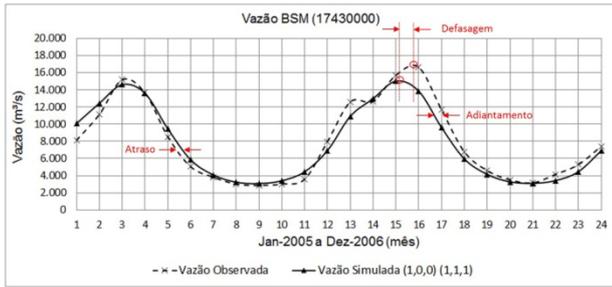


Figura 11 - Hidrogramas das vazões observadas e simuladas na estação Barra do São Manuel - Calibração

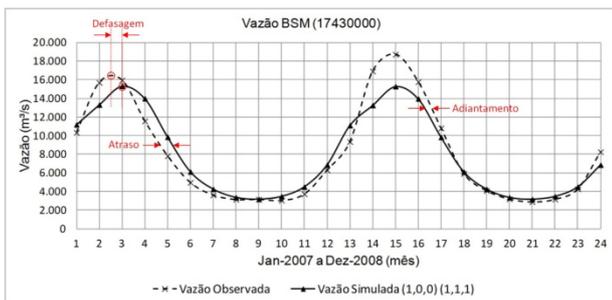


Figura 12 - Hidrogramas das vazões observadas e simuladas na estação Barra do São Manuel - Validação

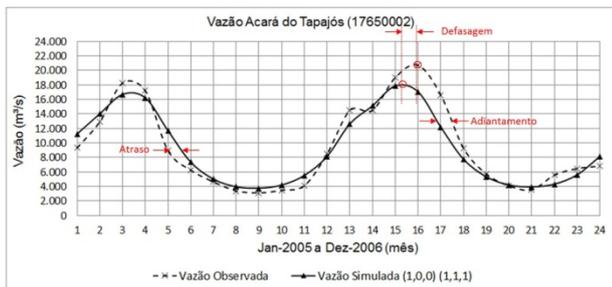


Figura 13 - Hidrogramas das vazões observadas e simuladas na estação Acará do Tapajós - Calibração

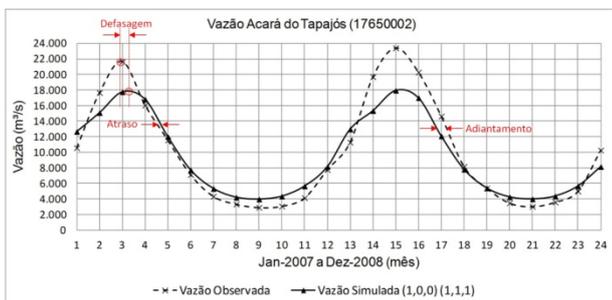


Figura 14 - Hidrogramas das vazões observadas e simuladas na estação Acará do Tapajós - Validação

Os hidrogramas evidenciaram, tanto na fase de calibração quanto na de validação, que o modelo apresentou melhor desempenho na simulação das vazões mínimas do que das máximas, uma vez que o ajuste dos parâmetros do modelo foi direcionado às estiagens. Sob a ótica das vazões máximas, no primeiro período sazonal da fase de validação, observou-se uma significativa defasagem das vazões simuladas em relação às observadas. Na calibração, essa defasagem fez-se presente no segundo período sazonal. Na defluência das vazões do segundo período sazonal, tanto na calibração quanto na validação, ocorreu um adiantamento das vazões observadas em relação às simuladas, com inversão de fase dos hidrogramas.

As Figuras 15 a 28 apresentam os cotogramas dos níveis de água observados e simulados para as fases de calibração e de validação em cada estação. Observou-se, tanto na fase de calibração quanto na de validação, que o modelo apresentou desempenho satisfatório na simulação dos níveis de água mínimos, com atrasos uniformes dos níveis de água observados em relação aos simulados nas vazantes do primeiro período sazonal e adiantamentos nas do segundo.

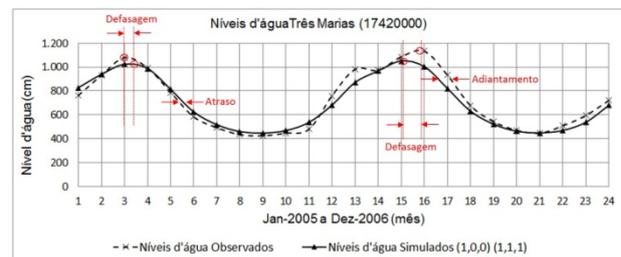


Figura 15 - Cotogramas dos níveis de água observados e simulados na estação Três Marias - Calibração

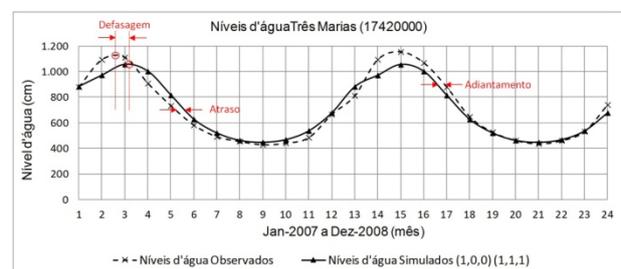


Figura 16 - Cotogramas dos níveis de água observados e simulados na estação Três Marias - Validação

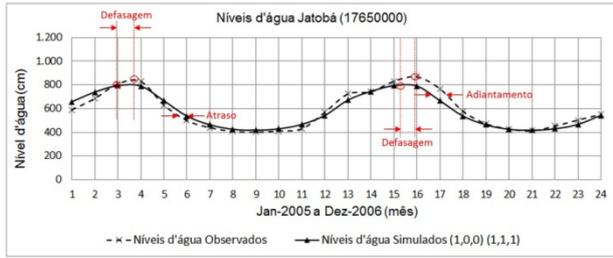


Figura 17 - Cotagramas dos níveis de água observados e simulados na estação Jatobá - Calibração

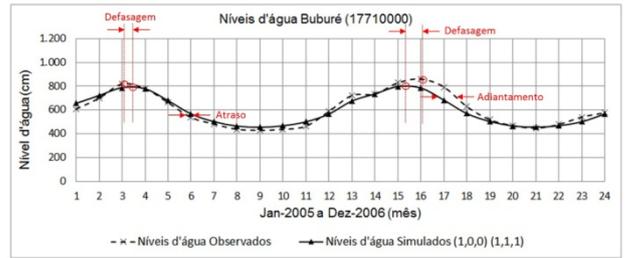


Figura 21 - Cotagramas dos níveis de água observados e simulados na estação Buburé - Calibração



Figura 18 - Cotagramas dos níveis de água observados e simulados na estação Jatobá - Validação

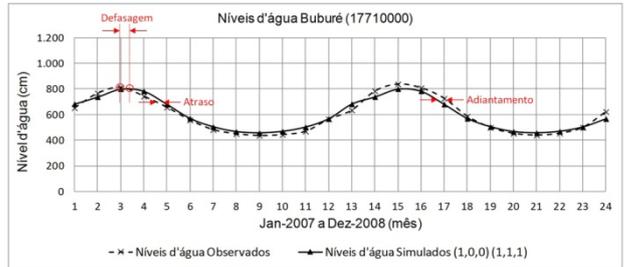


Figura 22 - Cotagramas dos níveis de água observados e simulados na estação Buburé - Validação

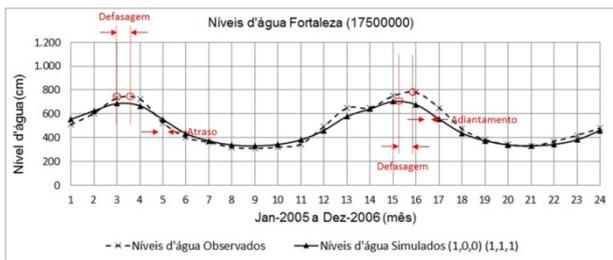


Figura 19 - Cotagramas dos níveis de água observados e simulados na estação Fortaleza - Calibração

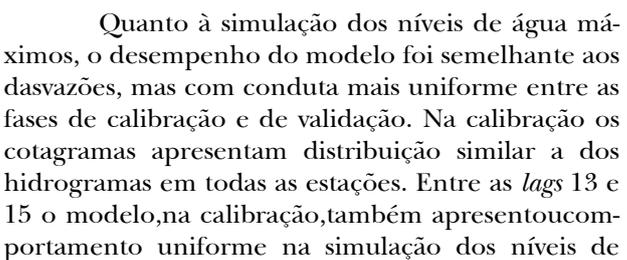


Figura 23 - Cotagramas dos níveis de água observados e simulados na estação Barra do São Manuel - Calibração

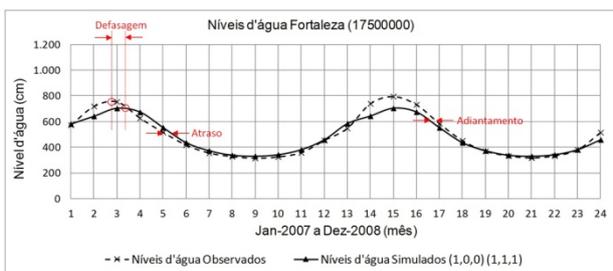
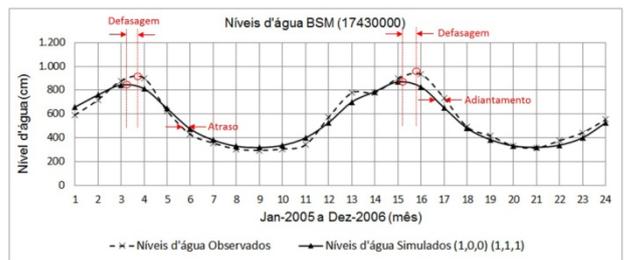


Figura 20 - Cotagramas dos níveis de água observados e simulados na estação Fortaleza - Validação



Quanto à simulação dos níveis de água máximos, o desempenho do modelo foi semelhante aos das vazões, mas com conduta mais uniforme entre as fases de calibração e de validação. Na calibração os cotagramas apresentam distribuição similar a dos hidrogramas em todas as estações. Entre as lags 13 e 15 o modelo, na calibração, também apresentou comportamento uniforme na simulação dos níveis de água junto à inflexão do ponto correspondente ao lag 14.

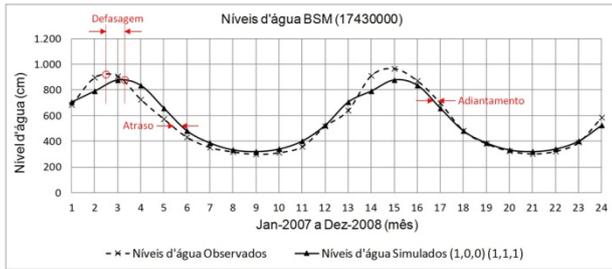


Figura 24 - Cotogramas dos níveis de água observados e simulados na estação Barra do São Manuel - Validação



Figura 27 - Cotogramas dos níveis de água observados e simulados na estação Acará do Tapajós - Calibração

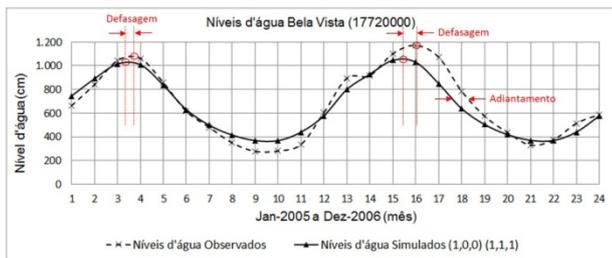


Figura 25 - Cotogramas dos níveis de água observados e simulados na estação Bela Vista - Calibração

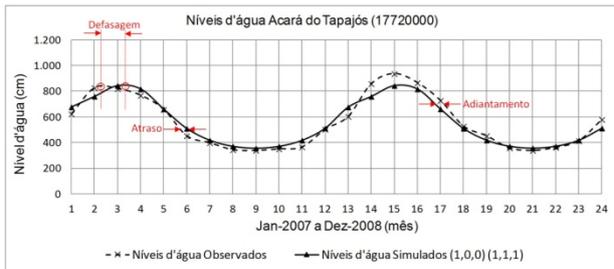


Figura 28 - Cotogramas dos níveis de água observados e simulados na estação Acará do Tapajós - Validação

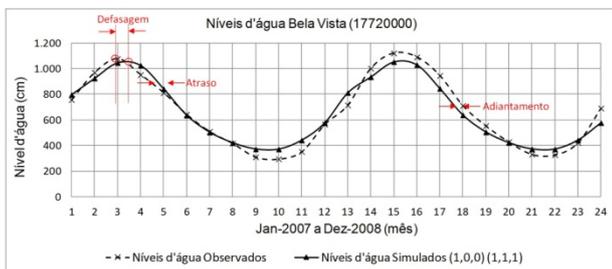


Figura 26 - Cotogramas dos níveis de água observados e simulados na estação Bela Vista - Validação

A análise e a simulação dos dados das séries e a topologia da equação de previsão do modelo (Equação 11), evidenciaram que as defasagens existentes entre os picos máximos das simulações são resultantes do fato de que o modelo até o ponto correspondente a *lag* 12, adota resíduos relacionados com valores observados. A partir do ponto correspondente a *lag* 13, o modelo afere seus resíduos, tendo como referência valores simulados e não mais os observados. As defasagens do segundo período sazonal, apresentadas de forma localizada no período da calibração, decorrem do ajuste linear do modelo à inflexão pontual da afluência de vazão junto do ponto correspondente a *lag* 14.

Os hidrogramas e cotogramas das fases de calibração e de validação e os valores dos critérios de desempenho da Tabela 11 evidenciaram que o modelo apresentou maior robustez na simulação dos níveis de água do que das vazões. Um aumento de desempenho, na fase de validação, é observado apenas para os níveis de água. Uma explicação plausível para essa melhor simulação, baseia-se na maior não linearidade da vazão em relação aos níveis de água. O que matematicamente pode ser observado pelas maiores inclinações das curvas de vazão, representando maiores gradientes dessa grandeza em relação aos níveis de água, esse fato favorece a simulação dos níveis de água através de modelos lineares como os ARIMA utilizados nesse trabalho.

Por outro lado, segundo Mine e Tucci (1999), modelos do tipo ARIMA apresentam respostas atrasadas nas subidas dos hidrogramas, pelo fato de dependerem apenas das informações ocorridas no próprio local. O erro do modelo é maior na subida dos hidrogramas, pois as previsões, nas defasagens iniciais, estão sempre um pouco atrasadas em relação às observações.

#### Previsão e Verificação do Modelo

As equações de previsão de vazão e de níveis de água  $\hat{Y}_t$  foram obtidas, respectivamente, com a substituição dos coeficientes das Tabelas 12 e 13 na

Equação 11. A sequência de Equações 12 a 15 mostram as equações de previsão de vazão para as estações Três Marias, Fortaleza, BSM e Acará, respectivamente e as Equações 16 a 22 as equações de previsão de níveis de água para as estações Três Marias, Jatobá, Fortaleza, Buburê, BSM, Bela Vista e Acará, respectivamente.

As Tabelas 10 e 11 mostram os coeficientes auto regressivo e de média móvel, não sazonal e sazonal e sazonal e sazonal, obtidos via aplicação do modelo ARIMA (1,0,0) (1,1,1)<sub>12</sub> às séries de vazão e de níveis de água, em cada estação.

**Tabela 10 - R<sup>2</sup> e Coeficientes auto regressivo e média móvel, não sazonal e sazonal para as séries de vazão.**

Estação	Coeficientes ARIMA (0,0,0) (1,1,1) - Vazão		
	$\phi_1$	$\Phi_1$	$\theta_1$
Três Marias	0,6196	0,0097	0,9698
Fortaleza	0,7493	-0,0436	0,9520
BSM	0,6571	0,0090	0,9660
Acará	0,7911	-0,0416	0,9581

**Tabela 11 - R<sup>2</sup> e Coeficientes auto regressivo e média móvel, não sazonal e sazonal para as séries de níveis de água.**

Estação	Coeficientes ARIMA (0,0,0) (1,1,1) – Níveis de água		
	$\phi_1$	$\Phi_1$	$\theta_1$
Três Marias	0,6364	-0,0189	0,9438
Jatobá	0,4038	0,0407	0,9673
Fortaleza	0,6510	-0,0275	0,9458
Buburê	0,6750	0,0234	0,9531
BSM	0,6537	-0,0157	0,9477
Bela Vista	0,6868	0,6868	0,9538
Acará	0,6559	0,0230	0,9477

$$\hat{Y}_t = 0,0097.Y_{t-12} + Y_{t-12} - 0,0097.Y_{t-24} + 0,6196.Y_{t-1} - 0,6196.(0,0097).Y_{t-13} - 0,6196.Y_{t-13} + 0,6196.(0,0097).Y_{t-25} - 0,9698.a_{t-12} \quad (12)$$

$$\hat{Y}_t = -0,0436.Y_{t-12} + Y_{t-12} + 0,0436.Y_{t-24} + 0,7493.Y_{t-1} - 0,7493.(-0,0436).Y_{t-13} - 0,7493.Y_{t-13} + 0,7493.(-0,0436).Y_{t-25} - 0,9520.a_{t-12} \quad (13)$$

$$\hat{Y}_t = 0,0090.Y_{t-12} + Y_{t-12} - 0,0090.Y_{t-24} + 0,6571.Y_{t-1} - 0,6571.(0,0090).Y_{t-13} - 0,6571.Y_{t-13} + 0,6571.(0,0090).Y_{t-25} - 0,9660.a_{t-12} \quad (14)$$

$$\hat{Y}_t = -0,0416.Y_{t-12} + Y_{t-12} + 0,0416.Y_{t-24} + 0,7911.Y_{t-1} - 0,7911.(-0,0416).Y_{t-13} - 0,7911.Y_{t-13} + 0,7911.(-0,0416).Y_{t-25} - 0,9581.a_{t-12} \quad (15)$$

$$\hat{Y}_t = -0,0189.Y_{t-12} + Y_{t-12} + 0,0189.Y_{t-24} + 0,6364.Y_{t-1} - 0,6364.(-0,0189).Y_{t-13} - 0,6364.Y_{t-13} + 0,6364.(-0,0189).Y_{t-25} - 0,9438.a_{t-12} \quad (16)$$

$$\hat{Y}_t = 0,0407.Y_{t-12} + Y_{t-12} - 0,0407.Y_{t-24} + 0,4038.Y_{t-1} - 0,4038.(0,0407).Y_{t-13} - 0,4038.Y_{t-13} + 0,4038.(0,0407).Y_{t-25} - 0,9673.a_{t-12} \quad (17)$$

$$\hat{Y}_t = -0,0275.Y_{t-12} + Y_{t-12} + 0,0275.Y_{t-24} + 0,6510.Y_{t-1} - 0,6510.(-0,0275).Y_{t-13} - 0,6510.Y_{t-13} + 0,6510.(-0,0275).Y_{t-25} - 0,9458.a_{t-12} \quad (18)$$

$$\hat{Y}_t = 0,0234.Y_{t-12} + Y_{t-12} - 0,0234.Y_{t-24} + 0,6750.Y_{t-1} - 0,6750.(0,0234).Y_{t-13} - 0,6750.Y_{t-13} + 0,6750.(0,0234).Y_{t-25} - 0,9531.a_{t-12} \quad (19)$$

$$\hat{Y}_t = -0,0157.Y_{t-12} + Y_{t-12} + 0,0157.Y_{t-24} + 0,6537.Y_{t-1} - 0,6537.(-0,0157).Y_{t-13} - 0,6537.Y_{t-13} + 0,6537.(-0,0157).Y_{t-25} - 0,9477.a_{t-12} \quad (20)$$

$$\hat{Y}_t = -0,0053.Y_{t-12} + Y_{t-12} + 0,0053.Y_{t-24} + 0,6868.Y_{t-1} - 0,6868.(-0,0053).Y_{t-13} - 0,6868.Y_{t-13} + 0,6868.(-0,0053).Y_{t-25} - 0,9538.a_{t-12} \quad (21)$$

$$\hat{Y}_t = 0,0230.Y_{t-12} + Y_{t-12} - 0,0230.Y_{t-24} + 0,6559.Y_{t-1} - 0,6559.(0,0230).Y_{t-13} - 0,6559.Y_{t-13} + 0,6559.(0,0230).Y_{t-25} - 0,9477.a_{t-12} \quad (22)$$

A verificação dos modelos foi feita mediante a análise dos limites de significância de 95% demarcados nos gráficos da ACF e da PACF dos resíduos. Como exemplo as Figuras 29 e 30 mostram a ACF e a PACF dos resíduos resultantes da aplicação do modelo ARIMA(1,0,0) (1,1,1)<sub>12</sub> às séries de vazão e de níveis de água da estação Acará do Tapajós.

Os gráficos das funções da ACF e da PACF dos resíduos mostraram, em praticamente todas as defasagens, que os coeficientes de correlação encontram-se entre os limites críticos de confiança de 95%. Dessa forma, constata-se que o modelo ARI-

MA(1,0,0)(1,1,1)<sub>12</sub>, incorporou de forma satisfatória as características das séries na previsão das observações de vazões e de níveis de água.

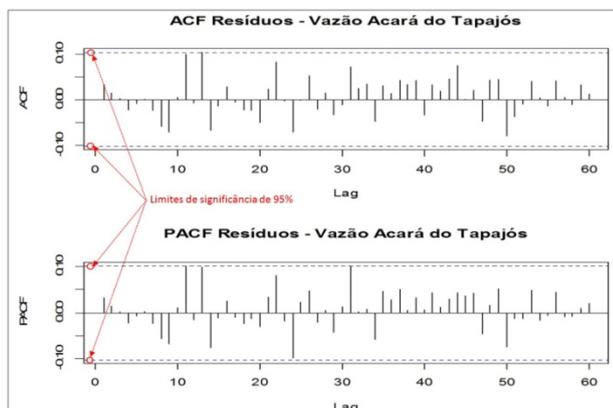


Figura 29 – ACF e PACF dos resíduos das vazões simuladas na estação Acará

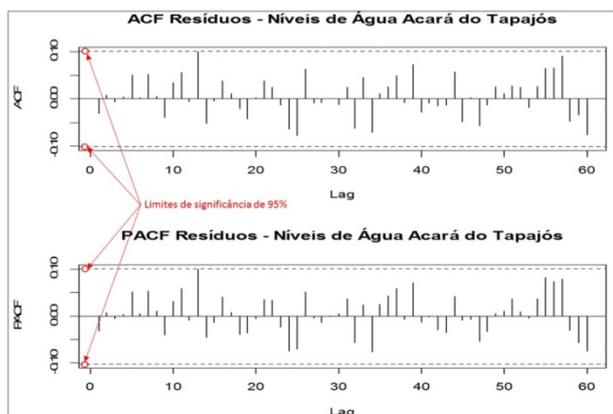


Figura 30 – ACF e PACF dos resíduos dos níveis de água simulados na estação Acará

## CONCLUSÃO

O modelo ARIMA (1,0,0)(1,1,1)<sub>12</sub> foi aplicado e verificado tendo por base a metodologia proposta por Box e Jenkins. Modelos auto regressivos e de médias móveis foram ajustados às séries de vazão e de níveis de água do Rio Tapajós.

A calibração e a validação do modelo, que tiveram como critério de análise de desempenho o coeficiente de Nash e Sutcliffe R<sup>2</sup> e a raiz do erro relativo quadrático médio RMSE, apresentaram R<sup>2</sup> médio acima de 0,90 e RMSE abaixo de 0,16. Tais resultados evidenciaram que o modelo captou de

forma satisfatória o comportamento das séries de níveis de água e de vazão.

A aplicação do modelo às séries de níveis de água mostrou-se mais eficiente que as de vazão, com um razoável aumento de desempenho do modelo entre as fases de calibração e de validação. As simulações nas estiagens, em função de sua menor variabilidade interquartil, apresentaram melhores resultados. No período de cheia, o modelo deve ser aplicado com cautela, uma vez que grande parte dos eventos extremos foi subestimada.

A verificação dos modelos, feita mediante a análise dos limites de significância de 95% demarcados nos gráficos da ACF e da PACF dos resíduos, evidenciaram que os coeficientes de correlação e de correlação parcial ficaram dentro dos limites de significância de 95%, evidenciando que o modelo apresentou desempenho satisfatório.

De maneira geral, conclui-se que o modelo ARIMA concebido apresentou um comportamento plenamente satisfatório aos objetivos da pesquisa, a qual, no contexto de uso futuro da água da bacia do Tapajós, é uma ferramenta interessante, podendo ser usada para analisar o regime hidrológico da bacia, diante das UHE a serem instaladas e, conseqüentemente, dos possíveis impactos ambientais e conflitos pelo uso da água.

## REFERÊNCIAS

- AGÊNCIA NACIONAL DE ÁGUAS – ANA. (2002). *A Evolução da Gestão dos Recursos Hídricos no Brasil / The Evolution of Water Resources Management in Brazil*. Brasília.
- ALVES, A. KOBAYAMA, M. SILVA, R. V. CHECCHIA, T. (2006). *Análise de dados hidrológicos na região de município de Alfredo Wagner/SC*. In: Simpósio de Recursos Hídricos do Sul, I, 2006, Curitiba. Anais, Porto Alegre, Associação Brasileira de Recursos Hídricos.
- BAYER, D. M., CASTRO, N. M. R., BAYER, F. M., (2012). *Modelagem e Previsão de Vazões Médias Mensais do Rio Potiribu*. Revista Brasileira de Recursos Hídricos. V. 17, nº 2. pp. 61-72.
- BERTONI, J. C., TUCCI, C. E. M. (2007). *Precipitação*. In: Tucci, C. E. M. Hidrologia: Ciência e Aplicação. Porto Alegre: UFRGS, 2007. P. 177-241.
- BIRINCI, V.; AKAY, O. (2010). *A Study on Modeling Daily Mean Flow with MLR, ARIMA and RBFNN*. In Anais do BALWOIS: Water observation and information system for

- decision support. Ohrid, República da Macedônia. 25-29 de maio de 2010.
- BOX, G. E. P. and JENKINS, G. M. (1976). *Time series analysis: forecasting and control, 2nd ed.* San Francisco: Holden-Day.
- COLLISCHONN, B. (2006). *Uso de precipitação estimada pelo satélite TRMM em modelo hidrológico distribuído*, Dissertação de mestrado, IPHUFGRS, 193 páginas, anexos.
- DeLURGIO, S. A. (1998). *Forecasting principles and applications*. 1st Edition, Singapore: McGraw-Hill. 802p.
- FIERING, M.B. (1964). *Multivariate technique for synthetic hydrology*. Journal of the Hydraulic Division ASCE 90, 43–60.
- GAN, T.Y., DLAMINI, E.M., BIFTU, G.F. (1997). *Effects of model complexity and structure, data quality and objective function on hydrologic modeling*. Journal of Hydrology 192, 81-103.
- KOUTROUMANIDIS, T.; SYLAIOS, G.; ZAFEIRIOU, E.; TSIHRINTZIS, V. A. (2009). *Genetic modeling for the optimal forecasting of hydrologic time-series: Application in Nestos River*, Journal of Hydrology, Volume 368, Issues 1–4, 30 April 2009, Pages 156-164, ISSN 0022-1694, 10.1016/j.jhydrol.2009.01.041.
- LEVENBACH, H. e CLEARY, J. P. (1984). *“The Modern Forecaster: The Forecasting Process Through Data Analysis”*, Lifetime Learning Publications, Belmont, Califórnia
- LOHANI, A.K.; KUMAR, R.; SINGH, R.D. (2012). *Hydrological time series modeling: A comparison between adaptive neuro-fuzzy, neural network and autoregressive techniques*, Journal of Hydrology, Volumes 442–443, 6 June 2012, Pages 23-35, ISSN 0022-1694, 10.1016/j.jhydrol.2012.03.031.
- LUCAS, E.W.M.; SOUSA, F.A.S.; SILVA, F.D.S; LUCIO, P.S. (2009). *Modelagem hidrológica determinística e estocástica aplicada à região hidrográfica do Xingu – Pará*. Revista Brasileira de Meteorologia, v.24, n.3, 308-322.
- MINISTÉRIO DE MINAS E ENERGIA (MME), EMPRESA DE PESQUISA ENERGÉTICA ( EPE) (2007). *Matriz Energética Nacional 2030*. Brasília, MME/EPE.
- MINE, M.R.; TUCCI, C.E. (1999). *Previsão em tempo real de vazões afluentes a reservatório de usinas hidrelétricas*. Revista Brasileira de Recursos Hídricos, Vol. 9, n. 2, p. 73-99.
- NOURANI, V.; BAGHANAM, A. H.; ADAMOWSKI, J.; GEBREMICHAEL, M. (2013). *Using self-organizing maps and wavelet transforms for space-time pre-processing of satellite precipitation and runoff data in neural network based rainfall-runoff modeling*, Journal of Hydrology, Volume 476, 7 January 2013, Pages 228-243, ISSN 0022-1694, 10.1016/j.jhydrol.2012.10.054.
- OSPINA, R.; CRIBARI-NETO, F.; VASCONCELLOS, K. L. P. (2006). *Improved point and interval estimation for a beta regression model*. Comput. Stat. Data Anal., Amsterdam, v.51, n.2, p.960-981.
- PERRIN, C.; MICHEL, C.; ANDRÉASSIAN, V. (2001). *Does a large number of parameters enhance model performance? Comparative assessment of common catchments model structures on 429 catchments*. Journal of Hydrology, v. 242, n. 3/4, p. 275-301.
- PRUSKI, F. F.; PEREIRA, S. B., NOVAES, L. F.; SILVA, D. D.; RAMOS, M. M. (2004) *Precipitação média anual e vazão específica média de longa duração, na Bacia do São Francisco*. Revista Brasileira de Engenharia Agrícola e Ambiental, v. 8, n. 2/3, p. 247-253.
- SIMAS, A. B.; BARRETO-SOUZA, W.; ROCHA, A. V. (2010). *Improved estimators for a general class of beta regression models*. Comput. Stat. Data Anal., Amsterdam, v.54, n.2, p.348-366.
- SRINIVAS, V. V. and SRINIVASAN, K. (2006). *Hybrid Matched-Block Bootstrap for Stochastic Simulation of Multi-season streamflows*. Journal of Hydrology, Elsevier, Vol.329, Issues 1-2, pp.1-15.
- VASCONCELLOS, K. L. P.; CRIBARI-NETO, F. (2005) *Improved maximum likelihood estimation in a new class of beta regression models*. Braz. J. Probab. Stat., São Paulo, v.19, p.13-31.
- YE, W., BATES, B.C., VINEY, N.R., SILVAPAN, M., JAKEMAN, A.J. (1997). *Performance of conceptual rainfall-runoff models in low-yielding ephemeral catchments*. Water Resources Research 33 (1), 153±166.
- WU, C.L.; CHAU, K.W. (2010). *Data-driven models for monthly streamflow time series prediction*, Engineering Applications of Artificial Intelligence, Volume 23, Issue 8, December 2010, Pages 1350-1367, ISSN 0952-1976, 10.1016/j.engappai.2010.04.003.

***Simulation Of Monthly Average Flows And Water Levels For Tapajós River, Using ARIMA Models***

**ABSTRACT**

*This paper presents an analysis of predictions of monthly average flows and water levels, in advance of 24 months, for the river Tapajós, in Brazil, using stochastic models such as ARIMA. The model identification was performed through the analysis of the autocorrelation coefficient and function (ACF) and the analysis of the partial autocorrelation coefficient and function (PACF). The likelihood criteria showed lower values for the ARIMA (1,0,0)(1,1,1)<sub>12</sub>. The performance criteria used were the coefficient of Nash and Sutcliffe R<sup>2</sup> and RMSE (Root Mean Square Error). The ARIMA model (1,0,0)(1,1,1)<sub>12</sub> showed a better performance for the series of flow and water levels, both in the calibration and validation phases. For calibration, the model presented an average R<sup>2</sup> equal to 0.94 for water levels and an average R<sup>2</sup> equal to 0.92 for flow rates; and average RMSE equal to 0.08 for water levels and average RMSE equal to 0.15 for flow rates. For validation, the average R<sup>2</sup> was equal to 0.95 for water levels and average R<sup>2</sup> was equal to 0.91 for flow rates; and average RMSE equal to 0.07 for water levels and average RMSE equal to 0.15 for flow rates. It was observed, due to the larger linearity of the temporal behavior of water levels, that these were better simulated than the flow rates. This result can be explained by the linear conception of the ARIMA models. In the context of future use of the water in the Tapajós basin, the model considered, is an interesting tool and can be used to analyze the hydrological regime of the basin, through predictions of water levels, against the HPP to be installed and, consequently, the potential environmental impacts and water use conflicts.*

**Key-words:** Hydrologic modeling, stochastic model ARIMA, water levels forecast, streamflow forecast, Amazon.