



**Modelos de Regressão Aplicados à  
Previsão de Nível Fluviométrico no Rio Sapucaí em Itajubá-MG**  
Regression Models Applied to Fluviometric  
Level Forecast in the Sapucaí River in Itajubá-MG

Michelle Simões Reboita; Benedito Cláudio da Silva & Mateus Vitoriano da Silva

*Universidade Federal de Itajubá, Instituto de Recursos Naturais  
Avenida BPS, 1303, Pinheirinho, 37500-903, Itajubá, MG, Brasil  
E-mails: reboita@gmail.com; silvabenedito@gmail.com; mateus\_5555@hotmail.com  
Recebido em: 18/03/2019 Aprovado em: 14/05/2019  
DOI: [http://dx.doi.org/10.11137/2019\\_2\\_217\\_229](http://dx.doi.org/10.11137/2019_2_217_229)*

### Resumo

Tendo em vista que o município de Itajubá, localizado no sul do Estado de Minas Gerais, é atingido por inundações devido ao aumento de nível do rio Sapucaí, principalmente no verão, o objetivo deste estudo é obter modelos de previsão de nível para a estação fluviométrica Santa Rosa, situada no início da área urbana de Itajubá. Para isso, foram utilizados dados de outras três estações fluviométricas, a montante da estação Santa Rosa, no ano de 2013 e modelos de regressão linear simples e múltipla e de regressão polinomial cúbica. Os dados do mês de janeiro de 2013 foram utilizados para o ajuste dos modelos, enquanto que os dados do restante do ano para validação. Em geral, o aumento de nível na estação Santa Rosa ocorre entre três e sete horas após esse evento ter ocorrido nas demais estações. Então, através da técnica de regressão e usando dados com diferentes defasagens temporais, obtiveram-se vários modelos preditivos que apresentaram boa performance (medida em termos de índices estatísticos) em prever o nível fluviométrico na estação Santa Rosa. Entretanto, um dos melhores ajustes estatísticos foi obtido com a regressão linear múltipla envolvendo três estações como preditoras com dados defasados em 4 horas. Assim, sugere-se a utilização desse modelo para monitoramento operacional do nível na estação Santa Rosa, para fins de emissão de alertas de cheia na área urbana de Itajubá.

**Palavras-Chave:** previsão de nível fluviométrico; sul de Minas Gerais; análise de regressão

### Abstract

Considering that the municipality of Itajubá, located in the south of Minas Gerais state, is affected by flooding due to the increase in the level of the Sapucaí river, mainly in summer, the purpose of this study is to obtain forecast models for the river level at the Santa Rosa fluviometric station located at the beginning of the urban area of Itajubá. For this study, data from three other fluviometric stations, located upstream of the Santa Rosa station, of 2013-year and simple and multiple linear regressions and cubic polynomial regression were used. Data from January 2013 were used to adjust the models, while the data from February to December for validation. In general, the increase in the fluviometric level at Santa Rosa station occurs between three and seven hours after those in the other stations. Then, through the regression technique and using data with different temporal lags, several forecast models were obtained. Although they presented good performance (measured in terms of statistical indices) in predicting the fluviometric level at the Santa Rosa station, the best statistical adjustments were obtained with multiple linear regression involving three stations as predictors and to the lag of four hours. Thus, it is suggested to use this model for operational level monitoring at the Santa Rosa station in order to provide flood alerts in the urban area of Itajubá.

**Keywords:** fluviometric level forecast; south of Minas Gerais State; regression analysis

## 1 Introdução

As inundações, definidas como o extravasamento de um curso d'água a sua cota máxima e que atinge a planície de inundação (Amaral & Gutjahr, 2014), estão entre os fenômenos naturais que mais provocam danos no globo, incluindo, além da perda de vidas, prejuízos econômicos, sociais e ambientais (Alfieri & Thielen, 2015; Ye et al., 2016). Apesar disso, os sistemas de previsão e alerta de inundações, que sejam efetivos na redução desses impactos, ainda são um grande desafio para os governos, principalmente para as autoridades municipais (Mercuryeva et al., 2015). Fato que pode estar relacionado à necessidade de investimento financeiro para implementação e manutenção desses sistemas bem como na ineficiência da divulgação das informações.

Um sistema de alerta é formado por três componentes: i) monitoramento hidrometeorológico, ii) modelagem e previsão de vazões e/ou níveis e iii) armazenamento e divulgação dos dados. Em relação ao segundo componente, a previsão de eventos de inundações pode ser feita em diferentes escalas temporais (Silva et al., 2006), isto é, com antecedência que pode variar de alguns minutos ou horas (nowcasting), alguns dias (curto prazo, até 14 dias), semanas ou meses (sub-sazonal ou sazonal, até 9 meses). A escala temporal das previsões dependerá das características da bacia hidrográfica, como área da bacia, declividade, uso e cobertura do solo, além do tipo de evento meteorológico que provoca as inundações. A combinação das características físicas condiciona o tempo de concentração da bacia, que é o tempo gasto para a água percorrer a drenagem principal, desde a nascente mais distante até o ponto de interesse. De maneira geral, melhores resultados são obtidos se a escala temporal do modelo de previsão for menor do que o tempo de concentração da bacia.

Quanto aos modelos que podem ser empregados na previsão de vazões e níveis, dois tipos principais podem ser considerados (Silva et al. 2006; Mercuryeva et al., 2015): i) modelos empíricos, que utilizam apenas uma relação entre entrada e saída, sem considerar os processos físicos envolvidos. Alguns exemplos desses modelos são as redes neurais artificiais, equações de regressão e modelos estocás-

ticos; ii) modelos conceituais, que simulam os processos físicos intermediários através de equações parametrizadas ou não. Como modelos dessa categoria têm-se os modelos hidrológicos de transformação chuva-vazão ou de propagação do escoamento em rios. Os modelos conceituais possuem a vantagem de se adequarem melhor às condições fora da faixa de valores para os quais foram ajustados e estarem menos sujeitos a gerar previsões incoerentes. Por outro lado, os modelos empíricos possuem estruturas mais simples e exigem menor quantidade de dados. Por possuírem ajustes mais simples, os modelos empíricos podem ser recalibrados com maior frequência (Silva et al., 2006), reduzindo a chance do surgimento de erros sistemáticos durante a operação.

Os dados de entrada para os modelos de previsão dependem do tipo de modelo empregado, mas os principais são a chuva e a vazão, ou nível d'água. A variável chuva pode ser obtida através de estações de monitoramento instaladas na bacia hidrográfica ou através de estimativas de radar ou satélite. Adicionalmente, podem ser utilizadas previsões de precipitação de modelos atmosféricos, que possibilitam gerar previsões com maior horizonte (usualmente até 14 dias a frente, se a variável for obtida de modelos de previsão de tempo). No caso da utilização de chuva observada, a previsão se estende a poucas horas e é dependente do tempo de concentração da bacia; porém, os erros de previsão são menores em comparação com o uso de chuva prevista (Silva et al., 2006). Para se obterem erros ainda menores na previsão, podem ser utilizados modelos do tipo vazão-vazão, ou nível-nível. Nesse caso, são utilizados dados de estações fluviométricas localizadas ao longo do rio principal e seus afluentes. A desvantagem é que as antecedências possíveis são ainda menores, uma vez que dependem do tempo de deslocamento da onda de cheia ao longo do trecho de rio monitorado.

No Brasil, ainda são poucas as bacias hidrográficas que possuem um sistema de previsão em operação e, principalmente, com emissão de alertas de inundação. O setor de geração hidrelétrica é o que tradicionalmente mais utiliza previsões de vazão, mas para fins de geração de energia (Santos & Oliveira, 2016; Fan et al., 2017; Silveira et al.,

2017), através de modelos tanto empíricos quanto conceituais. Visando a produção de alertas de inundações, a Companhia de Pesquisa de Recursos Minerais (CPRM, 2018) monitora atualmente 16 bacias em todo país. Para algumas dessas bacias são geradas previsões, como no caso dos rios Amazonas e Doce, mas para a maioria apenas são monitoradas as condições dos níveis dos rios, que são divulgadas no sítio da CPRM e com uma classificação de alerta. Alguns trabalhos relatam as experiências de previsão e alerta em bacias brasileiras, tais como: Reis et al. (2016); Negrão et al. (2017); Rodriguez et al. (2017); Souza et al. (2017).

Um rio que necessita de monitoramento contínuo é o rio Sapucaí que está localizado na bacia do Rio Grande no sul do Estado de Minas Gerais (Figura 1). O rio Sapucaí nasce na serra da Mantiqueira, na cidade de Campos do Jordão (SP) e percorre cerca de 343 km até desaguar no lago da usina hidrelétrica de Furnas (IGAM, 2010). Ao longo do seu percurso, o rio cruza a área urbana do município de Itajubá (que é o foco deste trabalho) que, constantemente, é afetado por inundações. Pinheiro (2005) apresentou um histórico das inundações ocorridas em Itajubá desde a sua fundação em 1819 até o ano de 2000. Barbosa et al. (2015) realizaram um estudo similar, mas estendido até 2011. Esses trabalhos revelam que os episódios de inundações se concentram entre os meses de outubro a março, que correspondem ao período chuvoso da região sudeste do Brasil (Reboita et al., 2010, 2017). Nesse período, as chuvas estão associadas principalmente à convecção e à influência da Zona de Convergência do Atlântico Sul – ZCAS (Reboita et al., 2010; 2017). A ZCAS é uma banda persistente de nuvens e precipitação que se estende do sul da Amazônia ao sudeste do Brasil e oceano Atlântico (Kodama, 1992). Esse sistema pode causar precipitação durante dias consecutivos bem como chuvas intensas em curto período de tempo, ambos podendo sobrecarregar a bacia do Rio Grande. Como as águas dessa bacia escoam pelo rio Sapucaí podem causar problemas associados a inundações em Itajubá, já que o rio atravessa a área urbana desse município (Figura 2). Reboita et al. (2017) relatam que as inundações ocorridas em 16 de janeiro de 1991 e 02 de janeiro de 2000 foram devido à atuação da ZCAS.

O primeiro ponto de inundação em Itajubá ocorre no bairro Santa Rosa, local onde se inicia o trajeto do rio pela área urbana. Entre os anos de 2009 e 2015 foram coletados dados de nível em alguns pontos ao longo do rio Sapucaí, inclusive no bairro Santa Rosa (Marciano et al., 2018). Com base em alguns períodos desses dados, Reis et al. (2016) foram os primeiros autores a propor a elaboração de um modelo de previsão do nível fluviométrico em Santa Rosa baseado na técnica de regressão polinomial cúbica. Entretanto, esses autores focaram apenas a fase de ascensão da onda de cheia não se preocupando com outras características do nível em tal localidade. Dentro deste contexto, o presente trabalho tem como objetivo o desenvolvimento de modelos empíricos de previsão de nível na estação fluviométrica localizada no bairro Santa Rosa em Itajubá-MG (e, por esse motivo, recebe a denominação estação Santa Rosa), a fim de antecipar eventos de inundações na área urbana. Para tanto, são utilizadas medições de nível d'água em estações fluviométricas localizadas no curso principal do rio Sapucaí e em seus afluentes, que servem para o ajuste de diferentes modelos através da técnica de regressão.

## 2 Metodologia

### 2.1 Área de Estudo e Dados

O presente estudo destina-se a prever o nível do rio Sapucaí ao passar pelo bairro Santa Rosa (Figura 2) em Itajubá – MG. Tal município localiza-se no sul de MG (Figura 1) e, segundo estimativas do IBGE (2017), abriga cerca de 97.000 habitantes. É caracterizado como um dos municípios mais importantes da região sul de MG, devido ao seu parque industrial e educacional. Grande parte da sua população com menor poder aquisitivo reside nas planícies do rio Sapucaí, estando sujeita a episódios frequentes de inundação.

Os dados de nível fluviométrico utilizados no estudo foram obtidos do Sistema de Monitoramento de Enchentes (SME) do Rio Sapucaí que operou entre os anos de 2009 e 2015 (Marciano et al., 2018). O SME possuía 18 estações, sendo 7 localizadas a montante de Itajubá. Entretanto, para este trabalho foram utilizadas somente 4 estações fluviométricas, por serem as que registravam dados dos rios que

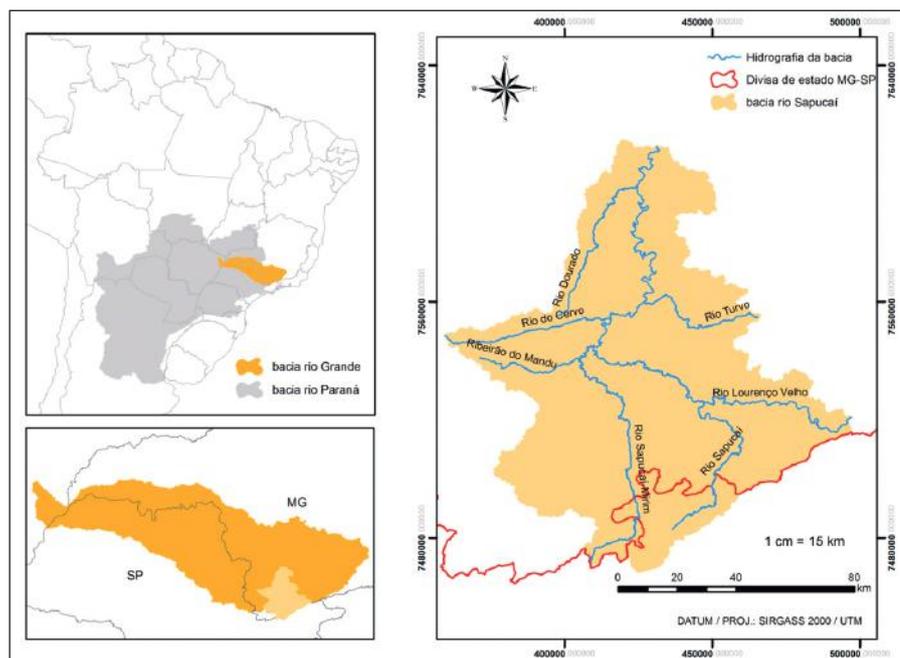


Figura 1 Localização da bacia do rio Sapucaí. Fonte: Extraído de Siqueira (2017).

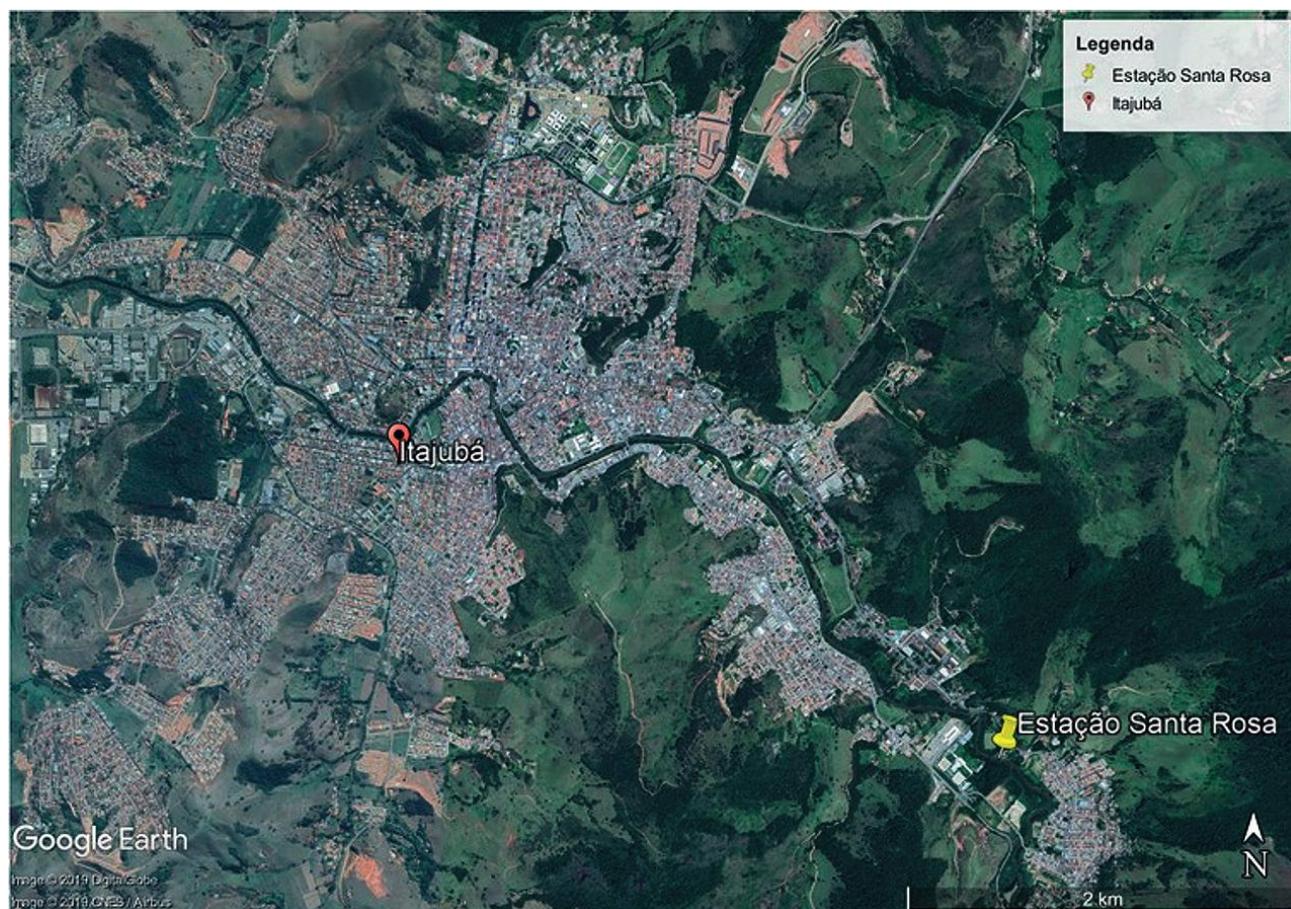


Figura 2 Vista aérea do rio Sapucaí atravessando a cidade de Itajubá – MG. O bairro Santa Rosa aparece no canto inferior direito da imagem com a localização da estação fluviométrica de mesmo nome. A sede administrativa do município é indicada com o marcador vermelho. Fonte: Google Earth 2019.

influenciam diretamente a área urbana de Itajubá (Figura 3). Atualmente, a prefeitura de Itajubá mantém convênio com a Universidade Federal de Itajubá (UNIFEI) em um projeto intitulado Sistema de Previsão, Monitoramento, Modelagem e Mitigação de Eventos Extremos na Bacia Hidrográfica do Alto Rio Sapucaí – SP3M, sendo mantidas 6 estações fluviométricas (meteorologia.unifei.edu.br).

A Figura 3 apresenta a localização das estações fluviométricas utilizadas no estudo: estação Santa Rosa (22,45° S e 45,42° W), localizada no rio Sapucaí e dentro da área urbana de Itajubá; estação Água Limpa (22,47° S, 45,36° W), localizada no rio Santo Antônio, afluente do Sapucaí; estação Santana (22,50° S, 45,23° W), localizada no rio Bicas, afluente do rio Sapucaí; e estação São Pedro (22,51° S, 45,39° W), localizada no rio Sapucaí

As estações do SME possuíam transmissão via sinal de celular, sendo a frequência de transmissão em condições normais de 30 minutos. Em caso de ocorrência de elevação do nível do rio, automaticamente a frequência se alterava para 10 minutos.

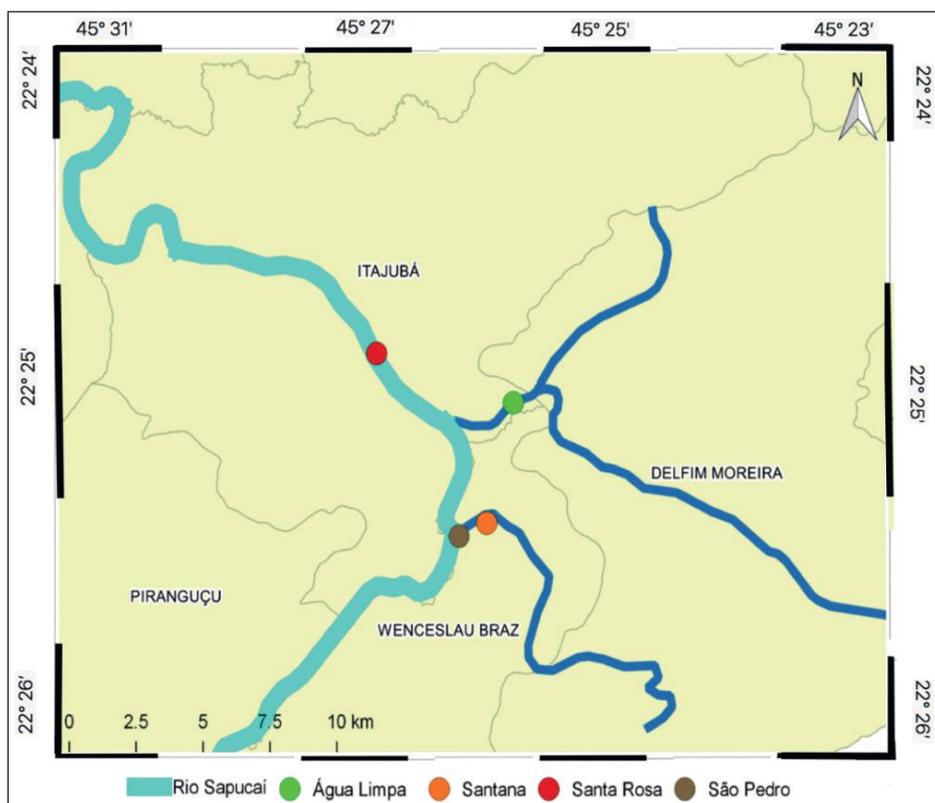
Para o ajuste dos modelos de previsão propostos neste estudo foram utilizados dados em intervalos temporais de 1 hora, uma vez que o tempo de propagação das vazões entre as estações de montante e a cidade de Itajubá é sempre maior do que esse valor (Reis et al., 2016).

As séries temporais do nível fluviométrico passaram inicialmente por um controle de qualidade que envolveu a identificação de erros e falhas. Em virtude do grande percentual de falhas e algumas inconsistências identificadas nos dados, o presente estudo teve que se restringir ao período de janeiro a dezembro de 2013, o qual os dados quase não apresentaram inconsistências.

## 2.2 Ajuste e Validação dos Modelos Preditivos

Como o objetivo deste estudo é obter modelos preditivos para a cota fluviométrica na estação Santa Rosa, a metodologia foi constituída de duas etapas: a obtenção de modelos preditivos e a validação desses. Para tanto, as séries temporais foram divididas em dois períodos: um para a obtenção dos modelos

Figura 3 Localização das estações fluviométricas utilizadas no estudo e do rio Sapucaí (azul claro), seus afluentes (azul escuro) e da divisão de municípios.



de regressão numérica, que descrevessem o comportamento das séries temporais (fase de ajuste ou calibração), e outro para aplicação dos modelos, a fim de validá-los (fase de validação). Para o ajuste dos modelos matemáticos utilizaram-se dados fluviométricos do mês de janeiro de 2013, enquanto que para a validação (comparação entre os valores estimados e observados) foram usados dados de fevereiro a dezembro de 2013.

### 2.2.1 Modelos de Regressão

Existem diversos tipos de regressões numéricas que podem ser utilizadas para descrever um conjunto de dados (Hair et al., 2009; Lohmann, 2011). O presente estudo testou as regressões lineares simples e múltiplas, além da regressão polinomial.

Basicamente, a regressão linear simples leva em consideração os dados de uma única estação fluviométrica a montante, enquanto que a regressão múltipla utiliza dados de várias estações. Considerando que as previsões devem ser feitas para a estação Santa Rosa (SR) com base em dados de numa única estação, por exemplo a Água Limpa (AL), a equação de regressão linear simples pode ser escrita de acordo com a Equação 1. O mesmo pode ser feito considerando as demais estações a montante, Santana (SA) e São Pedro (SP). Já na regressão linear múltipla, utilizam-se duas ou mais estações e a equação obtida pode ser expressa conforme a Equação 2.

$$SR(t+h) = A_0 + AL(\text{nível}(t)) \quad (1)$$

onde: SR (t+h) indica o nível previsto (calculado) na estação Santa Rosa, considerando um intervalo de tempo t+h;  $A_0$  é um coeficiente linear de ajuste; AL é um coeficiente angular que indica o ângulo de inclinação da reta ajustada pelo modelo e (nível(t)) caracteriza o nível do rio observado na estação de Água Limpa no instante t.

$$SR(t+h) = A_0 + AL(\text{nível}(t)) + SA(\text{nível}(t)) + SP(\text{nível}(t)) \quad (2)$$

onde  $A_0$  é um coeficiente de ajuste; AL, SA e SP são coeficientes angulares obtidos para as estações Água Limpa, Santana e São Pedro respectivamente, e nível(t) é o nível do rio observado em cada uma das estações no instante t.

$$SR(t+h) = A_0 + [A_1 * AL(\text{nível}(t))] + [A_2 * AL(\text{nível}(t))^2] + [A_3 * AL(\text{nível}(t))^3] \quad (3)$$

sendo a descrição das variáveis similares às equações anteriores.

Na utilização da técnica de regressão linear, é importante verificar a significância estatística do modelo obtido (equação). Para isso, é utilizada a distribuição F que testa a hipótese de que nenhum dos coeficientes de regressão tenha significado (Lapponi, 2000). Nessa estatística é obtido o p-valor e, no caso, de se considerar a significância de 0,05, os valores de p-valor < 0,05 indicarão que o modelo de regressão pode ser aceito.

O modelo de regressão polinomial leva em consideração o nível de uma única estação fluviométrica para o cômputo do valor previsto na estação de SR, porém como engloba mais 2 coeficientes possui ordem polinomial 3. Esse modelo é escrito como na Equação 3.

### 2.2.2 Estatísticas

Para identificar a hora (ou horas) de defasagem com maior relação entre o nível fluviométrico das diferentes estações e o nível na estação Santa Rosa, foi calculado o coeficiente de correlação de Pearson (r):

$$r = \frac{\sum(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{(\sum(x_i - \bar{x})^2)(\sum(y_i - \bar{y})^2)}} \quad (4)$$

onde  $x_i$  e  $y_i$  são os dados de duas séries temporais, enquanto que  $\bar{x}$  e  $\bar{y}$  correspondem ao valor médio das respectivas séries temporais. O coeficiente r assume valores entre -1 e 1. Quando  $r=1$  a correlação entre duas variáveis é perfeita, quando  $r=-1$  há uma correlação perfeita, mas de forma inversa, e quando  $r=0$  significa que não dá dependência linear entre as variáveis.

Para o cálculo do coeficiente de correlação as séries temporais devem possuir o mesmo número de dados. Quando a correlação é calculada sem defasagem, isso significa que a primeira linha do conjunto

de dados possui a informação do tempo  $t$  de ambas as séries temporais, a segunda linha do tempo  $t+1$  e, assim, por diante. A correlação defasada significa deslocar os dados de uma das séries temporais. Como no presente estudo deseja-se relacionar o nível de três estações fluviométricas com o nível da estação Santa Rosa, a série temporal da estação Santa Rosa não será deslocada temporalmente, enquanto que a das outras sim. Por exemplo, para uma correlação defasada em 3 horas, a primeira linha do conjunto de dados da estação Santa Rosa deve possuir informação do tempo  $t$  enquanto a da outra estação do tempo  $t+3$ . Nota-se que os três primeiros tempos foram excluídos da série temporal. Como o número de dados deve ser igual em ambas as séries para o cálculo do coeficiente de correlação, eliminam-se, os três tempos finais da série temporal da estação Santa Rosa. O mesmo procedimento é aplicado para defasagens com diferentes tempos.

Uma vez aplicada a técnica de regressão (Hair et al., 2009) e obtidos os modelos de previsão (equações de regressão) é importante avaliar a qualidade do ajuste desses modelos. Para isso é calculado o coeficiente de determinação ( $R^2$ ). De acordo com Stevenson (2001), a equação para obter  $R^2$  é:

$$R^2 = 1 - \frac{[\sum(P_i - \bar{P})^2]/(N-2)}{[\sum(P_i - P)^2]/(N-2)} \quad (5)$$

onde  $P_i$  são os valores previstos por um modelo de regressão,  $\bar{P}$  é a média da série dos valores previstos,  $P_c$  são os valores que compõem a reta de regressão e  $N$  é o número de dados da série  $z$ .  $R^2$  pode variar de 0 a 1; quanto mais próximo de 1 indica um melhor ajuste do modelo aos dados observados. Também é válido lembrar que o  $R^2$  pode ser obtido simplesmente pelo coeficiente de correlação elevado ao quadrado.

Na fase de validação do desempenho dos modelos, foram computados dois coeficientes estatísticos: o erro médio (EM) e o coeficiente de eficiência de Nash Sutcliffe (NS):

$$EM = \frac{1}{N} \sum_{i=1}^N (P_i - O_i) \quad (6)$$

em que:  $P_i$  é a série de dados modelada (prevista com as equações de regressão),  $O_i$  é a série de dados observada e  $N$  é o número de dados da série temporal.

O NS é um dos critérios estatísticos mais importantes utilizados para avaliar a performance de modelos hidrológicos (Machado & Vettorazzi, 2003). De acordo com Silva et al. (2008), se  $NS > 0,75$  o desempenho de um modelo é considerado bom, se  $0,75 \geq NS \geq 0,36$  o desempenho é aceitável e se  $NS < 0,36$  o desempenho é considerado inaceitável.

### 3 Resultados

Os resultados serão apresentados em duas seções: a primeira, chamada de fase de ajuste, enfoca a obtenção dos modelos de regressão usando um período específico da série de dados (06/01/2013 a 31/01/2013) enquanto a segunda, chamada de fase de validação, avalia a performance dos modelos obtidos utilizando um outro período da série temporal (01/02/2013 a 31/12/2013).

#### 3.1 Fase de Ajuste

A Figura 4 mostra o nível fluviométrico por hora, no período de 06/01/2013 a 31/01/2013, medido nas estações Santa Rosa, Água Limpa, Santana e São Pedro. Há duas informações relevantes na figura: ocorrência de três episódios de maior aumento do nível do rio no período em estudo (dias 11/01/2013, 14/01/2013 e 17/01/2013) e que o aumento do nível na estação Santa Rosa ocorre após ao das demais estações em cerca de poucas horas.

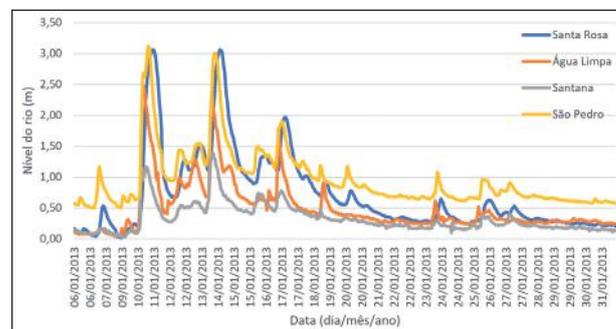


Figura 4 Nível fluviométrico (m) observado por hora nas estações de Santa Rosa (linha azul), Água Limpa (linha laranja), Santana (linha cinza) e São Pedro (linha amarela) no período de 06/01/2013 a 31/01/2013.

A análise mais detalhada da Figura 4 revela que o tempo de defasagem entre o nível na estação Santa Rosa e o das demais varia, em geral, de 3 a 7

horas. Esses diferentes tempos de defasagem ocorrem em função de características relacionadas ao escoamento ao longo do rio: a) em condições de níveis mais elevados o escoamento tende a ser mais rápido do que em níveis menores; b) extravasamentos para as planícies laterais tendem a reduzir a velocidade e o pico a jusante e c) mudanças na calha principal do rio, seja no sentido de obstruí-la ou desobstruí-la, também irão alterar as defasagens.

Devido aos motivos descritos, para verificar o tempo (ou os tempos) de defasagem mais relevante, foi calculado o coeficiente de correlação de Pearson ( $r$ ) entre o nível fluviométrico na estação Santa Rosa com o nível das outras três estações. Para isso, a correlação foi computada considerando a hora atual na estação Santa Rosa com a informação medida a 3, 4, 5, 6 e 7 horas antes nas demais estações. A Tabela 1 mostra que todas as correlações foram superiores a 0,90 e mais elevadas quando considerada a estação São Pedro. Para essa estação, a maior correlação ( $r = 0,98$ ) ocorreu para 3 e 4 horas de defasagem. Considerando a Água Limpa, o maior valor de correlação ( $r = 0,94$ ) foi obtido para 4, 5 e 6 horas de defasagem e para Santana o maior ( $r = 0,96$ ) foi obtido para 4 e 5 horas. Nota-se, portanto, que todas as estações concordam que há maior correlação para 4 horas de defasagem. Em outras palavras, o aumento no nível fluviométrico na estação Santa Rosa poderá ocorrer a cerca de 4 horas após ao registrado nas demais estações.

Estações Fluviométricas	Horas de Defasagem				
	3 h	4 h	5 h	6 h	7 h
Água Limpa	0,93	0,94	0,94	0,94	0,92
Santana	0,95	0,96	0,96	0,95	0,94
São Pedro	0,98	0,98	0,97	0,95	0,92

Tabela 1 Coeficiente de correlação de Pearson ( $r$ ) obtido entre o nível fluviométrico (m) medido na estação Santa Rosa e nas demais estações, no período de 06/01/2013 a 31/01/2013, considerando diferentes tempos de defasagem.

Para se obterem modelos (equações) para a previsão do nível fluviométrico na estação Santa Rosa, a técnica de regressão foi utilizada. Embora já se conheça que há correlação alta entre o nível fluviométrico na estação Santa Rosa com o nível registrado com quatro horas de antecedência nas

demais estações, a técnica de regressão foi aplicada para cinco horários de defasagem (de 3 a 7 horas) e considerando a regressão linear simples, a regressão linear múltipla e regressão polinomial cúbica. Em todos os casos, a estação Santa Rosa é a variável a ser prevista ou o que também pode ser chamado de variável dependente. Já as demais estações são as variáveis preditoras, o que também pode ser chamado de variáveis independentes.

Foram obtidos 15 modelos com a regressão linear simples, 20 com a regressão linear múltipla e 15 com a regressão polinomial cúbica. Por brevidade, a Tabela 2 mostra os modelos obtidos com regressão linear considerando apenas a defasagem de 4 horas; já a Tabela 3 mostra todos os modelos obtidos com a regressão polinomial cúbica. No caso dos modelos obtidos com a regressão linear, todas as estatísticas do p-valor foram inferiores ao nível de significância de 0,05, indicando que os modelos podem ser aceitos.

Estações Fluviométricas	Equações
Água Limpa	$SR = -0,01726 + AL (1,48093)$
Santana	$SR = -0,16779 + SA (2,676957)$
São Pedro	$SR = -0,56627 + SP (1,319586)$
Água Limpa + Santana	$SR = -0,13923 + AL (0,512081) + AS (1,822582)$
Água Limpa + São Pedro	$SR = -0,46624 + AL (0,357456) + SP (1,034541)$
Santana + São Pedro	$SR = -0,4669 + SA (0,845068) + SP (0,930247)$
Água Limpa + Santana + São Pedro	$SR = -0,43381 + AL (0,2129) + SA (0,619834) + SP (0,864244)$

Tabela 2 Equações prognósticas do nível fluviométrico na estação de Santa Rosa com um intervalo de antecedência de 4 horas obtidas com a técnica de regressão linear simples e múltipla. As iniciais SR, AL, SA e SP indicam o nível fluviométrico nas estações Santa Rosa, Água Limpa, Santana e São Pedro, respectivamente.

De forma resumida, a aplicação da técnica de regressão ajusta uma reta ou uma curva nos dados e fornece uma equação para tais formas geométricas. Uma maneira de avaliar a qualidade do ajuste obtido é através do coeficiente de determinação ( $R^2$ ), que aqui foi computado com a Equação 5. As Tabelas 4

Estações	Intervalo de Tempo (horas)	Equações
	3	$SR = 10,4326 + 1,0058(AL)^2 + 1,0739(AL) \parallel 0,0047$
	4	$SR = 10,3681 + 0,8814 + 1,1066 \parallel 0,0033$
Água Limpa	5	$SR = 10,2877 + 0,6884 + 1,1937 \parallel 0,0083$
	6	$SR = 10,2026 + 0,4671 + 1,3017 \parallel 0,0145$
	7	$SR = 10,1108 + 0,2136 + 1,4331 \parallel 0,0223$
	3	$SR = 13,1913 + 5,4345 + 0,3451 + 0,0736$
	4	$SR = 13,0631 + 5,4355 + 0,2354 + 0,0951$
Santana	5	$SR = 12,8684 + 5,264 + 0,2051 + 0,1111$
	6	$SR = 12,6479 + 4,999 + 0,2177 + 0,1255$
	7	$SR = 12,404 + 4,6478 + 0,273 + 0,1373$
	3	$SR = 10,2154(SP)^3 + 0,8717(SP)^2 + 0,4024(SP) \parallel 0,304$
	4	$SR = 10,1887(SP)^3 + 0,7669(SP)^2 + 0,5089(SP) \parallel 0,3334$
São Pedro	5	$SR = 10,1491(SP)^3 + 0,592(SP)^2 + 0,7081(SP) \parallel 0,3939$
	6	$SR = 10,1041(SP)^3 + 0,3837(SP)^2 + 0,952(SP) \parallel 0,4684$
	7	$SR = 10,0607(SP)^3 + 0,1742(SP)^2 + 1,2016(SP) \parallel 0,544$

Tabela 3 Equações prognósticas do nível fluviométrico na estação de Santa Rosa com os diferentes intervalos de antecedência obtidas com a regressão polinomial cúbica. As iniciais SR, AL, SA e SP indicam o nível fluviométrico nas estações Santa Rosa, Água Limpa, Santana e São Pedro, respectivamente.

e 5 mostram os valores de R<sup>2</sup> obtidos na regressão linear e polinomial, respectivamente. Quanto mais próximo de 1 for o R<sup>2</sup>, melhor é o ajuste do modelo em relação aos dados observados. A maioria dos modelos obtidos com a regressão linear forneceram valores de R<sup>2</sup> superiores a 0,9, indicando um ótimo ajuste, mas quando a variável independente foi a defasada em 4 horas, os valores de R<sup>2</sup> ainda foram ligeiramente maiores (Tabela 4). Resultados similares foram obtidos com o uso da regressão polinomial (Tabela 5).

É importante enfatizar que r e R<sup>2</sup> estão relacionados uma vez que R<sup>2</sup> também pode ser obtido através do coeficiente de correlação r elevado ao quadrado. Portanto, seria desnecessário apresentar o R<sup>2</sup> dos modelos obtidos com a regressão linear simples na Tabela 4, uma vez que a Tabela 1 mostrou os valores de r. Por outro lado, preferiu-se mostrar o R<sup>2</sup> de todos os modelos a fim de deixar a Tabela 4 completa. Ao se comparar os valores da Tabela 1 elevados ao quadrado com os da Tabela 4 podem ocorrer

pequenas diferenças entre os valores, fato que está associado a utilização da Equação 5 para cálculo do R<sup>2</sup> e não do coeficiente de correlação.

Estações Fluviométricas	Horas de Defasagem				
	3 h	4 h	5 h	6 h	7 h
Água Limpa	0,86	0,89	0,89	0,88	0,85
Santana	0,90	0,92	0,93	0,91	0,87
São Pedro	0,95	0,96	0,94	0,90	0,85
Água Limpa + Santana	0,91	0,94	0,94	0,92	0,89
Água Limpa + São Pedro	0,96	0,97	0,95	0,92	0,88
Santana + São Pedro	0,96	0,97	0,96	0,93	0,89
Água Limpa + Santana + São Pedro	0,96	0,97	0,96	0,93	0,89

Tabela 4 R<sup>2</sup> obtido nos modelos de regressão linear. Os maiores valores encontrados em cada modelo estão destacados em negrito.

Estações Fluviométricas	Horas de defasagem				
	3 h	4 h	5 h	6 h	7 h
Água Limpa	0,90	0,91	0,91	0,89	0,85
Santana	0,92	0,94	0,94	0,92	0,89
São Pedro	0,97	0,97	0,95	0,91	0,86

Tabela 5 R<sup>2</sup> obtido nos modelos de regressão polinomial cúbica. Os maiores valores encontrados em cada modelo estão destacados em negrito.

A partir dos modelos apresentados nas Tabelas 2 e 3, pode-se prever o nível fluviométrico na estação Santa Rosa. Embora a validação dos modelos (Seção 3.2) seja realizada com um período da série temporal diferente do usado na fase de ajuste, aqui, apenas a fim de ilustração, também são apresentadas as séries simuladas no mesmo período dos dados usados na construção das equações. A Figura 5 é um exemplo usando o modelo de regressão linear simples em que a variável independente é o nível fluviométrico na estação São Pedro com defasagem de 4 horas (Tabela 2). A figura também mostra o erro obtido entre os dados observados e modelados e que, em geral, esse é de cerca de  $\pm 0,2$  m. Já a Figura 6 é um exemplo usando o modelo de regressão linear múltipla considerando como variáveis independentes Água Limpa, Santana e São Pedro com defasagem de 4 horas. Nesse exemplo, as maiores diferenças entre as séries observada e modelada ocorrem durante os maiores níveis registrados (dias 11 e 14

de janeiro de 2013). Algumas estatísticas calculadas entre o nível observado e simulado com os diferentes modelos de regressão considerando as variáveis independentes com defasagem de 4 horas são mostradas na Tabela 6.

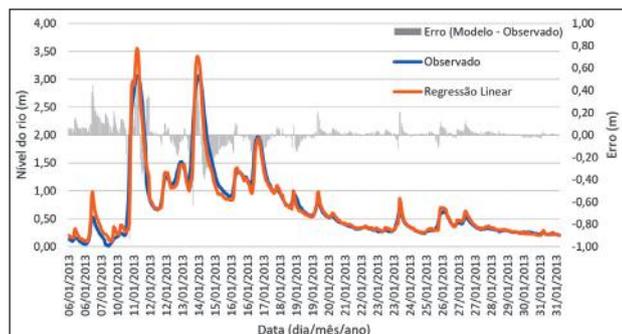


Figura 5 Nível fluviométrico observado (linha azul) e modelado (linha laranja) na estação Santa Rosa pela regressão linear simples (variável independente nível na estação São Pedro com defasagem de 4 horas) durante o período de calibração que foi de 06/01/2013 a 31/01/2013. O eixo vertical da esquerda indica o nível fluviométrico (m) e o da direita, a diferença (erro) obtida entre as duas séries (modelada – observada).

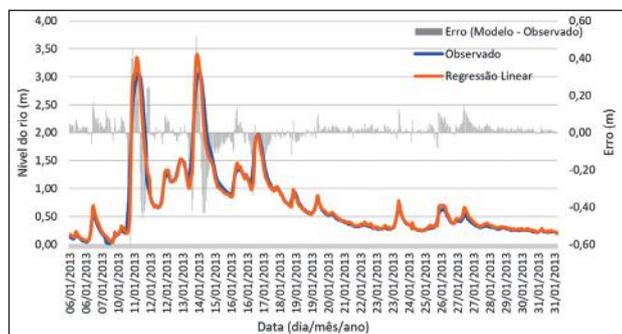


Figura 6 Nível fluviométrico simulado na estação Santa Rosa com o modelo de regressão linear múltipla com três variáveis independentes considerando a defasagem de 4 horas. O eixo vertical da esquerda indica o nível fluviométrico (m) e o da direita, a diferença (erro) obtida entre as duas séries (modelada – observada).

### 3.2 Fase de Validação

A segunda etapa do estudo corresponde à validação dos modelos de previsão usando um período da série temporal distinto ao usado na calibração. Portanto, o período de validação foi de 01/02/2013 a 31/12/2013. As análises dessa seção focam as simulações considerando como variável independente o nível fluviométrico com defasagem de 4 horas. Inicialmente, a fim de ilustração, são mostradas algumas das séries modeladas. A Figura 7 apresenta a comparação do nível fluviométrico observado na

estação Santa Rosa com o previsto com o modelo de regressão linear múltipla considerando como variáveis independentes o nível nas estações Água Limpa e São Pedro com defasagem de 4 horas e a regressão polinomial cúbica considerando o nível na estação São Pedro, também com defasagem de 4 horas. Já a Figura 8 mostra a comparação com a série simulada com o modelo de regressão linear múltipla com três variáveis independentes (de Água Limpa, Santana e São Pedro) e defasagem de 4 horas. As duas figuras também apresentam as diferenças entre os dados modelados e observados e, da mesma forma que na seção anterior, essas diferenças variam entre  $\pm 0,2$  m. Durante os períodos em que o nível fluviométrico está baixo, os modelos, em geral, subestimam ligeiramente o nível observado. As maiores diferenças (em módulo) ocorrem quando são registrados os maiores níveis fluviométricos.

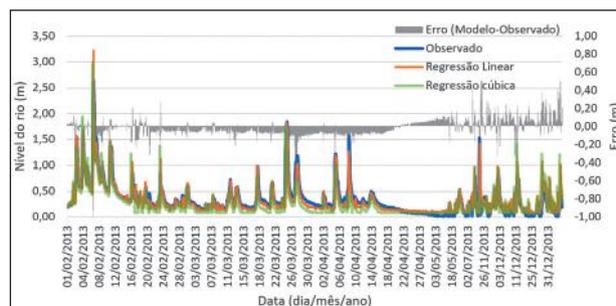


Figura 7 Nível fluviométrico observado (linha azul) e modelado na estação Santa Rosa pela regressão linear múltipla (variáveis independentes são o nível nas estações Água Limpa e São Pedro com defasagem de 4 horas – linha laranja) e cúbica (variável independente São Pedro com defasagem de 4 horas, linha verde) durante o período de validação que foi de 01/02/2013 a 31/12/2013. O eixo vertical da esquerda indica o nível fluviométrico (m) e o da direita, a diferença (erro) obtida entre a série observada e modelada com a regressão linear (modelada – observada).

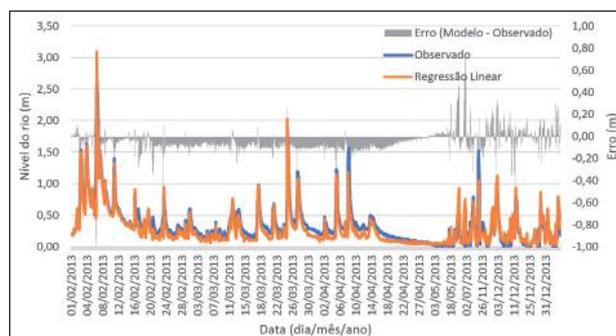


Figura 8 Nível fluviométrico observado (linha azul) e modelado na estação Santa Rosa considerando o modelo de regressão linear múltipla com três variáveis independentes (de Água Limpa, Santana e São Pedro) e defasagem de 4 horas.

A performance dos modelos em simular o nível fluviométrico é indicada através de algumas estatísticas. Dessa forma, foram calculados o coeficiente de correlação, o erro médio e coeficiente de eficiência de NS entre o nível fluviométrico observado na estação Santa Rosa e o simulado com os modelos obtidos com a regressão linear (simples e múltipla) e a regressão polinomial cúbica considerando as séries das variáveis independentes defasadas em 4 horas (Tabela 6). Considerando a regressão linear, os modelos que produziram o nível fluviométrico mais próximo do observado foram os obtidos com regressão linear múltipla e quando considerado no cálculo as variáveis independentes Santana e São Pedro, pois as correlações foram iguais ou maiores do que 0,9, o erro médio variou entre -0,09 e -0,05 e o coeficiente de NS foi superior a 0,7 (Tabela 6). Utilizando a regressão polinomial cúbica, o modelo que produziu dados mais similares aos observados foi o que utilizou o nível fluviométrico da estação São Pedro ( $r=0,88$  e  $NS=0,6932$ ). Embora o modelo que utilizou a estação Água Limpa tenha indicado boa correção ( $r=0,83$ ), o coeficiente de eficiência de NS foi pequeno ( $NS=0,1824$ ). De acordo com Silva et al. (2008), a performance de um modelo é considerada aceitável quando NS apresenta valores entre 0,36 e 0,75 e bom para valores maiores do que 0,75.

É importante mencionar que as estações Santana e São Pedro se encontram entre o rio Sapucaí e um tributário enquanto a Água Limpa em outro tributário. Portanto, para o monitoramento do nível na estação Santa Rosa, é necessário a previsão com as equações que utilizam tais variáveis de forma independente ou

com regressão múltipla. Destaca-se, ainda, que o modelo utilizando a estação São Pedro é o que apresenta os melhores resultados quando se analisam cada um individualmente. Isso ocorre porque registra as vazões da bacia contribuinte de maior área, entre as estações de montante. Por esse motivo, essa estação deve ser a principal no sistema de previsão, sendo recomendável que esteja sempre em operação.

Para finalizar a análise estatística, também se apresenta o coeficiente de correlação calculado para todos modelos de regressão linear usando variáveis preditoras com diferentes tempos de defasagem (Tabela 7). Fica evidente que as melhores correlações obtidas entre a observação e simulação ocorrem com os modelos usando variáveis defasadas em 4 horas em relação à estação Santa Rosa.

Estações	Horas de Defasagem				
	3 h	4 h	5 h	6 h	7 h
Água Limpa	0,81	0,80	0,78	0,74	0,69
Santana	0,74	0,75	0,74	0,72	0,70
São Pedro	0,88	0,88	0,85	0,82	0,78
Água Limpa + Santana	0,84	0,85	0,83	0,81	0,77
Água Limpa + São Pedro	0,93	0,94	0,92	0,88	0,82
Santana + São Pedro	0,90	0,90	0,87	0,83	0,78
Água Limpa + Santana + São Pedro	0,93	0,94	0,91	0,87	0,82

Tabela 7 Coeficiente de correlação obtido entre o nível fluviométrico observado na estação Santa Rosa e o simulado com modelos de regressão linear usando variáveis com diferentes tempos de defasagem entre 01/02/2013 a 31/12/2013. O maior coeficiente de correlação obtido em cada modelo é destacado em negro.

Modelos	Ajuste		Validação		
	R	NS	r	EM	NS
$SR = -0,01726 + (AL \cdot 1,48093)$	0,94	0,8869	0,80	0,18	0,0396
$SR = -0,16779 + (SA \cdot 2,676957)$	0,96	0,9233	0,75	-0,09	0,2032
$SR = -0,56627 + (SP \cdot 1,319586)$	0,98	0,9589	0,88	-0,08	0,7122
$SR = -0,13923 + (AL \cdot 0,512081) + (SA \cdot 1,822582)$	0,97	0,9351	0,85	-0,01	0,5877
$SR = -0,46624 + (AL \cdot 0,357456) + (SP \cdot 1,034541)$	0,98	0,9658	0,94	-0,03	0,8859
$SR = -0,4669 + (SA \cdot 0,845068) + (SP \cdot 0,930247)$	0,98	0,9474	0,90	-0,09	0,7216
$SR = -0,43381 + (AL \cdot 0,2129) + (SA \cdot 0,619834) + (SP \cdot 0,864244)$	0,98	0,9693	0,94	-0,05	0,8481
$SR = -0,3681(AL)^3 + 0,8814(AL)^2 + 1,1066(AL) - 0,0033$	0,96	0,9121	0,83	0,16	0,1824
$SR = -3,0631(SA)^3 + 5,4355(SA)^2 + 0,2354(SA) + 0,0951$	0,97	0,9443	0,70	-0,03	0,2641
$SR = -0,1887(SP)^3 + 0,7669(SP)^2 + 0,5089(SP) - 0,3334$	0,99	0,9708	0,88	0,03	0,6932

Tabela 6 Coeficiente de correlação (r), erro médio (EM; unidade metros) e coeficiente de eficiência de Nash-Sutcliffe (NS) calculados entre o nível fluviométrico (m) observado na estação Santa Rosa e o simulado com os modelos obtidos com a regressão linear e cúbica considerando as séries defasadas em 4 horas.

Esta seção é finalizada com uma comparação entre os resultados obtidos com três modelos considerando defasagem de 4 horas: um com o método da regressão linear múltipla usando as três estações fluviométricas como variáveis preditoras e os outros dois com a regressão polinomial cúbica usando a estação São Pedro como variável preditora. Entretanto, um desses modelos com regressão polinomial foi obtido de Reis et al. (2016), que consideraram na fase de ajuste apenas dados de nível elevado.

A análise foi realizada para o mês de fevereiro de 2013, que compreende um dos meses mais chuvosos na região de estudo e, conseqüentemente, de maiores registros do nível do rio. De fato, em fevereiro de 2013 ocorreram 4 eventos de aumento do nível fluviométrico como mostra a Figura 9. Verificaram-se algumas oscilações entre níveis menores (~0,5 m) até o nível mais alto registrado que foi de 2,64 m no dia 07/02/2013. Em geral, os três modelos representam com boa acurácia as oscilações no nível na estação Santa Rosa. Entretanto, o modelo obtido com a regressão linear múltipla consegue simular os máximos da série temporal de forma mais similar à observação. Além disso, o coeficiente de correlação para esse período de dados também mostra maior valor quando considerado o modelo obtido com a regressão linear (Figura 9).

Quando se monitora o nível de um rio em um determinado local não basta ter a informação do nível no instante  $t$ , também é importante conhecer o nível máximo do rio, pois a partir do momento em que o rio atinge esse nível, alterações de poucos centímetros podem significar o alagamento de extensas áreas. Reis et al. (2016) ressaltam que se o nível na estação Santa Rosa chegar a 3,2 metros, isso pode ser considerado estado de atenção enquanto que o nível de 4,2 metros é estado de alerta. Na Figura 9, o pico máximo observado chegou a 2,62 m enquanto que os valores previstos pelas regressões chegaram a cerca de 3 m. Como esse valor previsto é menor do que o limiar usado para indicar estado de atenção, numa prática operacional, ainda não seria emitido nenhum boletim à população.

#### 4 Conclusões

Este estudo obteve modelos de previsão para o nível fluviométrico na estação Santa Rosa localizada no percurso do rio Sapucaí ao entrar na área urbana

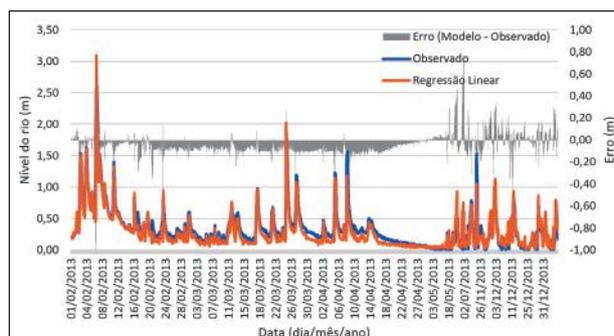


Figura 9 Nível do rio observado (linha azul) e modelado (linha laranja) na estação Santa Rosa pela regressão linear múltipla (Água Limpa + Santana + São Pedro) e regressão polinomial cúbica (linha amarela, estação de São Pedro) contendo o intervalo de 4 horas. A figura também inclui o resultado obtido com o modelo de regressão polinomial cúbico de Reis et al. (2016) (linha verde), que é  $SR(4h) = -0,0177 - 0,127(SP)^2 + 1,7982(SP) - 0,8146$ .

do município de Itajubá, MG. Para obtenção dos modelos, foram utilizados dados de três estações fluviométricas, consideradas como variáveis preditoras, e as técnicas de regressão linear simples e múltipla e de regressão polinomial cúbica. Verificou-se que:

- as equações de regressão obtidas são capazes de prever com grande acurácia o nível do rio observado na estação de Santa Rosa, sejam em condições de vazões baixas, médias ou elevadas;

- que os melhores resultados estiveram associados ao uso das variáveis preditoras com intervalos de 4 horas de antecedência em relação ao nível observado na estação Santa Rosa. Apesar do melhor resultado ser obtido para a defasagem de 4 horas, as equações para os demais tempos de antecedência (entre 3 e 7 horas) também podem ser aplicadas, uma vez que as estatísticas calculadas indicaram bons ajustes dos modelos preditivos. Embora os resultados desses possam apresentar erros ligeiramente maiores, ainda assim terão precisão suficiente para o planejamento e tomada de ações em situações de risco e

- que a combinação de mais de uma estação fluviométrica nos modelos possibilita um melhor prognóstico para o nível do rio na estação Santa Rosa. A combinação envolvendo as estações Água Limpa e São Pedro obteve o melhor  $R^2$  dentre todas as combinações possíveis. Portanto, o modelo obti-

do com a regressão linear múltipla é o melhor a ser usado na previsão de nível na estação Santa Rosa.

Por fim, destaca-se que se o estudo tivesse incluído um maior número de dados, o que não foi possível devido à inconsistência das séries temporais, as equações prognósticas poderiam ter mostrado ainda melhores resultados. Entretanto, como mencionado na introdução, os modelos de regressão podem ser continuamente atualizados com a inclusão de novos dados, e isso será uma meta do sistema de monitoramento de Itajubá, pois dados de nível começaram a ser registrados pelo novo sistema de monitoramento do projeto SP3M.

## 5 Agradecimentos

Os autores agradecem à FAPEMIG e CNPq pelo auxílio financeiro e a coordenação do SME pelo fornecimento das séries temporais de níveis das estações fluviométricas.

## 6 Referências

Alfieri, L. & Thielen, J. 2015. A European precipitation index for extreme rain-storm and flash flood early warning. *Meteorological Applications*, 22: 3-13.

Amaral, R. & Gutjahr, M.R. 2014. Cadernos de Educação Ambiental (n. 8): Desastres Naturais. Governo do Estado de São Paulo, Secretaria do Meio Ambiente, Instituto Geológico, 1ª reimpressão, 94p.

Barbosa, A.A.; Oliveira, G.M. & Oliveira, T.J. 2015. Histórico de Enchentes em Itajubá/MG. *Revista Meio Ambiente e Sustentabilidade*, 9(4): 125-140.

CPRM. 2018. Serviço Geológico do Brasil. SACE – Sistema de Alerta de Eventos Críticos. Disponível em: <<http://www.cprm.gov.br/sace/>>. Acesso: 15 jan. 2019.

Fan, F.M.; Pontes P.R.M.; Buarque D.C. & Collischonn, W. 2017. Evaluation of upper Uruguay river basin (Brazil) operational flood forecasts. *Brazilian Journal of Water Resources*, 22(37): online edition (without pages).

IBGE. 2017. Instituto Brasileiro de Geografia e Estatística. Estimativas populacionais para os municípios e para as Unidades da Federação brasileiros em 01.07.2017. Disponível em: <[https://ww2.ibge.gov.br/home/estatistica/populacao/estimativa2017/estimativa\\_dou.shtm](https://ww2.ibge.gov.br/home/estatistica/populacao/estimativa2017/estimativa_dou.shtm)>. Acesso em: 20 jan. 2019.

IGAM. 2010. Instituto Mineiro de Gestão das Águas. Plano Diretor de Recursos Hídricos da Bacia do Rio Sapucaí. Disponível em: < <http://www.cbhsapucaí.org.br/cbh/Pagina.do?idSecao=53>>. Acesso em: 20 jan. 2019.

Hair, J.F.; Black, W.C.; Babin, B.J.; Anderson, R.E. & Tatham, R.L. 2009. *Análise Multivariada de Dados*. Bookman, 6ªed., 688p.

Kodama, Y.M. 1992. Large-scale common features of subtropical precipitation zones (the Baiu frontal zone, the SPCZ, and the SACZ). Part I: Characteristics of subtropical frontal zones. *Journal of the Meteorological Society of Japan*, 70: 813–835.

Lapponi, J.C. 2000. *Estatística usando o Excel*. Lapponi Treinamento e Editora LTDA. 450p.

Lohmann, M. 2011. Regressão logística e redes neurais aplicadas à pre-

visão probabilística de alagamentos no município de Curitiba, PR. Programa de Pós-Graduação em Geografia. Universidade Federal do Paraná, Tese de Doutorado, 231p.

Machado, R.E. & Vettorazzi, C.A. 2003. Simulação da produção de sedimentos para a microbacia hidrográfica do Ribeirão dos Marins, SP. *Revista Brasileira de Ciência do Solo*, 27: 735-741.

Marciano, A.G.; Barbosa, A.A. & Silva, A.P.M. 2018. Estudo de cenários na simulação de eventos de cheia no rio Piranguçu e sua influência no distrito industrial de Itajubá – MG. *Revista Brasileira de Energias Renováveis*, 7: 1-15.

Merkuryeva, G.; Merkuryev, Y.; Boris V; Sokolov, B.V.; Potryasaev, S.; Zelenstov, V.A. & Lektuers, A. 2015. Advanced river flood monitoring, modelling and forecasting. *Journal of Computational Science*, 10: 77-85.

Negrão, A.C.; Rodriguez, D.A.; Chou, S.C.; Medeiros, G.S. & Camargo, C.P. 2017. Previsão de vazão de eventos extremos utilizando o modelo hidrológico distribuído MHD-INPE e previsões meteorológicas do modelo ETA-INPE para uma alta resolução espacial – estudo de caso: São Luís do Paraitinga (SP). *Revista Brasileira de Cartografia*, 69: 647-657.

Pinheiro, M.V. 2005. Avaliação técnica e histórica das enchentes em Itajubá – MG. Programa de Pós-Graduação em Engenharia de Energia, Universidade Federal de Itajubá, Dissertação de Mestrado. 104p.

Reboita, M.S.; Gan, M.A; Rocha, R.P. & Ambrizzi, T. 2010. Regimes de precipitação na América do Sul: uma revisão bibliográfica. *Revista Brasileira de Meteorologia*, 25(2):185-204.

Reboita, M.S.; Marietto, D.M.G.; Souza, A. & Barbosa, M. 2017. Caracterização atmosférica quando da ocorrência de eventos extremos de chuva na região sul de Minas Gerais. *Revista Brasileira de Climatologia*, 13(21): 20-37.

Reis, J.B.C., Pons, N.A.D. & Lopes, E.S.S. 2016. Monitoramento e alerta de inundação no município de Itajubá (MG) por regressão polinomial. *Geociências*, 35(1): 134-148.

Rodriguez, D.A., Lopes, L.G., Carriello, F., Siqueira Júnior, J.L., Pinto, G.L., Tomasella, J. & Chou, S.C. 2017. Previsões de cheias extremas nos horizontes sazonais e de curto e médio prazos na bacia do rio Madeira: estudo de caso da enchente de 30 de março de 2014 em Porto Velho. *Revista Brasileira de Cartografia*, 69: 827-836.

Santos, P.G.P. & Oliveira, T.F. 2016. Modelagem e previsão de vazão afluente média mensal no rio Tocantins, usina hidrelétrica Tucuruí-Pará, Amazônia, Brasil. *Biota Amazônia*, 6: 9-16.

Silva, B.C.; Tucci, C.E.M. & Collischonn, W. 2006. Previsão de Vazão com Modelos Hidroclimáticos. *Revista Brasileira de Recursos Hídricos*, 11: 15-29.

Silva, P.M.O.; Mello, C.R.; Silva, A.M. & Coelho, G. 2008. Modelagem da hidrografia de cheia em uma bacia hidrográfica da região Alto Rio Grande. *Revista Brasileira de Engenharia Agrícola e Ambiental*, 12: 258-265.

Silveira, C.S.; Alexandre, A.M.B.; Souza Filho, F.A.; Junior, F.C.V. & Cabral, S.L. 2017. Monthly streamflow forecast for National Interconnected System (NIS) using Periodic Autoregressive Endogenous Models (PAR) and Exogenous (PARX) with climate information. *Brazilian Journal of Water Resources*, 22: online.

Siqueira, P.C. 2017. Simulação hidrológica de cenários de uso do solo na bacia do rio Sapucaí. Programa de Pós-Graduação em Meio Ambiente e Recursos Hídricos, Universidade Federal de Itajubá, Dissertação de Mestrado. 83p.

Souza, A.S.; Curvello, A.M.L.; Souza, F.L.S. & Silva, H.J. 2017. A flood warning system to critical region. *Procedia Computer Science*, 109: 1104-1109.

Stevenson, W.J. 2001. *Estatística Aplicada à Administração*. São Paulo, Harbra, 495p.

Ye, J.; Shao, Y. & Li, Z. 2016. Flood forecasting based on TIGGE precipitation ensemble forecast. *Advances in Meteorology*, 2016, 1-9.