

# Modelando o Sistema Cantareira com Equações Diferenciais Estocásticas

Renato Mendes Coutinho\*, Roberto André Kraenkel†  
Instituto de Física Teórica - UNESP

Paulo Inácio de Knecht López de Prado‡  
Departamento de Ecologia, Instituto de Biociências - USP

## Águas Futuras

3 de julho de 2015

### Resumo

Este documento apresenta de forma resumida os métodos empregados para gerar projeções para a quantidade de água armazenada no sistema Cantareira, utilizadas no *site* **Águas Futuras** - <http://cantareira.github.io>.

## 1 A equação diferencial estocástica

Para descrever a dinâmica do volume de água no reservatório da Cantareira, escrevemos uma equação diferencial estocástica que representa o balanço de entradas e saídas de água no sistema, sujeito a ruído ambiental:

$$dV = (\alpha R^\beta V^\gamma - E)dt + \sigma V dW \quad (1)$$

Na equação acima,  $E$  é a retirada de água ( $m^3 s^{-1}$ ),  $R$  é a pluviosidade média nos 30 dias anteriores (mm),  $V$  é o volume de água armazenado no sistema ( $m^3$ ) e  $W$  é um processo de Wiener. O termo  $\sigma V dW$  expressa um ruído gaussiano instantâneo no volume armazenado, com média zero e desvio padrão proporcional ao volume. Assume-se, ademais, um erro de medida: os valores observados do volume seguem uma distribuição gaussiana com valor esperado igual ao valor esperado pela equação estocástica e um desvio padrão desconhecido  $\epsilon$ . As constantes  $\alpha$ ,  $\beta$ ,  $\gamma$  e  $\sigma$  são parâmetros a serem determinados no ajuste aos dados.

O termo  $\alpha R^\beta V^\gamma$  representa uma relação precipitação-vazão, ou seja, expressam quanto de água entra no sistema devido a uma precipitação dada por  $R$ . A forma desta relação utiliza a ideia que o próprio volume serve de substituto (*surrogate*) para a descrição estado do solo na área de bacia hidrográfica, na linha dos *data-based mechanistic models* discutidos por Beven, [1].

---

\*renatomc@ift.unesp.br

†kraenkel@ift.unesp.br

‡prado@ib.usp.br

## 2 Ajuste do Modelo

Os dados utilizados para ajustar o modelo são séries temporais de precipitação, volume armazenado, vazões afluentes e defluentes, dia-a-dia. São obtidos de *sites* públicos, e a os dados e seus detalhes podem ser obtidos no *site* **Águas Futuras**, na seção “Fontes”.

Fizemos a estimativa dos parâmetros em três etapas:

1. O erro de observação  $\epsilon$  é estimado a partir de ajuste de trajetórias (*trajectory matching*), [2].
2. O expoente  $\gamma$  foi estimado como a inclinação de uma regressão gaussiana linear da área do reservatório em função de seu volume, em escala logarítmica.
3. Os parâmetros  $\alpha$ ,  $\beta$  e  $\sigma$  foram estimados com um filtro de partículas bayesiano (*Bayesian Sequential Monte Carlo*), [3], aperfeiçoado por [2].

Inicialmente fizemos uma comparação entre o ajuste do modelo dado pela Equação 1 e de um modelo em que a vazão afluente não depende do volume, em outras palavras, com  $\gamma = 0$ . O ajuste do modelo dado pela equação 1 foi muito mais plausível para a descrição da série temporal ( Critério de informação de Akaike,  $\Delta AIC > 7$ ).

Na Figura 1 mostramos, a título de exemplo, a curva resultante do ajuste do modelo para o período de maio/2014 e maio/2015.

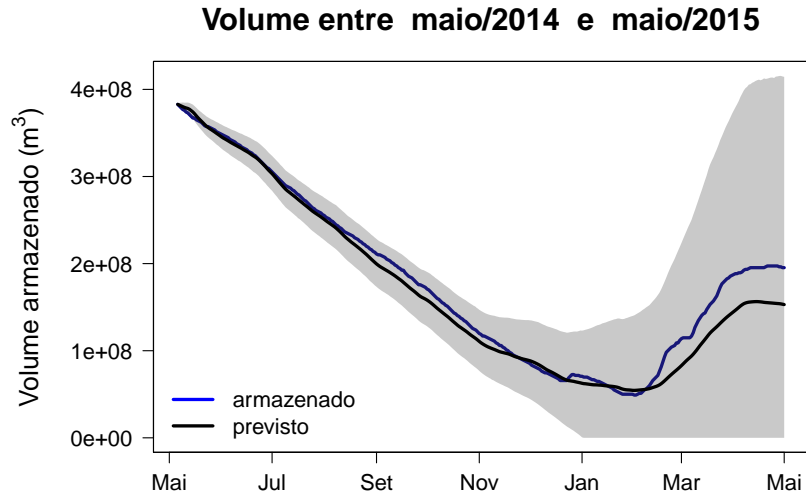


Figura 1: Séries temporais para o volume de água armazenado no sistema Cantareira entre maio/2014 e maio/2015. A linha azul mostra o volume de fato armazenado, dia-a-dia, neste período. A linha preta mostra o volume previsto pelo modelo. A área sombreada mostra a região de credibilidade bayesiana. Trajetória e intervalos são calculados a partir de 4000 simulações numéricas do modelo, com parâmetros amostrados da distribuição posterior do ajuste BSMC ( *Bayesian Sequential Monte Carlo*). Dados do *site* da SABESP: <http://www2.sabesp.com.br/mananciais/divulgacaopcj.aspx>. Coeficientes deste ajuste:  $\alpha = 5,930$ ,  $\beta = 1,043$ ,  $\gamma = 0,590$ ,  $\epsilon = 3,2 \times 10^7$ ,  $\sigma = 3,192 \times 10^3$ .

## 3 Calibração e Validação

Para fazer as projeções, fazemos a calibração do modelo pelo processo de ajuste de parâmetros explicado acima, para períodos de seis meses. De fato, os valores dos pa-

râmetros estimados podem variar de acordo com o intervalo de tempo considerado, mostrando que o modelo não tem uma validade global, devendo ser ajustado a intervalos dos dados. Embora não tenhamos verificado, especulamos que isto acontece por variações ao longo do tempo de variáveis que não incluímos no modelo, como por exemplo a temperatura.

Para validar as projeções, realizamos estes ajustes para vários intervalos de seis meses ao longo da série temporal que vai de janeiro de 2012 a fevereiro de 2015. Para cada um destes ajustes fizemos projeções para os meses seguintes, que comparamos com os volumes observados. Observamos que as projeções assim feitas foram sempre muito acuradas para o período de um mês. Projeções para períodos maiores muitas vezes resultam corretas, mas não sempre. Deste modo, por um princípio de precaução, estabelecemos o período de um mês como o limite de validade das projeções.

## 4 Projeções

A cada dia ajustamos o modelo aos dados, acrescidos no novo valor do dia. Para as projeções, necessitamos tecer algumas hipóteses sobre:

- pluviosidade futura;
- vazão defluente futura.

Com respeito à vazão defluente, usamos a média dos trinta dias passados, dado que é pouco comum a retirada de água do sistema sofrer variações bruscas. Tal hipótese pode ser modificada se os órgãos gestores vierem a anunciar políticas distintas.

Já com respeito à pluviosidade futura, as projeções disponíveis no *site Águas Futuras* usam dois tipos de cenários. Para projeções de curtíssimo prazo (5 dias) usamos a previsão meteorológica para a bacia PCJ (disponível no site do Comitê PCJ), que por sua vez utiliza previsões do INPE (Instituto Nacional de Pesquisas Espaciais). Para as projeções de um mês usamos três cenários baseados na média climatológica da pluviosidade (na média, 25% acima ou 25% abaixo da média). Dado que a média é da pluviosidade acumulada mensal, fazemos uma interpolação de modo a obter valores diários.

Vale lembrar, por fim, que o parâmetro  $R$ , na equação (1), é a pluviosidade média dos trinta dias anteriores, o que dá ao modelo uma certa “inércia”. Esse fato, que é necessário para produzir resultados acurados, deriva de mecanismos hidrológicos que não são modelados explicitamente, em linha com a visão fenomenológica que está na base do modelo.

## 5 Para saber mais

Todas as análises são realizadas no ambiente de código aberto R [5]. Os códigos dessas análises e os dados são públicos e estão disponíveis no *site Águas Futuras* (<http://cantareira.github.io>). Caso precise de mais detalhes sobre os ajustes e projeções, ou sobre como usar estes códigos entre em contato conosco. Críticas e sugestões são muito bem vindas também.

## Referências

- [1] K. Beven, *Rainfall-Runoff Modelling: the Primer* (Wiley, 2012). Ver sobretudo o capítulo 4 e em particular a discussão sobre *data-based mechanistic models*.

- [2] A. A. King, *et alli.*: pomp: Statistical inference for partially observed Markov processes (R package), 2014. URL <http://pomp.r-forge.r-project.org>
- [3] J. Liu e M. West. *Combining Parameter and State Estimation in Simulation-Based Filtering*, pp. 197–224, (Springer, New York, 2001).
- [4] K. P. Burnham e D. R. Anderson , *Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach*, (Springer, New York, 2010).
- [5] R Core Team. *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria, 2015. URL <http://www.R-project.org/>.