

# Um Algoritmo Genético Epidêmico para Calibração Multiobjetivo do Modelo Hidrológico IPH2

Amarísio da S. Araújo<sup>1</sup>, Haroldo F. de Campos Velho<sup>2</sup>, Vitor C. F. Gomes<sup>1</sup>

<sup>1</sup>Programa de Pos-Graduação em Computação Aplicada – CAP

Instituto Nacional de Pesquisas Espaciais – INPE

<sup>2</sup>Laboratório Associado de Computação e Matemática Aplicada – LAC

Instituto Nacional de Pesquisas Espaciais – INPE

{amarisio.araujo,haroldo,vitor.gomes}@lac.inpe.br

**Abstract.** *Hydrological models that simulate the process of transformation of rainfall into runoff are based on simplified mathematical equations that describe the behavior of a watershed, with some parameters to be adjusted. This adjustment, called calibration, can be done with numerical techniques of optimization based on the concept of Pareto front. This paper presents an epidemic genetic algorithm for multi-objective calibration of hydrologic model IPH2.*

**Resumo.** *Modelos hidrológicos que simulam o processo de transformação de chuva em vazão baseiam-se em equações matemáticas simplificadas que descrevem o comportamento de uma bacia hidrográfica, com alguns parâmetros que devem ser ajustados. Tal ajuste, chamado de calibração, pode ser feito com técnicas numéricas de otimização multiobjetivo baseadas no conceito de frente de Pareto. Este trabalho apresenta um algoritmo genético epidêmico para a calibração multiobjetivo do modelo hidrológico IPH2.*

**Palavras-chave:** *modelo hidrológico, calibração multiobjetivo, algoritmo genético.*

## 1. Introdução

Os diversos processos do ciclo hidrológico que atuam em uma bacia hidrográfica podem ser representados por modelos matemáticos simplificados que simulam a transformação da chuva em vazão - modelos chuva-vazão. Tais modelos possuem parâmetros que caracterizam o sistema, sendo que alguns deles, por representarem abstrações da realidade, não podem ser medidos [Tucci, 2005]. A vazão estimada  $\hat{Q}_t$  depende, em cada intervalo de tempo  $t$ , do valor da precipitação  $P_t$ , do valor da evaporação  $E_t$  e dos valores de tais parâmetros,  $\theta$ , ou seja:  $\hat{Q}_t(\theta) = H(P_t, E_t, \theta)$ , onde  $H(\cdot)$  representa o modelo hidrológico. A diferença entre os valores simulados e os  $NT$

dados observados é definida por:  $\varepsilon_t(t) = Q_t - \hat{Q}_t$ ,  $t = 1, 2, \dots, NT$  (estas diferenças representam os erros na estimação do modelo em cada intervalo de tempo). Com as informações hidrológicas da bacia, passa-se à calibração (ajuste) dos parâmetros  $\theta$ , que pode ser feita de formas diversas através de funções objetivo  $F(\varepsilon(\theta))$ , que medem o desempenho do modelo. Devido às simplificações do modelo e da representatividade dos parâmetros, e às incertezas nos dados, na prática, não existirá um conjunto único de parâmetros que seja capaz de representar os processos hidrológicos. Assim, a calibração é um problema matemático de muitas soluções, podendo ocorrer soluções que não representem a realidade do problema e nem mesmo dos valores esperados. Das soluções aceitáveis tecnicamente, pode não ser possível distinguir a mais aceitável ou a melhor.

A calibração multiobjetivo de um modelo hidrológico é um problema de otimização multiobjetivo com a formulação geral:  $\min_{\theta \in \Omega} F(\theta) = \{F_1(\theta), \dots, F_k(\theta)\}$ , onde  $F_1(\theta), \dots, F_k(\theta)$  são as  $k$  diferentes funções objetivo a serem simultaneamente minimizadas e  $\theta$  é o vetor de parâmetros em um espaço factível de parâmetros  $\Omega \subset \mathbb{R}^n$ . A função  $F$  mapeia o espaço factível de parâmetros  $\Omega \subset \mathbb{R}^n$  no espaço objetivo  $Z \subset \mathbb{R}^k$ . Com a formulação seguinte [Yapo et al.], tem-se a divisão de  $\Omega$  em dois conjuntos: o das soluções “boas”, não dominadas (*conjunto de Pareto*) e o das soluções “ruins”, dominadas. Por definição, toda solução  $x_P$  pertencente ao conjunto de Pareto deve ser tal que:  $F(x_P)$  é estritamente menor do que  $F(x_D)$ , para toda solução  $x_D$  não contida no conjunto de Pareto, e não existe uma solução  $x_P^*$  no conjunto de Pareto tal que  $F(x_P^*)$  seja estritamente menor do que  $F(x_P)$ . O conjunto de Pareto (alvo da otimização) é mapeado por  $F$  em um conjunto no espaço objetivo chamado de *frente de Pareto*.

Neste artigo, é apresentado um algoritmo genético epidêmico (AGE) para a calibração multiobjetivo do modelo hidrológico de pequenas bacias IPH2 [Tucci, 1998]. Tal algoritmo é uma variante do algoritmo evolucionário MOCOM [Yapo et al, 1998], o qual já vem sendo usado no IPH2. Além dos operadores genéticos de cruzamento e mutação, ele utiliza um operador epidêmico (Chiwiacowsky, Campos Velho, 2003).

## 2. O Algoritmo AGE

Os cinco primeiros passos do AGE são idênticos ao do MOCOM. No *primeiro passo*, é estabelecida a região factível para os  $k$  parâmetros. No *segundo passo*, usando uma distribuição uniforme, é gerada uma população na região factível de parâmetros com  $ns$  pontos (indivíduos) formados por  $k$  coordenadas. No *terceiro passo*, os pontos da população são avaliados com as  $nf$  funções objetivo. No *quarto passo*, é feito o seguinte *ordenamento de Pareto*: identificam-se todos os indivíduos na população que são não dominados, e a esses indivíduos atribui-se o ranking 1. Estes indivíduos com ranking igual a 1 são temporariamente retirados da população. Na população restante, são identificados os indivíduos não dominados, e a esses indivíduos atribui-se o ranking 2. Retiram-se temporariamente esses indivíduos com ranking 2 dessa população. Estes passos se repetem até que a cada indivíduo da população tenha sido atribuído algum ranking. Os piores indivíduos (mais distantes da frente de Pareto) possuem o maior valor de ranking, o qual será denotado por  $R_{max}$ . Desta forma, são atribuídos, para cada ponto  $i$  da população, valores de ranking  $r_i$ , que variam de 1 a  $R_{max}$ , sendo  $R_{max} \leq ns$ .

*Quinto passo*: Cada indivíduo com ranking  $R_{max}$  dá origem a um conjunto denominado *Complexo* que é formado por um indivíduo com ranking  $R_{max}$  e mais  $k$

outros indivíduos selecionados aleatoriamente entre os indivíduos da população que não possuem ranking  $R_{max}$ , usando a seguinte equação de probabilidade associada a cada ponto  $i$ :  $p_i = \frac{R_{max}^{-r_i+1}}{\sum_{i=1}^{ns} (R_{max}^{-r_i+1})}$ . A probabilidade de um indivíduo ser selecionado depende de seu ranking, sendo favorecidos os indivíduos de menor ranking (melhores). Cada complexo possui, portanto,  $k + 1$  elementos.

**Sexto passo:** Aqui é feita a evolução dos complexos com o objetivo de fazer com que os pontos da população convirjam para a frente de Pareto. Dado um complexo, aplicam-se operadores genéticos de mutação e cruzamento em seus indivíduos, gerando-se  $k + 1$  novos elementos (filhos) no complexo até que, considerado um ordenamento de Pareto (descrito no quarto passo) feito no complexo, o indivíduo de maior ranking tenha um ranking menor do que o ranking máximo no complexo antes das operações. Feito isto, estes  $k + 1$  novos elementos substituem os elementos iniciais do complexo e diz-se que o complexo evoluiu. Colocou-se um parâmetro que estabelece o número máximo de tentativas de gerações em cada complexo até que se alcance uma evolução. Aqui, é considerada a aplicação de um operador de epidemia. Para tal, introduz-se um parâmetro no intervalo (0,1) (um percentual) de modo que se o percentual de complexos sem evolução alcança o valor deste parâmetro, ativa-se uma epidemia. Esta epidemia é aplicada a todos os indivíduos da população cujos valores de ranking sejam maiores ou iguais a um determinado valor pré-fixado (que também é um parâmetro), isto é, todos os indivíduos com valores de ranking maiores ou iguais a esse valor pré-fixado são eliminados da população e são gerados, do mesmo modo que a população inicial, outros tantos indivíduos que irão substituí-los, voltando-se ao terceiro passo do algoritmo. A aplicação de epidemias é limitada a um número máximo de iterações.

**Sétimo passo:** Feita a evolução de todos os complexos, os indivíduos são devolvidos à população.

**Oitavo passo:** O critério de parada: o processo finaliza quando  $R_{max} = 1$  ou quando um número máximo de iterações é atingido.

### 3. Alguns Resultados de Calibração do IPH2 com o AGE e com o MOCOM

Foram efetuados testes de calibração do modelo IPH2 com o AGE e com o MOCOM, usando duas funções objetivo:  $F_1 = \sqrt{\frac{\sum_{t=1}^{NT} (Q_t - \hat{Q}_t)^2}{NT}}$  (desvio das vazões) e

$F_2 = \sqrt{\frac{\sum_{t=1}^{NT} \left(\frac{1}{Q_t} - \frac{1}{\hat{Q}_t}\right)^2}{NT}}$  (desvio dos inversos das vazões). O modelo direto foi executado

com valores pré-fixados (Bravo et al., 2009) dos sete parâmetros do modelo e com dados reais de precipitação e evaporação (301 registros diários no período de 11/03/1983 a 05/01/1984) de uma bacia do rio Canoas, em Santa Catarina, cuja área é de 989 km<sup>2</sup>. Para simular as vazões observadas foi adicionado um ruído de 5% nos valores de saída do modelo IPH2, gerando uma série sintética de vazões observadas que foi usada para o procedimento de calibração.

A Figura 1 mostra as aproximações da frente de Pareto obtidas em uma calibração do IPH2 com as duas estratégias, considerando uma população de tamanho  $ns = 50$ . É importante observar que os melhores resultados de calibração se dão à medida que a frente de Pareto se aproxima da origem.

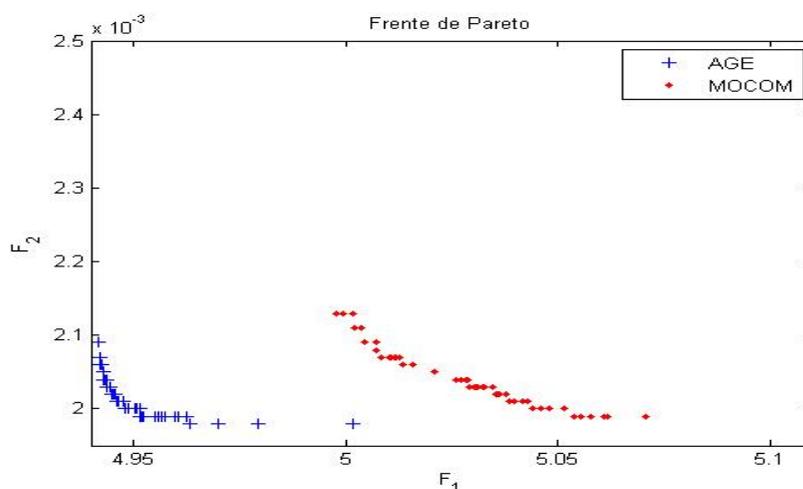


Figura 1. Resultados da calibração automática do modelo IPH2.

#### 4. Conclusões

Com o desenvolvimento dos computadores, cada vez mais rápidos e mais potentes, a calibração automática multiobjectivo dos parâmetros de modelos hidrológicos tem ganhado cada vez mais espaço, uma vez que ela permite uma melhor avaliação das incertezas inerentes aos processos, das imperfeições dos modelos e da representatividade dos parâmetros.

Este trabalho apresentou um algoritmo genético epidêmico (AGE) como uma possível alternativa para a calibração do modelo hidrológico de pequenas bacias IPH2, modelo este que tem se mostrado eficiente para simular várias bacias brasileiras, usando o algoritmo evolucionário MOCOM na sua calibração. Foram apresentados resultados de calibração usando as duas estratégias (AGE e MOCOM), a partir de uma série sintética de vazões, nos quais se pôde verificar a eficiência do AGE em gerar uma boa aproximação da frente de Pareto.

#### Referências

- Bravo, J. M., Collischon, W., Tucci, C. E. M. Verificação de eficiência de um algoritmo evolucionário multiobjectivo na calibração automática do modelo hidrológico IPH II. Revista Brasileira de Recursos Hídricos-RBRH, v.14, 37-60p.1999
- Chiwiacowsky, L., Campos Velho, H. F. Different approaches for the solution of a backward heat conduction problem, Inverse Problems in Engineering, v.11, n.6, 471-494p., 2003.
- Tucci, C. E. M. Modelos hidrológicos. 2 ed., Porto Alegre (RS): Editora da UFRGS, 2005. 669p. ISBN 85-7025-823-2.
- Yapo, P. O., Gupta, H. V., Sorooshian, S. Multi-objective global optimization for hydrological models. Journal of Hydrology, v.204, 83-97p.1998.