

# Um Algoritmo Genético Associado a uma Estratégia de Decisão Baseada em Lógica Fuzzy para Calibração Multi-objetivo de um Modelo Hidrológico

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 and a strategy based on fuzzy logic to choose the best solution among the Pareto solutions obtained by the algorithm.*

**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. Neste artigo, é apresentado um algoritmo genético epidêmico para a calibração multiobjetivo do modelo hidrológico IPH2 e uma estratégia baseada em lógica fuzzy para a escolha da melhor solução dentre as soluções de Pareto obtidas pelo algoritmo.*

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

## 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, 1998]. 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 de um vetor de parâmetros,  $\theta$ , ou seja:  $\hat{Q}_t(\theta) = H(P_t, E_t, \theta)$ , onde  $H(\cdot)$  representa o modelo hidrológico. Os erros na estimação do modelo em cada intervalo de tempo são obtidos pelas diferenças entre os valores simulados e os  $NT$  dados observados:  $\varepsilon_t(t) = Q_t - \hat{Q}_t$ ,  $t = 1, 2, \dots, NT$ . Com as informações da bacia hidrográfica, passa-se à calibração (ajuste) dos parâmetros  $\theta$ , que pode ser feita de formas diversas através de funções objetivo  $F(\varepsilon(\theta))$ , que quando otimizadas, espera-se, estabelecerão os valores dos parâmetros que gerem valores simulados de vazão próximos aos valores observados. As simplificações do modelo e da representatividade dos parâmetros, bem como as incertezas nos dados, são indicadores de que, na prática, não existirá um conjunto único de parâmetros que seja capaz de representar os processos hidrológicos, sendo necessário considerar duas ou mais funções objetivo, em geral, conflitantes. Assim, a calibração é um problema matemático de muitas (talvez infinitas) soluções igualmente possíveis, podendo ocorrer soluções que não representem a realidade do problema e nem mesmo dos valores esperados. Dentre as soluções aceitáveis tecnicamente, é importante que se saiba distinguir a melhor.

A calibração multiobjetivo de um modelo hidrológico é um problema de otimização multiobjetivo com a formulação geral:  $\min_{\theta \in \Omega} F(\theta) = \{F_1(\theta), \dots, F_m(\theta)\}$ , onde  $F_1(\theta), \dots, F_m(\theta)$  são as  $m$  diferentes funções objetivo a serem simultaneamente minimizadas e  $\theta = (\theta_1, \dots, \theta_k)$  é o vetor de parâmetros em um espaço factível de parâmetros  $\Omega \subset \mathbb{R}^k$ . A função  $F$  mapeia o espaço factível de parâmetros  $\Omega \subset \mathbb{R}^k$  no espaço objetivo  $Z \subset \mathbb{R}^m$ . 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:  $F_j(x_p)$  é estritamente menor do que  $F_j(x_D)$  para qualquer  $j$ , para toda solução  $x_D$  não contida no conjunto de Pareto, e não existe uma solução  $x_p^*$  no conjunto de Pareto tal que  $F_j(x_p^*)$  seja estritamente menor do que  $F_j(x_p)$  para todo  $j$ . O conjunto de Pareto (alvo da otimização) é mapeado por  $F$  em um conjunto no espaço objetivo chamado de *frente de Pareto*. Encontrado o conjunto de Pareto, métodos de tomada de decisão podem ser utilizados para determinar a melhor solução dentre as soluções de Pareto.

Neste artigo, é apresentado um algoritmo genético epidêmico (AGE) para a calibração multiobjetivo do modelo hidrológico de pequenas bacias IPH2 [Tucci, 1998] e uma estratégia, baseada em lógica fuzzy, para a escolha da melhor solução dentre as soluções de Pareto obtidas pelo AGE.

## 2. O Algoritmo AGE

O AGE 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). Com o AGE, obtém-se uma aproximação da frente de Pareto na calibração multiobjetivo do IPH2. Os cinco primeiros passos do AGE são idênticos ao do MOCOM.

**Primeiro passo:** Inicialmente, são estabelecidos os limites mínimos e máximos que os  $k$  parâmetros a serem ajustados podem assumir, definindo, assim, uma região factível  $\Omega$  para os parâmetros.

**Segundo passo:** Usando uma distribuição uniforme, é gerada uma população na região factível de parâmetros com  $ns$  pontos (indivíduos) formados por  $k$  coordenadas.

**Terceiro passo:** Os pontos da população são avaliados com as  $m$  funções objetivo.

**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 converjam 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 o ordenamento de Pareto feito no mesmo, o indivíduo de maior ranking tenha um ranking menor do que o ranking máximo no complexo antes das operações. Quando isto ocorre, este novo indivíduo de maior ranking substitui o indivíduo de ranking máximo do complexo, e diz-se, então, 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 um número máximo de iterações é atingido ou quando  $R_{max} = 1$ , ou seja, as soluções são todas não dominadas (frente de Pareto).

### 3. Uma estratégia de tomada de decisão baseada em lógica fuzzy

Depois de estabelecidas as soluções de Pareto (frente de Pareto) na calibração multiobjetivo do modelo hidrológico, é necessário uma tomada de decisão no sentido de escolher, dentre as soluções de Pareto, a solução de compromisso (ou de preferência). Neste trabalho será considerada para tal tarefa uma estratégia baseada na teoria de lógica fuzzy.

Em um ambiente de lógica fuzzy, em problemas que envolvem incertezas e imprecisão, a decisão é frequentemente definida baseada nas relações seguintes:

$$\text{encontre } x \text{ que satisfaça } f_i(x) \text{ op } f_i^*, i = 1, \dots, m, \text{ sujeito a } x \in G$$

onde  $f_i^*$  é a solução ideal da função objetivo  $f_i$  e  $G \subset \mathbb{R}^n$  é o conjunto que envolve as restrições do problema e  $op \in \{\lesssim, \cong, \gtrsim\}$ .

Os três tipos de relação fuzzy  $\lesssim$ ,  $\gtrsim$  e  $\cong$  são usados para indicar, respectivamente, que  $i$ -ésimo objetivo fuzzy é aproximadamente menor ou igual, maior ou igual, e na vizinhança de  $f_i^*$ .

Pela natureza imprecisa do julgamento do operador de decisão, usa-se uma função de pertinência  $\mu_i$  para representar a  $i$ -ésima função objetivo de uma solução no conjunto de Pareto. Os valores da função de pertinência indicam o grau de satisfação das funções objetivo do problema, estando estes valores entre 0 e 1. Há vários tipos de funções de pertinência como, por exemplo, a linear, a triangular, a trapezoidal e a exponencial [Sakawa et al., 1987]. Como o problema multiobjetivo aqui considerado é de minimização, usaremos a função de pertinência correspondente à relação fuzzy do tipo  $\lesssim$ , que é dada por:

$$\mu_i = \begin{cases} 1; & f_i \leq f_i^{min} \\ 0; & f_i \geq f_i^{max} \\ \frac{f_i^{max} - f_i}{f_i^{max} - f_i^{min}}; & f_i^{min} < f_i < f_i^{max} \end{cases}$$

onde  $f_i^{min}$  e  $f_i^{max}$  são, respectivamente, os valores mínimo e máximo da  $i$ -ésima função objetivo. Estes valores podem ser estabelecidos após a ordenação das soluções da frente de Pareto, sendo  $f_i^{max}$  a pior solução para a  $i$ -ésima função objetivo e  $f_i^{min}$  a melhor solução para a  $i$ -ésima função objetivo.

Para cada solução não dominada  $p$ , considerando  $N_D$  como sendo o número de soluções não dominadas da frente de Pareto e  $N_{obj}$  o número de funções objetivo, obtém-se uma função de pertinência com a seguinte equação:

$$\mu^{(p)} = \sum_{i=1}^{N_{obj}} \mu_i^{(p)}$$

A melhor solução de compromisso é encontrada por meio do valor máximo de  $\mu^{(p)}$ , ou seja, será aquela correspondente a um dos valores  $p = 1, 2, \dots, N_D$ , para a qual  $\mu^{(p)}$  é máximo [Lopes, 2010].

#### 4. Resultados de Calibração do IPH2 com o AGE e com uma tomada de decisão baseada em lógica fuzzy

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

$F_2 = \sqrt{\frac{\sum_{t=1}^{NT} \left(\frac{1}{\hat{Q}_t} - \frac{1}{Q_t}\right)^2}{NT}}$  (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.

Na Figura 1, temos uma aproximação da frente de Pareto obtida em uma das calibrações do IPH2 com o AGE, sendo indicada, no espaço objetivo, a solução de compromisso obtida como resultado da tomada de decisão baseada em lógica fuzzy aqui apresentada. Os valores dos parâmetros correspondentes a essa solução são apresentados na Tabela 1, na qual também é apresentada a faixa de variação de cada um dos parâmetros, consideradas todas as soluções do conjunto de Pareto.

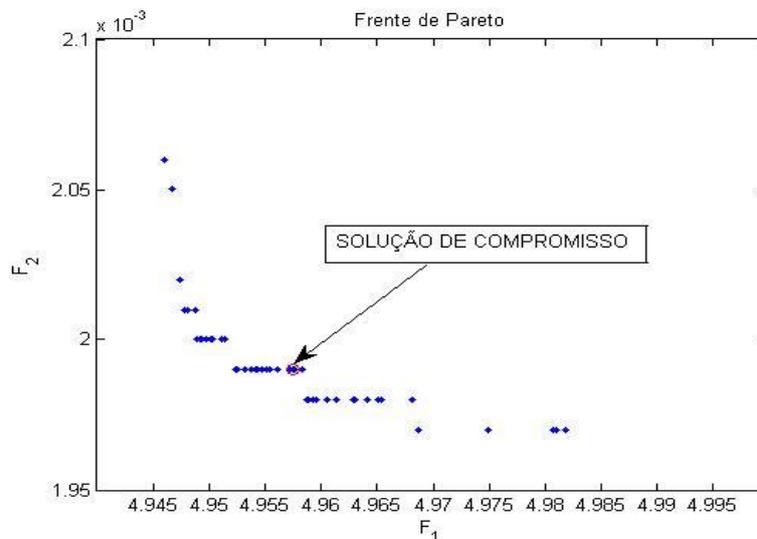


Figura 1. Resultado da calibração automática do modelo IPH2.

Tabela 1: Solução de compromisso obtida com a estratégia apresentada e faixa de variação dos parâmetros, resultantes de uma calibração como o AGE.

Parâmetros	Unidades	Valores Exatos	Solução de Compromisso	Faixa de variação dos parâmetros
$I_0$	$mm\Delta t^{-1}$	18,20	22,91	22,71 - 24,04
$I_b$	$mm\Delta t^{-1}$	2,95	3,03	2,90 - 3,04
$H$	--	0,13	0,06	0,055 - 0,070
$K_s$	$\Delta t$	5,72	5,64	5,61 - 5,70
$K_{sub}$	$\Delta t$	40,25	37,55	37,17 - 41,75
$RMAX$	$mm$	0,24	0,56	0,48 - 0,86
$\alpha$	--	12,66	14,62	12,78 - 18,01

## 5. Conclusões

Com o desenvolvimento dos computadores, cada vez mais rápidos e mais potentes, a calibração automática multiobjetivo 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) para a calibração multiobjetivo do modelo hidrológico de pequenas bacias IPH2 e um mecanismo de tomada de decisão baseado em lógica fuzzy para a escolha da melhor solução dentre as soluções de Pareto obtidas na calibração. Foram apresentados resultados de calibração do IPH2 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 e também a eficiência do mecanismo de tomada de decisão na obtenção de um bom ajuste dos parâmetros do modelo hidrológico IPH2.

Em trabalhos futuros, a estratégia aqui apresentada será comparada a outras técnicas já existentes usadas para encontrar a solução única do problema de calibração multiobjetivo do modelo hidrológico considerado.

## Referências

- Bravo, J. M., Collischon, W., Tucci, C. E. M. Verificação de eficiência de um algoritmo evolucionário multiobjetivo 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.

Lopes, P. S., Modelagem de problema inverso de detecção de danos por técnicas de identificação de parâmetros e de otimização. 2010. 136f. Tese (Doutorado) – Programa de Pós-Graduação em Engenharia Mecânica – Instituto de Engenharia Mecânica – Universidade Federal de Itajubá – Itajubá (MG), 2010.

Sakawa, M., Yano, H., Yumine, T., 1987, An interactive fuzzy satisficing method for multi-objective linear-programing problems and its application, IEEE Transactions on SMC, v.17, n.4, pp.654-661.

Tucci, C. E. M., 1998, “Modelos hidrológicos”, 669 pp., ABRH Editora da UFRGS, Porto Alegre, RS.

Yapo, P. O., Gupta, H. V., Sorooshian, S. Multi-objective global optimization for hydrological models. Journal of Hydrology, v.204, 83-97p.1998.