

## Proceeding Series of the Brazilian Society of Computational and Applied Mathematics

# Estimação de parâmetros em modelos estocásticos de vazões utilizando Algoritmos Genéticos

Leandro Furlam Turi<sup>1</sup>  
 Mônica de Souza Mendes Castro  
 Antonio Sérgio Ferreira Mendonça  
 Lucia Catabriga  
 Universidade Federal do Espírito Santo

## 1 Introdução e justificativa

Conhecer o comportamento de séries hidrológicas é de fundamental importância para diferentes finalidades no planejamento e gerenciamento de recursos hídricos. Geralmente, séries históricas de vazões são bastante curtas, o que torna necessária a utilização de séries sintéticas que contenham as mesmas propriedades estatísticas da série original, porém simulando sequências maiores de possíveis realizações de eventos de cheias e secas [3]. Mendonça, em [2], desenvolveu um modelo *multiplicative Periodic AutoRegressive Moving Average*, denominado PARMA multiplicativo ou PMIX, que permite a geração de vazões mensais e, por agregação, são geradas as vazões anuais. Tal modelo utiliza o processo de otimização não linear, denominado algoritmo de Powell, para determinação dos parâmetros periódicos. A forma matricial do PMIX está representada na Equação (1), onde  $p$ ,  $P$ ,  $q$ ,  $Q$  são as ordens do modelo;  $y$  é um valor de vazão medido e  $\epsilon$  é um resíduo aleatório, ambos de um ano  $v$  e período  $\tau$ . Demais variáveis são os parâmetros a serem estimados.

$$\begin{aligned} & [1 - \phi_{1,\tau}y_{v,\tau-1} - \dots - \phi_{p,\tau}y_{v,\tau-p}][1 - \Phi_{1,\tau} - \Phi_{P,\tau}y_{v-P,\tau}]y_{v,\tau} = \\ & [1 - \theta_{1,\tau}\epsilon_{v,\tau-1} - \dots - \theta_{q,\tau}\epsilon_{v,\tau-q}][1 - \Theta_{1,\tau}\epsilon_{v-1,\tau} - \Theta_{Q,\tau}\epsilon_{v-Q,\tau}]\epsilon_{v,\tau} \end{aligned} \quad (1)$$

Nessa pesquisa, busca-se a estimação ótima dos parâmetros no modelo estocástico, com aplicação de um Algoritmo Evolutivo multiobjetivo baseado no *fast Nondominated Sorting Genetic Algorithm* (NSGA-II), proposto por [1].

## 2 Metodologia

O algoritmo foi escrito na linguagem R<sup>2</sup>, com o auxílio de pacotes-padrão disponíveis no software. Na implementação do Algoritmo Genético foram utilizadas restrições na geração dos indivíduos a partir da análise de valores atípicos encontrados nas séries sintéticas geradas. Os parâmetros genéticos - tamanho da população, probabilidades de cruzamento, mutação e ciclo máximo - foram escolhidos empiricamente. A avaliação dos indivíduos foi

<sup>1</sup>leandrofturi@gmail.com

<sup>2</sup>Software livre, disponível em <https://www.R-project.org/>

feita pela comparação da Média Percentual Absoluta do Erro (MAPE), das séries históricas em relação às séries sintéticas. Foram definidas como funções objetivo as reduções dos valores de avaliação de médias, desvios-padrão, autocorrelações mensais e anuais, além da redução no somatório dos resíduos.

### 3 Discussões e Conclusão

Os parâmetros genéticos que apresentaram os melhores resultados foram: populações de 50 indivíduos (vetor de parâmetros), probabilidade de cruzamento de 50%, probabilidade de mutação de 5% e ciclo máximo de 10 mil. Em todos os testes realizados, foram encontrados resultados em que ocorreu a reprodução na série sintética da maioria das propriedades estatísticas da série histórica, sobretudo médias e desvios-padrão mensais, além de autocorrelações mensais e anuais em *lags* menores.

Tomando-se para análise a estação fluviométrica de Kinshasa (rio Congo, 1912 a 2010), em um indivíduo encontrado dentro do cenário de ótimos que se mostrou promissor a melhor solução, em até 10 mil ciclos do Algoritmo Genético, foram encontrados valores MAPE de médias de 0,213%, MAPE de desvios-padrão de 1,267%, MAPE de autocorrelações anuais *lag* 1 de 18,795% e mensais *lag* 1 de 0,886%, e somatório de resíduos de 283.369. Comparações entre parâmetros estatísticos são apresentadas na Figura 1.

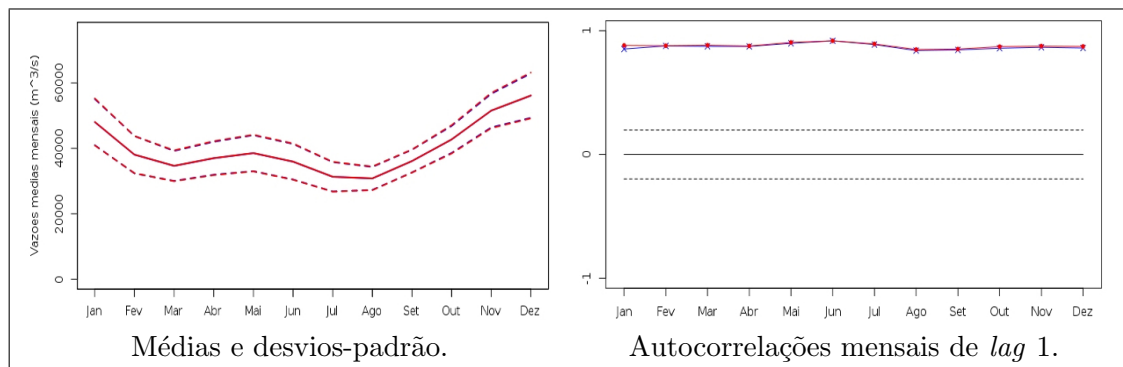


Figura 1: Comparação entre parâmetros históricos (em azul) e sintéticos (em vermelho) de um indivíduo dentro do cenário de ótimos. Ressalta-se o bom ajuste entre os dados.

### Referências

- [1] K. Deb, S. Agrawal, A. Pratap and T. Meyarivan. A fast elitist non-dominated sorting genetic algorithm for multi-objective optimization: NSGA-II. *International conference on parallel problem solving from nature*. Springer, Berlin, Heidelberg, 1:849-858, 2000. DOI: B10.1007/3-540-45356-3\_83.
- [2] A. S. F. Mendonça. Stochastic modeling of seasonal streamflow. Tese de Doutorado, Colorado State University, 1987.
- [3] J. Salas, J. W. Delleur and V. Yevjevich. *Applied modeling of hydrologic time series*. Book Crafters, Colorado, 1980.