

UM ALGORITMO GENÉTICO PARA CALIBRAÇÃO MULTI OBJETIVO DE UM MODELO HIDROLÓGICO

Amarísio da Silva Araújo

Universidade Federal de Viçosa - UFV
Departamento de Matemática - DMA
amarisio@ufv.br

Haroldo Fraga de Campos Velho

Instituto Nacional de Pesquisas Espaciais - INPE
Laboratório Associado de Computação e Matemática Aplicada - LAC
haroldo@lac.inpe.br

Vitor Conrado Faria Gomes

Instituto Nacional de Pesquisas Espaciais - INPE
Laboratório Associado de Computação e Matemática Aplicada - LAC
vitor.gomes@lac.inpe.br

Resumo

Modelos hidrológicos que simulam o processo de transformação de chuva em vazão são os mais utilizados na representação dos complexos processos do ciclo hidrológico. 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 automaticamente 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-Chaves: Modelo hidrológico; Calibração multiobjetivo; Algoritmo genético; Frente de Pareto.

Abstract

Hydrological models that simulate the process of transformation of rainfall into flow are the most used in the representation of the complex processes of the hydrological cycle. They are based on simplified mathematical equations that describe the behavior of a watershed, with some parameters which must be adjusted. This adjustment, called calibration can be done automatically with numerical techniques of multi-objective optimization based on the concept of Pareto front. This paper presents an epidemic genetic algorithm for multi-objective calibration of the hydrological model IPH2.

Keywords: Hydrological model; Multi-objective calibration; Genetic algorithm; Pareto front.

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. Com tais modelos, torna-se possível uma melhor compreensão do comportamento dos fenômenos hidrológicos na bacia, tais como precipitação, escoamento e armazenamento superficiais, infiltração e escoamento da água no solo, evapotranspiração, conformações geomorfológicas, transporte de sedimentos etc.. Eles são extremamente úteis na previsão de cenários hidrológicos, no preenchimento de falhas nos dados observados, na análise dos efeitos resultantes de modificação do uso do solo na bacia e na previsão de variáveis hidrológicas (vazão, por exemplo) em tempo real ([14]).

Os modelos hidrológicos chuva-vazão possuem parâmetros que caracterizam o sistema ([14]), sendo que alguns deles, por representarem abstrações da realidade, não podem ser diretamente medidos. Desse modo, levando-se em conta as informações hidrológicas existentes na bacia, os valores de tais parâmetros devem ser ajustados, processo este conhecido como calibração. 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. 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 mais aceitável ou a melhor.

A calibração manual por tentativa e erro é uma das técnicas pioneiras de calibração. Consiste em um procedimento em que o usuário altera os valores dos parâmetros do modelo a cada tentativa e compara os resultados obtidos com os valores medidos até se chegar a um conjunto de parâmetros para os quais, no seu entendimento, os resultados do modelo são os mais apropriados para o processo simulado. A comparação a cada tentativa é feita geralmente com alguma métrica ou estatística que avalia a medida do erro entre os valores de vazão calculados pelo modelo e os valores de vazão observados. Apesar de sua robustez e simplicidade, a calibração manual depende essencialmente da habilidade do usuário. Além de entender o modelo, é preciso possuir uma habilidade para se encontrar uma solução por tentativa, mesmo que, implicitamente, sejam consideradas múltiplas métricas.

O processo de calibração torna-se mais rápido e mais eficiente com o uso de técnicas de calibração automática, baseadas no uso de computação intensiva. Nestas técnicas, o procedimento básico para o ajuste dos parâmetros nos modelos hidrológicos considera os elementos seguintes: uma ou mais medidas de avaliação dos erros (as chamadas medidas de desempenho ou funções objetivo), um algoritmo de otimização e um critério de parada. Assim, a calibração se torna um problema de otimização de uma única função (calibração mono objetivo) ou de mais de uma função (calibração multiobjetivo).

Na calibração mono objetivo, considera-se uma única medida de desempenho, entre as várias possíveis, obtendo-se como resultado um conjunto único de parâmetros que otimizam tal medida. Na calibração multiobjetivo, considera-se a otimização simultânea de mais de uma medida de desempenho, que resultará em vários conjuntos de parâmetros que irão constituir um conjunto de soluções ótimas (conjunto de Pareto). Deste conjunto de soluções ótimas, espera-se, de forma apropriada, chegar-se a uma resposta desejada para o ajuste do modelo.

Dos diferentes algoritmos de otimização multiobjetivo que vêm sendo desenvolvidos, os algoritmos evolucionários têm recebido uma maior atenção nas últimas

quatro décadas. Em particular, os algoritmos genéticos têm-se mostrado bastante eficazes nesta tarefa.

Neste artigo, é apresentado um algoritmo genético multiobjetivo com epidemia para o ajuste de parâmetros no modelo hidrológico IPH2 ([14]). Tal algoritmo é uma variante de um algoritmo evolucionário MOCOM – UA ([15]), que já vem sendo usado na calibração do referido modelo.

1. CALIBRAÇÃO MULTIOBJETIVO DE MODELOS HIDROLÓGICOS

1.1. FORMULAÇÃO DO PROBLEMA

Em um modelo hidrológico chuva-vazão, 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 dos parâmetros a serem ajustados, θ , ou seja:

$$\hat{Q}_t = \mathcal{H}(P_t, E_t, \theta), \quad (1)$$

onde $\mathcal{H}(\cdot)$ representa o modelo hidrológico.

Os erros na estimação do modelo em cada intervalo de tempo são obtidos, então, pela diferença entre os valores simulados de vazão e os nt valores observados de vazão, Q_t :

$$\varepsilon_t = Q_t - \hat{Q}_t, t = 1, \dots, nt. \quad (2)$$

A medida de desempenho do modelo hidrológico pode ser feita, então, de formas diversas através de funções $F(\varepsilon(\theta))$ que avaliam o erro $\varepsilon(\theta)$ e permitem, conseqüentemente, estabelecer os valores dos parâmetros que gerem valores simulados de vazão próximos aos valores observados. Tais funções são chamadas de funções objetivo. Existe um grande número de funções objetivo que podem ser utilizadas na calibração de modelos hidrológicos chuva-vazão. Algumas dão maior peso aos erros de vazões máximas (cheias) e outras avaliam melhor o ajuste nas vazões mínimas (estiagens). Nas equações seguintes, temos quatro exemplos de funções objetivo usadas comumente no ajuste de parâmetros de modelos hidrológicos ([1]), quais sejam:

$$F_1 = \sqrt{\frac{\sum_{t=1}^{nt} (Q_t - \hat{Q}_t)^2}{NT}} \quad (\text{Desvio padrão das vazões}) \quad (3)$$

$$F_2 = \sqrt{\frac{\sum_{t=1}^{nt} (\frac{1}{Q_t} - \frac{1}{\hat{Q}_t})^2}{NT}} \quad (\text{Desvio padrão do inverso das vazões}) \quad (4)$$

$$F_3 = \frac{\sum_{t=1}^{nt} |Q_t - \hat{Q}_t|}{NT} \quad (\text{Desvio absoluto médio das vazões}) \quad (5)$$

$$F_4 = \sum_{t=1}^{nt} \frac{(Q_t - \hat{Q}_t)^2}{Q_t} \quad (\text{Desvio quadrático relativo médio das vazões}) \quad (6)$$

A calibração multiobjetivo de um modelo hidrológico pode ser descrita como um problema de otimização multiobjetivo com a seguinte formulação:

$$\min_{\theta \in \Omega} F(\theta) = \{F_1(\theta), \dots, F_m(\theta)\} \quad (7)$$

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 de avaliação F mapeia o espaço factível de parâmetros $\Omega \subset \mathbb{R}^k$ (variáveis de decisão) no espaço objetivo $Z = (F_1, \dots, F_m) \subset \mathbb{R}^m$.

1.2. CONJUNTO DE PARETO – FRENTE DE PARETO

A principal característica de um problema de otimização multiobjetivo é que a solução, em geral, não é única. É comum existirem várias soluções tais que, ao se passar de uma solução para outra, a melhora no valor de uma função objetivo é acompanhada pela piora nos valores de pelo menos uma das outras funções. Torna-se necessário, pois, saber identificar dentre tais soluções aquelas que são mais adequadas para resolver o problema modelado, considerando a comparação dos valores das funções objetivo nessas soluções. Para isto, tem-se a seguinte definição que estabelece o importante conceito de dominância: diz-se que uma solução \bar{x} é não dominada por uma solução \tilde{x} quando \bar{x} não é pior do que \tilde{x} em todos os objetivos e, além disso, \bar{x} é melhor do que \tilde{x} em pelo menos um dos objetivos. As soluções não dominadas são chamadas de soluções Pareto e formam um conjunto no espaço factível de parâmetros chamado de conjunto de Pareto. Uma importante característica deste conjunto é que ao mover-se de uma solução para outra, a melhora obtida em uma função objetivo é seguida pela piora em pelo menos uma das outras funções objetivo. Considerando o exemplo simples no problema de minimizar as funções $g(x) = x^2$ e $h(x) = (x - 2)^2$ em relação a um parâmetro real x ([12]), o conjunto de Pareto pode ser facilmente identificado, e consiste de todos os pontos no intervalo $[0,2]$ no espaço dos parâmetros, dado que para $0 \leq x < 1$, tem-se $h(x) > g(x)$, e para $1 < x \leq 2$, tem-se $g(x) > h(x)$.

A formulação seguinte ([15]) estabelece a divisão do espaço factível de parâmetros em duas partes: o das soluções não dominadas (conjunto de Pareto) e o das soluções dominadas, considerando a minimização de $F = \{F_1(x), \dots, F_m(x)\}$. Por definição, toda solução x_p pertencente ao conjunto de Pareto deve satisfazer as seguintes propriedades:

- i) $F_j(x_p) < F_j(x_D), j = 1, \dots, n$, para toda solução x_D não contida no conjunto de Pareto;
- ii) Não existe uma solução x_p^* no conjunto de Pareto tal que $F_j(x_p^*) < F_j(x_p), j = 1, \dots, n$.

O espaço factível de soluções fica, então, dividido em dois conjuntos, um de “boas” soluções (soluções Pareto) e um de soluções “ruins”, sendo impossível distinguir a melhor dentre as soluções “boas”. O conjunto de Pareto (das soluções não dominadas) é, portanto, o alvo da calibração multiobjetivo. Este conjunto, quando mapeado no espaço objetivo, forma uma superfície conhecida como frente de Pareto. Na Figura 1, é ilustrada a frente de Pareto para as funções $g(x) = x^2$ e $h(x) = (x - 2)^2$ do exemplo anterior.

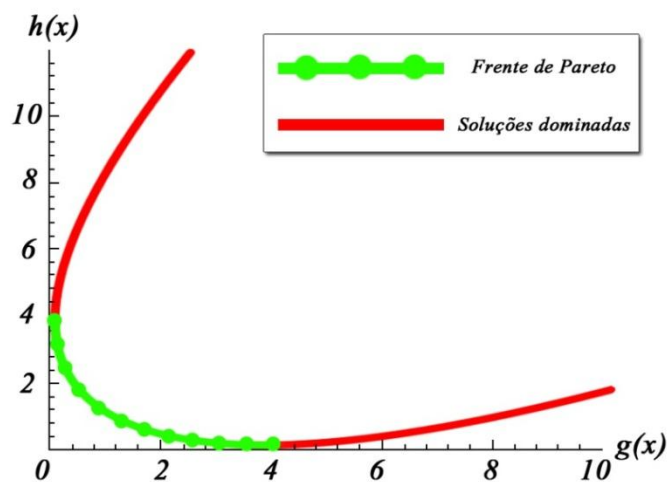


Figura 1: Espaço objetivo $Z=(g,h)$

Observa-se que, movendo-se ao longo da frente de Pareto, a melhora em uma função objetivo é acompanhada por uma piora na outra função. Neste caso, por se tratar de um problema de apenas duas funções objetivo, foi fácil identificar a frente de Pareto. Para problemas mais complexos, no entanto, não é possível fazer uma análise gráfica, sendo necessário o uso de técnicas computacionais para a obtenção de uma aproximação da frente de Pareto.

Várias técnicas têm sido propostas para gerar boas aproximações da frente de Pareto. Algumas delas, atribuindo diferentes pesos para as funções objetivo e reduzindo o problema a uma otimização mono objetivo ([7], [16]). Alguns métodos, usualmente iterativos, procuram gerar apenas um subconjunto de Pareto de interesse prévio. Outros requerem a escolha de uma solução dentre várias alternativas.

Técnicas mais eficientes foram apresentadas, por exemplo, por Erickson ([6]), Yapó ([15]), utilizando o conceito de ordenamento de Pareto (ou ranking de Pareto), que permitem encontrar, em apenas um procedimento de otimização, vários pontos na frente de Pareto. O algoritmo de Yapó, denominado MOCOM-UA (Multi-objective Complex Evolution-University of Arizona) combina técnicas de algoritmos evolucionários com o algoritmo simplex de Nelder e Mead ([11]). Deve-se mencionar aqui o algoritmo NSGA-II, desenvolvido por Deb e outros ([5]), que é considerado como o principal algoritmo genético multiobjetivo da atualidade.

1.3. ALGORITMOS GENÉTICOS

Um Algoritmo Genético (AG) ([8]) é uma metaheurística inspirada na Teoria da Evolução de Darwin que simula o processo de seleção natural e a sobrevivência dos indivíduos mais aptos de uma população.

Em um problema de otimização, por exemplo, um AG tenta encontrar uma boa solução, gerando, inicialmente, uma população aleatória, constituída de soluções viáveis para o problema, e manipulando essas soluções através de operadores matemáticos. Tais operadores usam as soluções já existentes para produzir novas soluções que, espera-se, sejam melhores que as anteriores. Diz-se que uma população de soluções é gerada, sendo cada solução identificada como um cromossomo (ou indivíduo da população). Para cada indivíduo, determina-se um valor de aptidão escalar que representa uma medida numérica da sua capacidade em bem resolver o problema. A ideia principal é selecionar para reprodução as soluções com maior aptidão e aplicar sobre elas operações que gerem melhores soluções para o problema.

Os entes básicos de um algoritmo têm a sua denominação associada a termos da Biologia. O conjunto de pontos a partir dos quais se deseja obter a melhor solução do problema é chamado de população, sendo tais pontos chamados de indivíduos ou cromossomos. Cada indivíduo possui uma unidade básica, denominada gene, que descreve uma certa variável do problema. Cada iteração do algoritmo é denominada geração. As operações usadas sobre os indivíduos da população para gerar indivíduos mais aptos, ou seja, melhores soluções para o problema são chamadas de operadores genéticos.

A representação ou codificação das variáveis em um AG pode ser binária, ou com números inteiros ou reais. A geração da população pode ser feita de forma aleatória ou usando alguma heurística de construção. A avaliação da população é feita usando uma ou mais funções de aptidão para avaliar a qualidade das soluções (no caso de uma otimização multiobjetivo, serão usadas as funções objetivo). A seleção dos indivíduos é um mecanismo usado para, depois de avaliados os indivíduos, permitir a hereditariedade entre as gerações e,

com isso, preservar as boas características dos indivíduos. Ela pode ser feita aplicando os seguintes métodos: da roleta, do torneio e do ranking . Os operadores genéticos básicos são os de cruzamento e mutação.

O operador de cruzamento faz alterações ou combinações de sequências de informações entre indivíduos da população, gerando novos indivíduos que herdam características dos indivíduos anteriores. Para manter a diversidade das soluções, os indivíduos com características muito semelhantes, que possivelmente tenham sido gerados por este operador, devem ser eliminados. As técnicas de elaboração do operador de cruzamento levam em conta a representação computacional das variáveis com a qual se trabalha. As seguintes técnicas, por exemplo, podem ser mencionadas: cruzamento de um ponto, cruzamento multipontos, cruzamento uniforme.

O operador de mutação consiste em alterar aleatoriamente o valor de um ou mais genes de um cromossomo. Ele não aumenta o tamanho da população, somente modificando os indivíduos existentes e transformando-os em indivíduos diferentes. Caso o resultado da mutação de um indivíduo seja um indivíduo que já existe, ele é descartado. O operador de mutação permite uma maior diversidade genética e uma maior exploração do espaço de busca e evita que o algoritmo fique estacionado em mínimos locais.

Com o objetivo aumentar a velocidade de convergência do algoritmo é importante preservar e utilizar em próximas gerações as melhores soluções encontradas em uma determinada geração, o que acarreta em elitismo. Um dos aspectos mais relevantes na estruturação de um AG é o controle de seus parâmetros: tamanho da população, taxa de cruzamento e taxa de mutação.

Além dos operadores genéticos acima mencionados, um novo operador pode ser considerado, e dele se fez uso no algoritmo genético aqui apresentado. Trata-se do Operador Epidêmico, ou, simplesmente, Epidemia ([3], [9]). Tal operador é ativado sempre que seja alcançado um número pré-fixado de gerações sem que se consiga obter uma melhora desejada na população. Ele pode também ser usado para evitar ou solucionar um problema de convergência prematura. Tal operador simula a ocorrência de uma epidemia que recai sobre a população, dizimando os indivíduos menos aptos, de modo que aqueles que apresentam melhores aptidões são preservados. Os indivíduos eliminados são, então, substituídos por novos elementos, gerados pelo mesmo processo de geração da população, e o processo de evolução é reiniciado.

Há uma vasta literatura sobre a teoria de algoritmos genéticos. Para uma introdução a esta teoria, bem como um bom entendimento dos vários aspectos técnicos relacionados à representação das variáveis em um AG e aos operadores genéticos, além da referência [8] mencionada acima, pode-se recomendar, por exemplo, os seguintes livros: “Introduction to Genetic Algorithms” ([13]) e “Genetic Algorithms + Data Structures” ([10]).

2. UM ALGORITMO GENÉTICO EPIDÊMICO PARA CALIBRAÇÃO MULTI OBJETIVO DO MODELO HIDROLÓGICO IPH2

O algoritmo genético que será aqui apresentado é uma variante do algoritmo evolucionário MOCOM-UA, mencionado anteriormente, e que pode ser visto com detalhes em [15]. Será usada a notação AGE para se fazer referência a este algoritmo.

A ideia central na elaboração do AGE foi substituir o processo de evolução no MOCOM (evolução combinada com um algoritmo de busca simplex, sem o uso operações genéticas) por uma evolução genética (com uso de operadores genéticos). Seus passos são descritos no tópico seguinte, sendo os cinco primeiros e os dois últimos análogos aos do

MOCOM.

Primeiro passo: Inicialmente, são estabelecidos os limites mínimos e máximos dos valores que os k parâmetros a serem ajustados podem tomar, definindo, assim, uma região factível Ω para os parâmetros.

Segundo passo: Usando uma distribuição uniforme, são gerados ns vetores de parâmetros do modelo dentro da região factível. Cada vetor de parâmetros é um indivíduo da população inicial, a partir da qual se buscará o conjunto das soluções não dominadas do problema. Portanto ns é o tamanho da população.

Terceiro Passo: Os pontos da população são avaliados com as m funções objetivo, gerando uma matriz de resultados $R(ns, m)$.

Quarto passo: Aplica-se o seguinte Ordenamento de Pareto: identificam-se todos os indivíduos na população que são não dominados e atribui-se a eles o ranking 1. Estes indivíduos com ranking igual a 1 são temporariamente retirados da população. Em seguida, na população restante, são identificados os indivíduos não dominados. A estes indivíduos atribui-se o ranking 2 e os mesmos são, temporariamente, retirados da população. Estes passos se repetem até que a cada indivíduo da população tenha sido atribuído algum valor de ranking. Os piores indivíduos da população são os que se encontram mais distantes da frente de Pareto, possuindo o maior valor de ranking, o qual é denotado por r_{max} .

Desta forma, são definidos valores de ranking r_i para cada ponto i da população, que variam de 1 a r_{max} , sendo $r_{max} \leq ns$. Os melhores indivíduos da população inicial (não dominados) possuem ranking igual a 1, e os piores possuem ranking igual a r_{max} . Com esta ordenação, são criadas, no espaço objetivo, várias frentes de dominância que indicam o nível de dominância de uma solução em relação às demais.

Quinto passo: Formação dos Complexos: cada indivíduo com ranking r_{max} dá origem a um conjunto denominado *complexo*. Um *complexo* é formado por um indivíduo com ranking r_{max} e mais k (número de parâmetros do modelo) outros indivíduos selecionados aleatoriamente dentre os indivíduos da população que não possuem ranking r_{max} , usando a seguinte equação de probabilidade associada a cada ponto: $p_i = \frac{r_{max}-r_i+1}{\sum_{j=1}^{ns}(r_{max}-r_j+1)}$.

A probabilidade de um indivíduo ser selecionado depende, pois, de seu ranking, sendo favorecidos os indivíduos de menor ranking (melhores). Os complexos podem possuir elementos comuns, e cada complexo possui $k + 1$ 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. Neste algoritmo é proposta uma evolução genética aplicada a cada complexo. A representação computacional das variáveis usada foi a real, sendo cada parâmetro a unidade básica (gene) de cada indivíduo. Dado um complexo, aplicam-se os operadores genéticos de mutação e cruzamento em seus indivíduos, gerando-se $k + 1$ novos indivíduos (filhos) no complexo até que, considerada a hierarquização feita no complexo, 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 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 começou a ser considerada a aplicação de um *operador de epidemia*. Para tal, foi introduzido 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. E esta epidemia será 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 também foi limitada a um número máximo de iterações.

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

Oitavo passo: Teste de convergência: se o critério de parada é satisfeito, o algoritmo para. O critério de parada para o algoritmo finaliza o processo quando todos os indivíduos da população são não dominados, ou seja, $r_{max} = 1$, ou quando um número máximo de iterações é atingido.

3. CALIBRANDO O MODELO HIDROLÓGICO IPH2 COM O AGE

3.1. UMA BREVE DESCRIÇÃO DO MODELO IPH2

O modelo hidrológico IPH2 é um modelo chuva-vazão do tipo concentrado, no qual os processos hidrológicos são representados por variáveis concentradas no espaço e a bacia hidrológica é representada por uma precipitação média ([14]). Ele é composto por três algoritmos: um para as perdas por evaporação e interceptação, um para a separação do escoamento e um para a propagação dos escoamentos superficial e subterrâneos.

O algoritmo de perdas por evaporação e interceptação utiliza um único parâmetro, $RMAX$, que descreve a capacidade máxima de armazenamento em um reservatório de perdas. Este parâmetro representa a lâmina de água interceptada pela vegetação ou armazenada em poças, e que fica facilmente disponível para a evaporação ([4]).

No algoritmo de separação do escoamento, há três parâmetros: I_0 e I_b , que representam, respectivamente, as capacidades inicial e mínima de infiltração no solo, e H , que representa a função de variação da capacidade de infiltração no solo, sendo dado pela equação $H = e^{-K}$, onde K é um parâmetro empírico relacionado ao tipo de solo, caracterizando o decaimento exponencial da curva de infiltração ([14]).

A propagação do escoamento superficial utiliza parâmetros de tempo de concentração que podem ser mantidos fixos ou calibrados, dependendo da disponibilidade de informações de características físicas da bacia. Utilizando a teoria do reservatório linear simples ([14]) para se considerar o efeito do armazenamento na bacia, tem-se o parâmetro K_s , que representa o tempo de retardo do escoamento superficial. Deve ser definido o percentual de área impermeável da bacia.

Na propagação do escoamento subterrâneo, utiliza-se também um modelo do reservatório linear simples com o parâmetro K_{sub} , que representa o tempo médio de esvaziamento do reservatório subterrâneo de escoamento. Para a simulação de séries contínuas de longo período, foi introduzido um parâmetro, α , que representa o escoamento direto de áreas impermeáveis.

Para a simulação do processo, além dos parâmetros acima, devem ser informadas as áreas da bacia e as variáveis de entrada do modelo: precipitação e evaporação ([2]).

3.2. ALGUNS RESULTADOS DE CALIBRAÇÃO DO IPH2

Foram efetuados testes de calibração do modelo IPH2 com o AGE e com o MOCOM. O modelo direto foi executado com os valores pré-fixados dos sete parâmetros do modelo ([1]) e com dados reais de precipitação e evaporação (301 registros) de uma bacia do rio Canoas, em Santa Catarina, cuja área é de 989 km^2 . Uma série de vazões, \bar{Q}_t , foi, então,

obtida. A esta série foi adicionado um ruído de 5%, ou seja, gerou-se uma nova série, \hat{Q}_t , através da equação $\hat{Q}_t = \bar{Q}_t(1 + \beta \cdot \gamma)$, onde $\alpha = 0.05$ e γ é um número aleatório no intervalo $[-1,1]$. A série \hat{Q}_t foi utilizada como a série de vazões observadas para o procedimento de calibração. O período dos dados foi de 11/03/1983 a 05/01/1984 e o intervalo de tempo considerado foi de um dia.

O AGE foi implementado em FORTRAN 90. Os testes foram realizados em um laptop HP Pavilion dv4, com processador AMD Turion(tm)II Dual-Core Mobile M520 de 2.30 GHz e 4GB de memória RAM, em ambiente Windows 7.

Foram consideradas nos testes as funções objetivo F_1 (desvio padrão das vazões) e F_2 (desvio padrão dos inversos das vazões), descritas nas equações 3 e 4, respectivamente.

Como visto na descrição do AGE, além do tamanho da população inicial, ns , que é o único parâmetro do algoritmo MOCOM, outros parâmetros foram inseridos. Os testes foram sendo feitos, partindo-se, inicialmente, de uma população com tamanho $ns=500$, considerado como bastante satisfatório para a calibração pretendida, utilizando o MOCOM ([1]) e, procurando-se ajustar, além do tamanho da população, os demais parâmetros do AGE. Com este tamanho da população, verificou-se uma dificuldade do AGE na aproximação da frente de Pareto, embora os valores obtidos para as funções objetivo fossem muito próximos dos obtidos com o MOCOM. Um número muito grande de interações se dava sem que se conseguisse chegar a uma população em que todos os indivíduos fossem não dominados. A Figura 2 mostra um dos resultados desses primeiros testes, ilustrando a dificuldade acima mencionada.

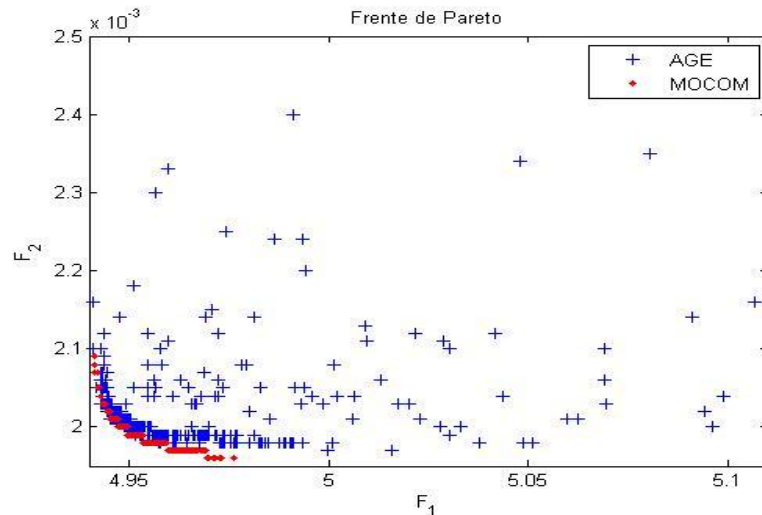


Figura 2: Espaço objetivo resultante da calibração do IPH2 ($ns=500$)

Resultados bastante satisfatórios para os testes com o AGE foram obtidos quando se considerou o tamanho da população $ns=50$. A partir de vários testes, estabeleceu-se o seguinte ajuste dos parâmetros do AGE: *número máximo iterações*: $MAX_{IT} = 5000$; *taxa de mutação*: $TAX_{MUT} = 0,03$; *máximo de tentativas de evolução de um complexo*: $MAX_{GER} = 100$; *taxa de ativação da epidemia* (percentual de complexos sem evolução a partir do qual a epidemia é ativada): $TAE = 0,70$; *taxa de epidemia* (indica em quais indivíduos a epidemia será aplicada): indivíduos com valores de ranking maiores ou iguais a $\frac{r_{max}}{2} + 1$ serão eliminados; *até quando serão aplicadas epidemias*: até que o número de iterações alcance a metade do valor do parâmetro MAX_{IT} .

A Figura 3 mostra as frentes de Pareto resultantes da calibração do IPH2 com o MOCOM e com o AGE, considerando uma população de tamanho $ns=50$. Observa-se que na calibração com o AGE a frente de Pareto está mais próxima da origem, o que é um bom indicador de sua melhor qualidade.

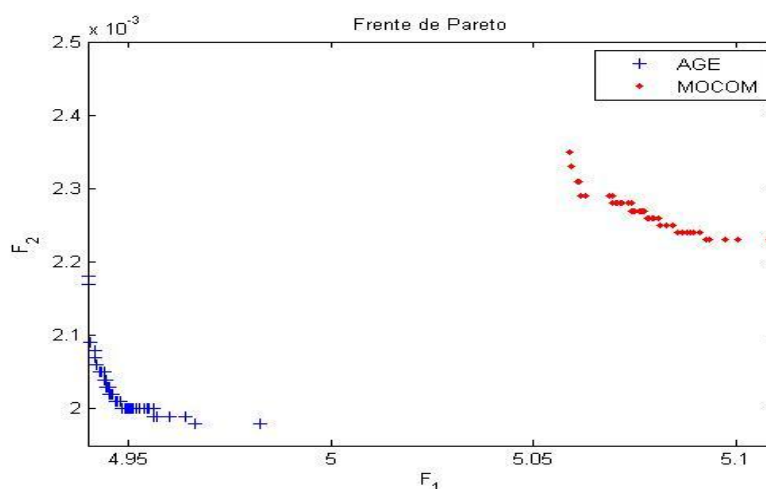


Figura 3: Espaço objetivo resultante da calibração do IPH2 ($ns=50$)

As duas estratégias, MOCOM e AGE se mostraram bastante satisfatórias na recuperação dos valores pré-definidos para os parâmetros do modelo IPH2.

Na Tabela 1, são apresentados os valores máximo e mínimo de cada um dos parâmetros obtidos com a calibração automática do modelo IPH2, tanto com o AGE quanto com o MOCOM, a partir de uma série de calibrações (todas considerando uma população de tamanho $ns=50$). Para ambas as estratégias a faixa de variação dos parâmetros foi relativamente estreita, o que é uma boa indicação da validade do modelo hidrológico.

Tabela 1: Faixas de variação dos parâmetros do IPH2 a partir de uma série de calibrações

Parâmetro	Unidade	Valores Exatos	Faixa de variação MOCOM		Faixa de variação AGE	
I_0	$mm\Delta t^{-1}$	18,20	11,72	23,45	22,60	27,04
I_b	$mm\Delta t^{-1}$	2,95	2,64	3,05	2,87	3,08
H	--	0,13	0,028	0,077	0,035	0,40
K_s	Δt	5,72	5,61	5,68	5,60	5,73
K_{sub}	Δt	40,25	35,89	47,53	35,76	43,29
$RMAX$	mm	0,24	0,01	1,53	0,44	0,86
α	--	12,66	4,89	9,94	17,04	20,00

A Figura 4 apresenta um gráfico gerado a partir de todos os parâmetros (normalizados) que constituem os indivíduos não dominados (conjunto de Pareto) obtidos ao final de uma das calibrações feitas com o AGE. A Figura 5 apresenta os hidrogramas (vazões pelo tempo), com a vazão observada e uma série de vazões estimadas, cada uma correspondente a um dos pontos no conjunto de Pareto da Figura 4. Como se pode observar na Figura 4, as soluções de Pareto tendem a se aglomerar com bastante proximidade. No entanto, o parâmetro α apresenta uma maior variação, indicando uma tendência de que ele seja

diferente nos dois objetivos considerados na calibração; este parâmetro certamente desempenha um importante papel na forma dos hidrogramas estimados.

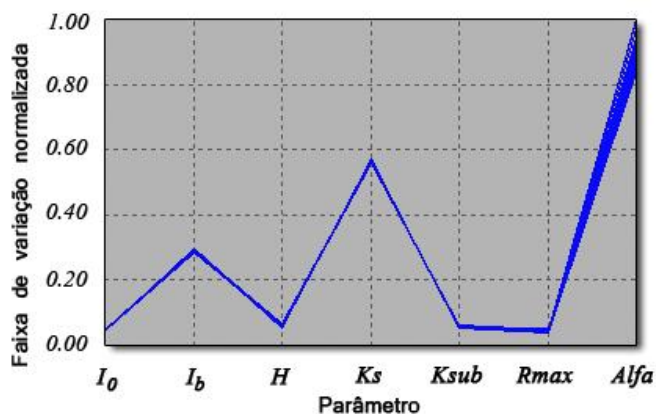


Figura 4: Conjunto de Pareto (calibração do IPH2 com o AGE)

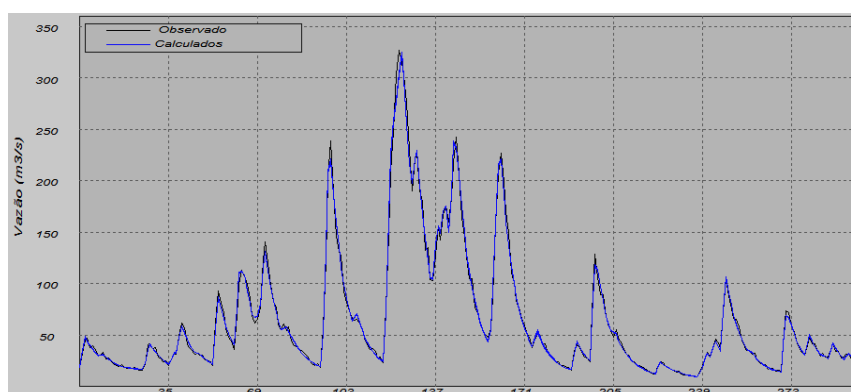


Figura 5: Vazões observadas e vazões calculadas com o conjunto de Pareto

4. CONCLUSÕES

O desenvolvimento dos computadores, cada vez mais rápidos e mais potentes, tem tornado possível o surgimento e o aprimoramento de várias técnicas computacionais para se resolver problemas de calibração dos parâmetros que surgem na modelagem de um dado fenômeno. No caso dos modelos hidrológicos, a calibração multiobjetivo tem ganhado cada vez mais espaço, uma vez que ela permite uma melhor avaliação das incertezas, das imperfeições do modelo e da representatividade dos parâmetros.

Este trabalho apresentou um algoritmo genético epidêmico (AGE) como uma possível alternativa para a calibração de modelos hidrológicos, tendo sido testado na calibração do modelo hidrológico IPH2, um modelo já estabelecido para pequenas bacias, eficiente na simulação dos processos hidrológicos e aplicado em várias bacias brasileiras, usando o algoritmo MOCOM.

Os resultados aqui apresentados foram produzidos a partir de testes baseados em uma série sintética de vazões. O AGE mostrou-se eficiente na recuperação dos valores pré-fixados dos parâmetros usados para gerar a série sintética de vazões e também para obter uma boa aproximação da frente de Pareto. Acredita-se, pois, que este algoritmo pode se constituir como uma boa alternativa para a calibração automática em outros modelos hidrológicos mais sofisticados, nos quais muito mais variáveis do ciclo hidrológico, tais como a evapotranspiração e o nível de água subterrâneo, possam ser levadas em conta.

5. REFERÊNCIAS BIBLIOGRÁFICAS

- [1] Bravo, J. M., Collischon, W., Tucci, C. E. M., 2009, “Verificação de eficiência de um algoritmo evolucionário multiobjectivo na calibração automática do modelo hidrológico IPH II”, Revista Brasileira de Recursos Hídricos-RBRH, Vol. 14, pp.37-60, Porto Alegre, RS.
- [2] Bravo, J. M., Allasia, D. G., Collischom, W., Tassi, R., Meller, A., Tucci, C. M., 2007 Avaliação visual e numérica da calibração do modelo IPH II com fins educacionais. In: XVII Simpósio Brasileiro de Recursos Hídricos, São Paulo (SP), 2007.
- [3] Chiwiacowsky, L., Campos Velho, H. F., 2003, “Different approaches for the solution of a backward heat conduction problem”, Inverse Problems in Engineering, Vol.11, n.6, pp.471-494.
- [4] Collischonn, W., Tucci, C. E. M., 2003 Ajuste multiobjetivo dos parâmetros de um modelo hidrológico. Revista Brasileira de Recursos Hídricos, v.8, n.3, 27-41p.2003.
- [5] Deb, K., Pratap, A., Agarwal, S., Meyarivan, T., 2002, “A Fast and Elitist Multi-objective Genetic Algorithm: NSGA-II”, IEEE Transactions on Evolutionary Computational, vol.6, no.2, pp.182-197.
- [6] Erickson, M., Mayer, A., Horn, J., 2002, “Multi-objective optimal design of groundwater remediation systems: application of the niched Pareto genetic algorithm (NPGA)”, Advances in Water Resources, Vol.25, pp.51-65.
- [7] Goicoechea, A., Hasen, D. R., Dukstein, L., 1982, “Multi-objective decision analysis with engineering and business applications”, 511 pp., John Wiley, New York, NY.
- [8] Goldberg, D. E., 1989, “Genetic algorithms in search, optimization, and machine learning,” Addison-Wesley Pub. Co., New York, NY.
- [9] Medeiros, F. L. L. Algoritmo genético híbrido como um método de busca de estados estacionários de sistemas dinâmicos. 2002. 191p. (INPE-13913-TDI/1058). Dissertação (Mestrado em Computação Aplicada) - Instituto Nacional de Pesquisas Espaciais (INPE), São José dos Campos (SP), 2002.
- [10] Michalewicz, Z., 1996, “Genetic Algorithms + Data Structures = Evolution Programs”, 388 p., Springer, London, UK, Third Revised and Extended Edition.
- [11] Nelder, J. A., Mead, R., 1965, “A simplex method for function minimization, Computer Journal, 7, pp.308-313.
- [12] Schaffer, J. D., 1984, “Multiple objective optimization with vector evaluated genetic algorithms”, Thesis (PhD), Vanderbilt University, Nashville, TN.
- [13] Sivanandam, S. N. and Deepa S. N., 2008, “Introduction to Genetic Algorithm”, 442 pp., Springer, New York, NY.
- [14] Tucci, C. E. M., 1998, “Modelos hidrológicos”, 669 pp., ABRH Editora da UFRGS, Porto Alegre, RS.
- [15] Yapo, P. O., Gupta, H. V., Sorooshian, S., 1998, “Multi-objective global optimization for hydrological models”, Journal of Hydrology, Vol. 204, pp. 83-97.
- [16] Zadeh, L. A., 1963, “Optimality and non-scalar valued performance criteria”, IEE Transactions, C8, 1.