

## EXTRAÇÃO DE REGRAS OPERACIONAIS ÓTIMAS DE SISTEMAS DE DISTRIBUIÇÃO DE ÁGUA ATRAVÉS DE ALGORITMOS GENÉTICOS E APRENDIZADO DE MÁQUINA

Ivaltemir (Tinil) Barros Carrijo<sup>1</sup>; Luisa Fernanda Ribeiro Reis<sup>1</sup>

**Resumo** - O crescimento desordenado das cidades, aliado à falta de financiamento para o setor de saneamento básico, tornaram os sistemas de abastecimento de água complexos e de difícil operacionalidade. A operação eficiente do sistema é uma ferramenta fundamental para que sua vida útil se prolongue o máximo possível, garantindo o perfeito atendimento aos consumidores, além de manter os custos com energia elétrica e manutenção dentro de padrões aceitáveis. Para uma eficiente operação, é fundamental o conhecimento do sistema, pois, através deste, com ferramentas como modelos de simulação hidráulica, otimização e definição de regras operacionais, é possível fornecer ao operador condições ideais para a operação das unidades do sistema, não dependendo exclusivamente de sua experiência pessoal, mantendo a confiabilidade do mesmo. Este trabalho propõe o desenvolvimento de um modelo computacional direcionado ao controle operacional ótimo de sistemas de macro distribuição de água potável, utilizando o simulador hidráulico *EPANET2*, os algoritmos genéticos multiobjetivo como ferramenta para a otimização e o algoritmo de aprendizado de máquina para extração de regras operacionais para o sistema. Pretende-se aplicar o modelo em uma parte do macro sistema distribuidor de água da cidade de Goiânia.

**Abstract** - The growth of cities, associated with the lack of investment in basic infrastructure, has rendered water supply systems complex and difficult to operate. The efficient operation of such systems is a fundamental tool for extending the system's service life as much as possible, thus ensuring good service to the consumers while keeping electrical energy and maintenance costs at acceptable levels. Efficient operation requires knowledge of the system, supported by tools such as models for hydraulic simulation, optimization, and definition of rules, provides the operator with proper conditions for the rational operation of the system's units without depending exclusively on personal experience while maintaining the system's reliability. This paper aims to develop a computational model for the optimal operational control of macro water distribution systems using the *EPANET2* hydraulic simulator, multi-objective genetic algorithms as an optimization tool, and machine learning to extract operational rules for the system. The studies were conducted on the macro system of the city of Goiânia, showing that solutions for its satisfactory operation can be quickly produced as a substitute to the personal judgment of the operator.

**Palavras-chave:** otimização multiobjetivo, algoritmos genéticos, otimização operacional, aprendizado de máquina.

## INTRODUÇÃO

A operação de um sistema urbano de abastecimento de água em tempo real é uma tarefa muito complexa, devido à necessidade de se garantir confiabilidade no atendimento dos serviços, economia no uso dos equipamentos (energia elétrica e manutenção), garantia de atendimento às

---

<sup>1</sup> Escola de Engenharia de São Carlos – Universidade de São Paulo – Dep. Hidráulica e Saneamento – São Carlos – S.P. - CEP: 13566-590 - Fone: 0xx16-3373-9545 Fax: 0xx16-3373-8269 E-mail: [tinil@sc.usp.br](mailto:tinil@sc.usp.br); [fernanda@sc.usp.br](mailto:fernanda@sc.usp.br)

demandas com pressões desejadas e retardamento de investimento para expansão das unidades. Além disso, a confiabilidade deve permitir o atendimento ao consumidor em condições anormais ocasionadas por avarias em unidades do sistema.

Segundo Zahed Filho (1990) o conceito de operação de sistemas é entendido por leigos como uma mera seqüência de comandos de equipamentos, que têm como objetivo o atendimento da demanda. Entretanto, na realidade, o problema é muito mais amplo, envolvendo aspectos de planejamento, controle e supervisão, serviços de infra-estrutura de apoio e atendimento ao usuário, todos considerados simultaneamente e interdependentes entre si.

Para que o planejamento possa definir a melhor regra de operação, são necessárias, pelo menos, quatro condições básicas: a) a definição clara dos objetivos a serem alcançados; b) a disponibilidade de modelos matemáticos de análises; c) os equipamentos para seu processamento; e d) o conhecimento do sistema.

Na busca de maior eficiência para os sistemas públicos de abastecimento de água, novas técnicas hidráulicas associadas a algoritmos de otimização têm sido desenvolvidas, buscando uma maior confiabilidade no estabelecimento de regras operativas do sistema.

De acordo com Righetto (2002), sistemas de médio e grande porte são projetados e operados com base numa eficiência global, envolvendo questões como: confiabilidade, distribuição de pressões e demandas, consumo de energia, minimização de perdas, etc. Nesta abordagem global, questões de eficiência numérica são associadas à obtenção de respostas às questões complexas operacionais, o que impulsiona os pesquisadores a desenvolverem técnicas numéricas apropriadas que permitam resolver problemas específicos ou gerais, de alta complexidade.

Em problemas relativos a operação de sistemas de distribuição de água, contendo várias elevatórias, adutoras, reservatórios e válvulas reguladoras, vários métodos podem ser utilizados para a simulação hidráulica e otimização do sistema. A interface entre os modelos de simulação hidráulica, de otimização e de definição de regras operacionais, deve ser cuidadosamente elaborada, no sentido de dar transparência ao modelo, de forma a facilitar a introdução de inequações restritivas e facilitar a obtenção dos valores da função objetivo, nos sucessivos passos requeridos pelo otimizador. Com a ampliação dos objetivos de análise de um sistema de distribuição, o algoritmo de otimização deve ser escolhido de forma a facilitar o seu uso e, ao mesmo tempo, permitir a análise de problemas complexos, envolvendo critérios de múltiplos objetivos.

Atualmente, algoritmos baseados em processos estocásticos de busca, vêm sendo utilizados com sucesso, com possibilidades de aplicação em várias áreas de conhecimento, devido à facilidade de implementação e a quase ilimitada condição de aplicabilidade de seus recursos. Os algoritmos genéticos (AGs) são algoritmos de busca baseados na seleção natural e na genética de evolução populacional e, especialmente adequados para tratar problemas complexos como os relativos aos sistemas de distribuição de água. Apresentam vantagem sobre as técnicas de otimização convencionais quanto à possibilidade de analisar diretamente populações de soluções e, através de um aprimoramento de populações sucessivas, chegar a soluções com alta performance quanto aos critérios múltiplos definidos pelo problema. Grande número de aplicações de AGs demonstra a sua adequação quanto à obtenção de solução, senão a ótima, pelo menos com elevada performance quanto aos critérios multiobjetivos impostos.

Objetiva-se, com este trabalho, apresentar uma metodologia para uma operação ótima de sistemas de distribuição de água, essencialmente o sistema macro adutor (esqueleto), composto por adutoras, elevatórias, reservatórios e válvulas de controle automático, envolvendo parâmetros relacionados com os custos de operação, confiabilidade operacional e satisfação quanto ao atendimento das demandas, além do fornecimento de regras adequadas ao atendimento desses requisitos.

Pretende-se simular hidraulicamente o sistema, utilizando o EPANET2 (Rossman, 2002), otimizá-lo operacionalmente através de AGs multiobjetivo e fornecer regras operacionais através do aprendizado de máquina (AM), objetivando fornecer ao operador uma condição de investigar as conseqüências das diferentes ações de operação antes da sua efetiva implementação. Se o sistema

de supervisão e análise estiverem interligados, o operador terá uma ferramenta apropriada para ajudá-lo a gerenciar o sistema.

## METODOLOGIA

Técnicas de otimização vêm sendo tradicionalmente usadas no dimensionamento de unidades de um sistema de distribuição de água, objetivando menor custo. Segundo Righetto (2002), a interface entre o modelo de simulação hidráulica e o modelo de otimização deve ser cuidadosamente elaborada no sentido de dar transparência ao modelo, de forma a facilitar o seu uso e, ao mesmo tempo, permitir a análise de problemas complexos, envolvendo múltiplos objetivos. Atualmente, algoritmos baseados em processo estocástico de busca, vêm sendo empregados com sucesso, devido à facilidade de aplicação e praticamente à ilimitada condição de uso. É um modelo de otimização que possibilita a introdução de regras heurísticas. O AG é um algoritmo de busca baseado na seleção natural e na genética de evolução populacional. Tem a vantagem sobre as demais técnicas, por analisar diretamente, uma população de soluções factíveis e, através de um aprimoramento de populações sucessivas, chegar a soluções com alta performance quanto aos critérios multiobjetivos definidos pelo problema.

Além de custos operacionais, um objetivo importante, quando se pretende otimizar operacionalmente um sistema de abastecimento de água potável, é a avaliação de sua confiabilidade. Segundo Bao e Mays (1990), a confiabilidade é a medida da performance (desempenho) do sistema. A performance depende das condições ideais que viabilizam fatores como: interação entre o sistema de transporte, conservação da capacidade de armazenamento, conservação da capacidade de distribuição e sistema de segurança, como o funcionamento adequado de válvulas reguladoras e bombas; confiabilidade dos componentes individuais do sistema; e variação espacial e temporal das demandas.

O objetivo da operação de sistemas de abastecimento de água é o atendimento, com riscos aceitáveis, das necessidades de consumo e minimização dos custos operacionais e, de maneira implícita, um melhor aproveitamento do sistema de transporte e reservação, de forma a retardar investimentos com ampliações. A operação, é a seqüência de manobras exercidas sobre os elementos ativos do sistema, como válvulas e bombas, de forma a atender os objetivos.

A otimização implementada neste trabalho, leva em consideração dois objetivos: a minimização de custos operacionais e a maximização dos benefícios hidráulicos, considerando o índice de atendimento da demanda, os níveis adequados de água nos reservatórios e as pressões mínimas e máximas nos pontos de demanda, para um período de análise de 24 horas.

### Simulação hidráulica do sistema

A simulação hidráulica avalia as respostas do sistema face às decisões operacionais em termos das suas variáveis de estado: pressão, vazão e níveis de água nos reservatórios. É, portanto, uma ferramenta essencial à rotina computacional que quantifica os objetivos estabelecidos. Utiliza-se, neste trabalho, o *EPANET2*, via *Toolkit* (Rossman, 2002) para este fim.

### Otimização operacional utilizando algoritmos genéticos multiobjetivo (MOGAs)

Durante décadas, considerável esforço tem sido utilizado para o desenvolvimento de algoritmos de otimização e modelos para sistemas de abastecimento de água. Em muitos casos, o primeiro objetivo desses modelos é minimizar os custos do sistema. Entretanto, na prática, um modelo ótimo para um sistema de água é um complexo processo de objetivos múltiplos envolvendo uma negociação (*trade-off*) entre os custos e seus benefícios. Operação integrada entre bombas, válvulas e reservatórios deve ser empregada para a obtenção da máxima performance do sistema, Xu e Goulter (1999).

As técnicas de otimização vêm sendo utilizadas na busca de soluções ótimas para problemas operacionais específicos. Se o objetivo for, por exemplo, o mínimo custo operacional, a função

custo estará associada à tarifa de energia elétrica, ao rendimento das bombas, às perdas de carga nas instalações, etc. Por outro lado, para o atendimento do objetivo de mínimo custo, há restrições impostas pelo próprio sistema, tais como, níveis máximos e mínimos dos reservatórios, limites de pressão e de potência e quantidade de água disponível.

Algumas dificuldades decorrem da utilização das técnicas de otimização para a solução do problema do controle. Estas dificuldades estão associadas ao grande número de equações a serem resolvidas, à dificuldade de formulação do problema geral, ao tratamento de restrições operacionais não explícitas e à aquisição e manutenção de dados.

Muitos problemas reais envolvem a otimização com base em mais de um critério ou objetivo. Uma abordagem simples para tratar de um conjunto de funções objetivo  $f_1, f_2, \dots, f_n$ , é compor uma nova função objetivo  $F$ , obtida da soma ponderada das funções objetivo, ou :

$$F = \sum_{i=1}^n w_i f_i \quad (1)$$

em que  $w_i$  é o peso atribuído ao objetivo  $i$  de maneira que ele assuma maior ou menor importância relativa, na avaliação da qualidade da solução. A minimização de  $F$  minimiza todas as funções  $f_i$ . No entanto, as funções objetivo podem ser conflitantes, no sentido de que, otimizar uma função, degrada a otimização da outra função.

As abordagens mais gerais para este tipo de problema envolvem a aplicação do conceito de não dominância de Pareto, segundo o qual uma solução  $x_1$  é dita não-dominada por uma solução  $x_2$ , quando ela é, no mínimo, tão boa quanto a solução  $x_2$ , de acordo com todas as funções e,  $x_1$  é melhor que  $x_2$ , de acordo com pelo menos uma das funções, ou seja:

$$\left( \forall_i : f_i(x_1) \geq f_i(x_2) \right) \text{ e } \left( \exists_j : f_j(x_1) > f_j(x_2) \right), \text{ no caso de maximização}$$

A meta do AG multiobjetivo é encontrar o conjunto Pareto-ótimo (também chamado de fronteira Pareto ótimo), que corresponde ao conjunto de soluções não-dominadas do espaço de busca. Seja  $S$  o espaço de busca, o conjunto Pareto-ótimo  $P$  é dado por:

$$P = \{x_i \in S \mid \nexists x_j \in S : x_i \text{ é dominado por } x_j\}$$

Uma solução não-dominada, também chamada de solução não inferior é uma solução segundo a qual um objetivo pré-definido não pode ser melhorado sem que o outro objetivo sofra uma degradação.

Os AGs são naturalmente apropriados à análise da frente Pareto, pelo fato de trabalharem com populações de solução, produzindo como respostas um conjunto de soluções ótimas ao invés de uma única solução.

A partir de 1993, diferentes algoritmos evolucionários foram sugeridos para resolver problemas de otimização multiobjetivo. Fonseca e Fleming (1993) com o *Multi-Objective Genetic Algorithm (MOGA)*, Horn et al. (1994) com o *Niched-Pareto Genetic Algorithm (NPGA)* e Srinivas e Deb (1994) com o *Nondominated Sorting Genetic Algorithm (NSGA)*, foram os precursores desta técnica, cujas características básicas são: avaliação dos membros de uma população com base no conceito de dominância Pareto e preservação da diversidade de soluções. Tais algoritmos mostraram-se eficientes na obtenção de múltiplas soluções não dominadas, para vários problemas de engenharia. No entanto, pesquisadores têm sugerido a introdução do elitismo para melhorar as suas propriedades de convergência. Dentre os algoritmos evolucionários multiobjetivo que

consideram o elitismo, destacam-se os algoritmos *SPEA* e *SPEAII* (*Strength Pareto Evolutionary Algorithm* - ZITZLER e THIELE, 1998 e ZITZLER et al., 2001), *PAES* (*Pareto-Archived Evolution Strategy* - KNOWLES e CORNE, 2000), AG elitista de Rudolph, 1998, *PESA* e *PESAII* (*Pareto Envelope-based Selection Algorithm*-CORNE et al., 2000 ) e *NSGAII* (*Nondominated Sorting Genetic Algorithm*-DEB et al., 2002).

Neste Trabalho, após testar o problema proposto utilizando os métodos *NSGA* , *NSGAII* e *SPEA*, decidiu-se empregar este último por ter apresentado os melhores resultados quando comparados os valores e a relação entre os objetivos. O método *SPEA*, baseado no elitismo, foi implementado através dos códigos apresentados por Andrzej Jaszkiewicz na *Multiobjective Methods Metaheuristic Library* (*MOMHLib++*) para C++.

### **Extração de regras operacionais utilizando aprendizado de máquina (AM)**

De acordo com Weiss e Kulikowski (1991), um sistema de aprendizado indutivo pode ser definido como um programa de computador que toma decisões baseadas na experiência contida em exemplos solucionados com sucesso. Os modelos de classificação construídos por estes sistemas podem ser desenvolvidos por dois caminhos principais. O primeiro obtém regras para o modelo através de entrevistas com peritos e inclusão do conhecimento prévio ao sistema. O segundo cria um modelo indutivo através da generalização de um grande registro de dados que foi coletado e classificado. Para aplicação dos algoritmos, devem ser analisadas algumas características específicas dos dados. As informações sobre cada caso (ou exemplo) do registro de dados são definidas através de atributos e cada caso é designado para pertencer a uma classe discreta e pré-definida. Toda informação sobre os casos deve ser apresentada na forma de atributos.

Um fator importante que deve ser considerado é que os classificadores devem fornecer uma descrição mais compacta do conceito embutido nos dados caso esses classificadores sejam analisados por seres humanos.

Uma das principais funções de um programa de aprendizado de máquina é construir os modelos de classificação na forma de árvores de decisão, para posterior aplicação. Neste trabalho, este não é o principal objetivo. O classificador denominado conjuntos de regras (*rulesets*) é utilizado para a seleção da(s) melhor(es) estratégia(s) e posterior extração de regras operacionais a partir de um conjunto de exemplos (casos) fornecidos pelo modelo de otimização através das soluções ótimas pertencentes à frente Pareto. Para tanto, é utilizado o indutor *SEE5*, que é a mais recente versão do indutor *C4.5* descrito por Quinlan (1993).

A figura 1 apresenta um fluxograma representativo da metodologia proposta.

### **DESCRIÇÃO DO PROBLEMA**

O propósito deste trabalho é desenvolver uma ferramenta que forneça um conjunto claro de regras operacionais, atendendo os objetivos pré-definidos e as condições de funcionamento de cada unidade do sistema. Para análise e avaliação dos resultados, foi utilizada uma parte do sistema de macro distribuição da cidade de Goiânia. Para melhor entendimento da aplicação proposta, apresenta-se um esquema do sistema estudado (figura 2), com suas características principais (tabelas 1 e 2).

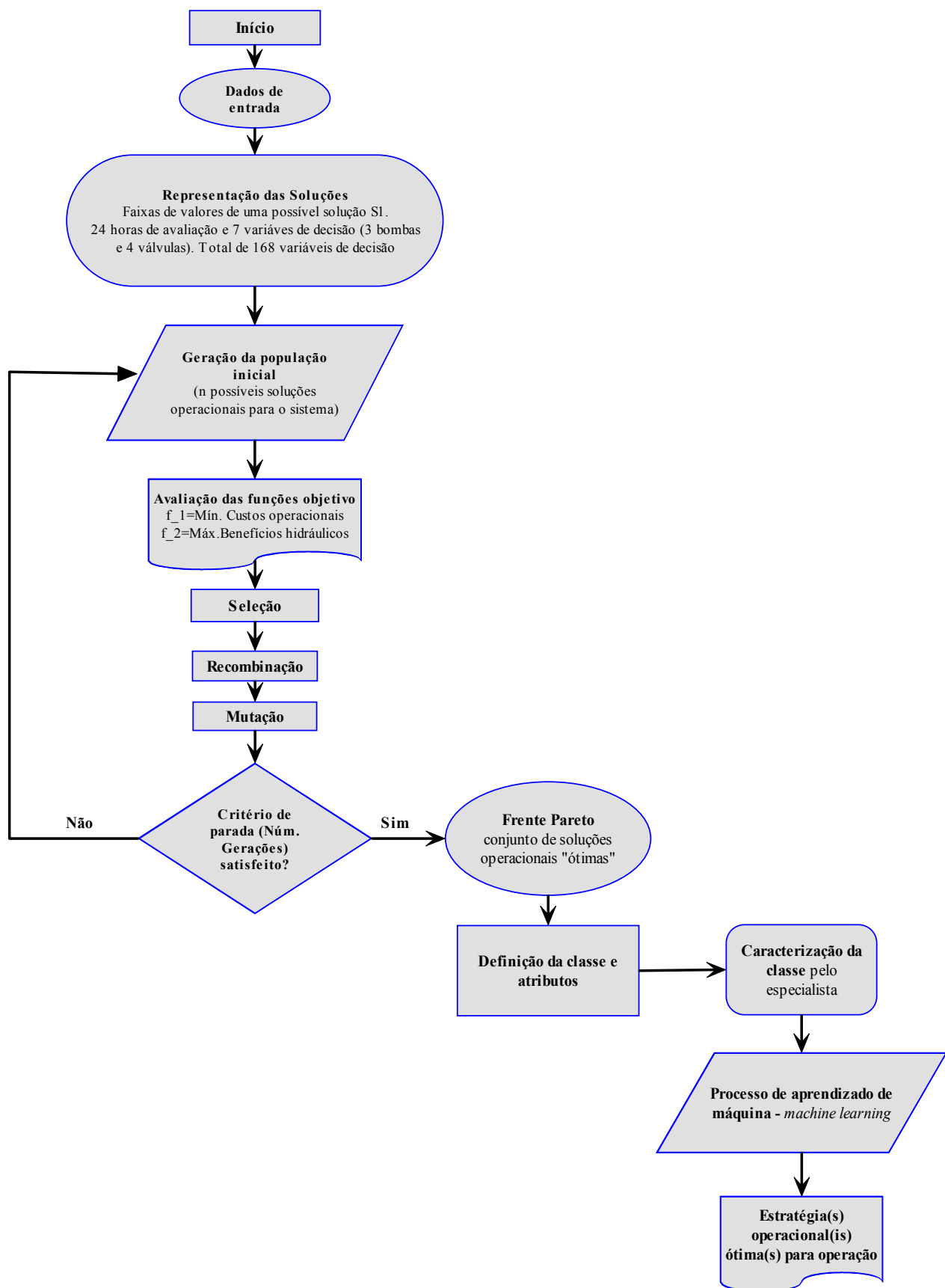


Figura 1 – Fluxograma representativo do modelo proposto

### Definição das funções objetivo

Vários objetivos podem ser listados quando avalia-se um sistema real de abastecimento de água, visando a sua operação ótima. Um deles é a minimização do custo operacional, incluindo a manutenção e a operação do sistema. Vários trabalhos já desenvolvidos demonstraram que, de todos os parâmetros relacionados a esta questão, o mais relevante é o custo relativo ao consumo de energia elétrica nas estações elevatórias de água. Um outro objetivo que pode ser avaliado é a confiabilidade do sistema em atender os consumidores de forma adequada. Neste caso, vários parâmetros podem ser enumerados.

A confiabilidade dos sistemas de abastecimento de água pode ser considerada sob o aspecto hidráulico ou mecânico. O primeiro está relacionado aos parâmetros físicos, mutáveis em função de mudanças operacionais do sistema, enquanto o segundo considera as possíveis intervenções sobre os equipamentos. Confiabilidade mecânica é a capacidade que os componentes do sistema têm de atender continuamente as necessidades de operação sem a necessidade de freqüentes reparos, modificações ou substituições de componentes ou sub-componentes. Confiabilidade hidráulica é definida como a capacidade que o sistema tem de atender as demandas dos usuários com um aceitável nível de interrupção, independente de condições anormais que possam ocorrer.

Este trabalho propõe avaliar dois objetivos básicos. O objetivo econômico e o objetivo de confiabilidade hidráulica, aqui denominado benefícios, relativos aos sistemas de macro distribuição de água potável, sendo que este último considera três parâmetros distintos: suprimento das vazões nos nós de demanda com pressões adequadas, atendimento de nível mínimo nos reservatórios e pressões mínimas nos referidos nós para garantia do abastecimento de água nos pontos de consumo.

No caso do objetivo econômico, pretende-se minimizar os custos relativos ao consumo de energia elétrica na estação elevatória de água. O custo diário para cada bomba de uma estação elevatória é dado pela soma do custo relativo ao fator de demanda máxima mais o custo do consumo medido.

De acordo com dados fornecidos pela Saneamento de Goiás S.A.-SANEAGO, para a estação elevatória Senac, os valores atualmente utilizados são os seguintes:

- para o consumo no horário de pico (18 – 21hs), R\$ 0,17076/kWh e fora deste horário, R\$ 0,0816/kWh;

- para a demanda máxima no período entre 18 e 21hs o valor total médio é de R\$ 26,38/kW e, fora deste intervalo, R\$ 8,66/kW.

Adotando como objetivo econômico a minimização dos custos de energia elétrica, relativos à operação de estações elevatórias, a função objetivo (FO\_1) pode ser expressa utilizando a seguinte equação:

$$FO\_1 = \sum_{t=1}^{24} \sum_{k=1}^{nb} Cu(t) * \frac{Q(k,t) * H(k,t) * \gamma}{\eta(k,t)} + D * Tarifa \quad (2)$$

onde :

k = número de bombas na estação elevatória;

t = hora do dia;

nb = número total de bombas na elevatória;

Cu(t) = custo unitário da tarifa(R\$/kWh);

Q(i, t) = vazão bombeada(m<sup>3</sup>/s);

H(i, t) = carga hidráulica(m);

η(i, t) = rendimento do conjunto(%);

D = fator de demanda máxima(kW);

Tarifa = tarifa correspondente ao fator de demanda máxima(R\$/kW)

A quantidade de água que um sistema de abastecimento pode efetivamente distribuir aos usuários, com uma adequada pressão, é um dos principais fatores na determinação da performance e da confiabilidade deste sistema. Para tanto, a relação entre as demandas e as cargas nodais deve ser incorporada na medida da confiabilidade.

Como citado anteriormente, a função objetivo benefício hidráulico será formulada para quantificar os benefícios resultantes de pressões adequadas nos nós de demanda, níveis adequados nos reservatórios de distribuição e garantia de atendimento às demandas horárias.

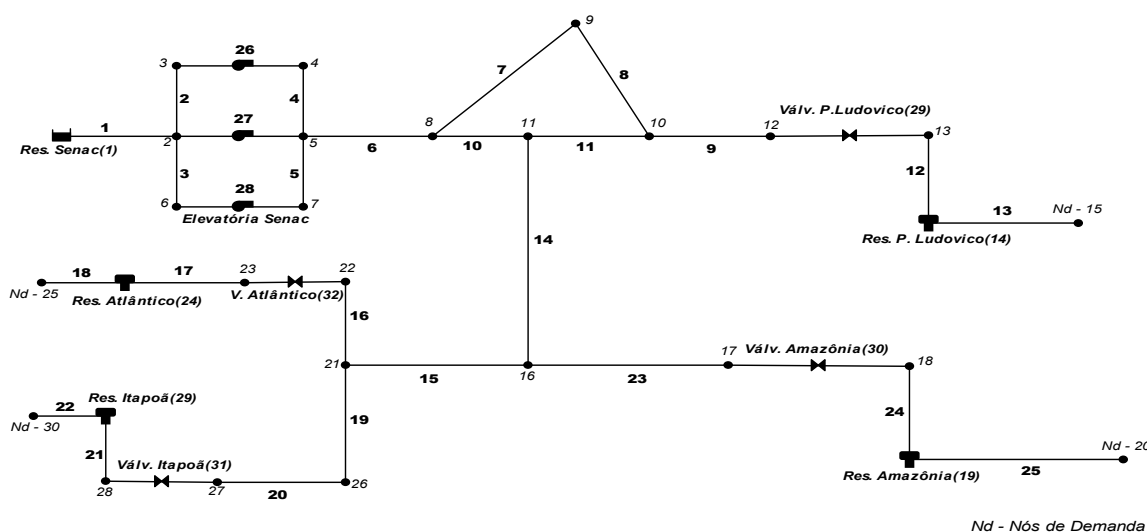


Figura 2 – Esquema do sistema de macro distribuição de Goiânia

Tabela 1 – Goiânia – dados dos reservatórios

| Reservatórios  | Volume( $m^3$ ) | Nív.Mín(m) | Nív.Máx(m) | Elevation(m) |
|----------------|-----------------|------------|------------|--------------|
| P.Ludovico(14) | 10.000          | 1,5        | 6,0        | 858,0        |
| Amazônia(19)   | 5.000           | 1,5        | 5,5        | 861,5        |
| Atlântico(24)  | 10.000          | 1,5        | 7,0        | 836,5        |
| Itapoã(29)     | 3.000           | 1,5        | 5,0        | 863,0        |

Fonte: Saneago S.A., 2002.

Tabela 2 – Goiânia - dados das bombas da estação Senac

| Bombas | Vazão(l/s) | Carga(m) |
|--------|------------|----------|
| 26     | 895        | 85       |
| 27     | 895        | 85       |
| 28     | 895        | 85       |

Fonte: Saneago S.A., 2002.

Baseado em estudos desenvolvidos por Tanyimboh et al. (2001), Gargano e Pianese (2000) e Righetto(2002), este trabalho propõe a avaliação dos benefícios hidráulicos através da adoção de



coeficientes ou índices de performance. Para consideração das pressões nos nós de demanda utilizou-se o coeficiente denominado de benefício de atendimento às pressões mínimas nodais( $\psi_{bp}$ ), de acordo com o seguinte equacionamento:

$$\psi_{bp(i,t)} = \left( \frac{P_{at(i,t)} - P_{mín}}{P_{req(t)} - P_{mín}} \right)^{1/2} \quad \text{se} \quad P_{mín} \leq P_{at(i,t)} \leq P_{req(t)} \quad (3)$$

$$\psi_{bp(i,t)} = 0 \quad \text{se} \quad P_{at(i,t)} < P_{mín} \quad \text{ou} \quad P_{at(i,t)} > P_{req(t)} \quad (4)$$

onde:

$P_{at(i,t)}$  = pressão no nó de demanda i no tempo t(fornecida pelo simulador hidráulico);

$P_{mín}$  = pressão mínima adotada no caso 15m;e

$P_{req(t)}$  = pressão máxima admitida no tempo t.

O benefício hidráulico relativo ao atendimento adequado da pressão nodal ( $BH_{PN}$ ) será dado pela seguinte equação:

$$BH_{PN} = \sum_{t=1}^{24} \sum_{i=1}^{nn} \psi_{bp(i,t)} \quad (5)$$

Para avaliação dos benefícios relativos aos níveis de água nos reservatórios, considerou-se o coeficiente denominado de benefício de adequação aos níveis de água nos reservatórios( $\psi_{bn}$ ), dado pela seguinte equação:

$$\psi_{bn(j,t)} = \left( \frac{N_{at(j,t)} - N_{mín(j)}}{N_{req(j,t)} - N_{mín(j)}} \right)^{1/2} \quad \text{se} \quad N_{mín} \leq N_{at(j,t)} \leq N_{req(j,t)} \quad (6)$$

$$\psi_{bn(j,t)} = 0 \quad \text{se} \quad N_{at(j,t)} < N_{mín} \quad \text{ou} \quad N_{at(j,t)} > N_{req(j,t)} \quad (7)$$

onde:

$N_{at(j,t)}$  = nível de água no reservatório j no tempo t(fornecido pelo simulador);

$N_{mín(j)}$  = nível mínimo de água no reservatório j ;e

$N_{req(j,t)}$  = nível de água requerido no tempo t.

O benefício hidráulico relativo à adequação dos níveis de água nos reservatórios ( $BH_{NR}$ ) é dado pela seguinte equação:

$$BH_{NR} = \sum_{t=1}^{24} \sum_{j=1}^{nr} \psi_{bn(j,t)} \quad (8)$$

Com relação à garantia de atendimento das demandas considerou-se o benefício ( $BH_{AD}$ ) de acordo com o seguinte equacionamento:

$$BH_{AD} = \sum_{t=1}^{24} \sum_{i=1}^{nn} \left( \frac{P_{at(i,t)} - P_{\min}}{P_{req(t)} - P_{\min}} \right)^{1/2} * \left( \frac{Q_{dem(i,t)}}{\sum_{i=1}^{nn} Q_{dem(i,t)}} \right) \quad (9)$$

onde :

$Q_{dem(i,t)}$  = demanda horária nó de demanda i no tempo t; e

$\sum_{i=1}^{nn} Q_{dem(i,t)}$  = demanda horária total no tempo t.

A função objetivo ( $FO_2$ ), relativa aos benefícios hidráulicos, seria expressa como o somatório desses três benefícios, ou seja:

$$FO_2 = BH_{PN} + BH_{NR} + BH_{AD} \quad (10)$$

O problema formulado anteriormente é de otimização multiobjetivo. As funções objetivo custo e benefício são conflitantes entre si. Assim, não há uma solução ótima simples que possa satisfazer o ótimo global de ambos os objetivos, mas um conjunto de soluções ótimas designadas por soluções não dominadas ou não inferiores que não podem ser melhoradas sem sacrificar o restante dos objetivos. Algoritmos genéticos têm se mostrado efetivos na identificação destas soluções. O conjunto destas é chamado frente ótima Pareto.

### Representação das soluções

Para o caso específico do sistema em estudo, cada vetor representativo de uma possível solução  $S_1$  para a estratégia operacional do sistema tem a seguinte característica:

$$S_1 = \left[ \underbrace{B_{(0,1)}, B_{(0,2)}, B_{(0,3)}, V_{(0,1)}, V_{(0,2)}, V_{(0,3)}, V_{(0,4)}}_{\text{Hora 0}}, \dots, \underbrace{B_{(23,1)}, B_{(23,2)}, B_{(23,3)}, V_{(23,1)}, V_{(23,2)}, V_{(23,3)}, V_{(23,4)}}_{\text{Hora 23}} \right] \quad (11)$$

Onde :

$S_1$  = vetor solução ;

B e V = variáveis de decisão relativas à operação de bombas e válvulas, respectivamente;

$B_{(0,1)}$  = funcionamento da bomba número 1 na hora zero (0/1 – desligada/ligada) ;

$V_{(0,1)}$  = funcionamento da válvula número 1 na hora zero (0/1 – fechada/aberta) .

### Arquivo de dados para o SEE5

Como visto anteriormente, existe um grupo de métodos conhecidos como sistemas *expert* ou sistemas baseados no conhecimento (*knowledge-based systems*). Estes sistemas criam modelos de classificação através da descoberta e análise de padrões que podem ser encontrados nos registros de dados.

O emprego do modelo otimizador através dos algoritmos genéticos multiobjetivo produz uma frente Pareto com as soluções operacionais de melhor compatibilização entre os objetivos pré-definidos. A utilização do aprendizado de máquina sobre este conjunto de soluções, aqui chamadas de exemplos ou casos, visa extrair as melhores regras (estratégias) operacionais do conjunto apresentado após a aplicação do método *SPEA*.

Cada exemplo é descrito por um vetor de valores de características, ou atributos, e o rótulo da classe associada. Em geral, o objetivo do algoritmo de indução é construir um classificador que possa determinar corretamente a classe de novos exemplos ainda não rotulados. O processo de classificação neste trabalho visa a identificação das melhores estratégias operacionais de um conjunto pré-definido através da frente Pareto. Para isto, é necessário verificar no arquivo de saída do modelo de otimização, os registros contendo os atributos que caracterizam a estratégia operacional. Cada registro deve ser rotulado por um especialista da área operacional. Este especialista tem várias opções em termos da melhor condição operacional para o sistema a ser operado, podendo ser destacadas a flexibilidade e a confiabilidade do sistema. Para este trabalho, considerou-se a confiabilidade mecânica representada pela condição operacional de bombas e válvulas como rótulo para definição das melhores regras. Consideraram-se os seguintes atributos: condição operacional boa e condição operacional ótima.

A escolha dos parâmetros a serem utilizados no modelo visou o atendimento dos objetivos pré-definidos, quais sejam, menor custo de energia elétrica e maior benefício hidráulico. Porém, a classe principal **condição operacional**, definida para o aprendizado de máquina, priorizou, através do conhecimento e de informações do operador (especialista), a maximização da confiabilidade mecânica em termos de uma operação mais eficaz das bombas na estação de bombeamento e das válvulas de controle na entrada dos reservatórios. Desta forma, a classe **ótima** representa as soluções, fornecidas pela frente Pareto, que fornecem os melhores valores em termos do atendimento das demandas, pressões nos nós e os melhores níveis de água no reservatório, com os menores custos de operação em termos de energia elétrica e, também, a melhor estratégia operacional em termos de preservação e manutenção mecânica das bombas e válvulas.

Righetto (2002) propôs cinco parâmetros distintos para avaliar a performance de uma determinada regra operacional num período de 24 horas. Um desses parâmetros era o índice de condição operacional de bombas. Neste trabalho seguiu-se a mesma metodologia. Porém, o julgamento do especialista para definição da classe condição operacional, foi implementado automaticamente através da criação de índices operacionais de bombas e válvulas.

O índice foi definido com base no número de bombas e válvulas em funcionamento num período de 24 horas de operação do sistema. Foi chamado de índice de condição operacional ( $K_{pv}$ ) e definido pela seguinte expressão:

$$K_{pv}(i) = \sum_{k=1}^{nb} \sum_{t=1}^{24} \lambda_b(k, t) + \sum_{j=1}^{nv} \sum_{t=1}^{24} \lambda_v(j, t) \quad (12)$$

Onde:

$t$  = intervalo de tempo considerado;

$nb$  = número de bombas na estação elevatória;

$nv$  = número de válvulas no sistema;

$i$  = solução(estratégia) operacional da frente Pareto; e

$\lambda_b$  e  $\lambda_v$  = parâmetros de funcionamento das bombas e válvulas, respectivamente;

Estes parâmetros têm as seguintes definições:

$\lambda_b$  ou  $\lambda_v = 1$  se  $\lambda_b(k, t) = \lambda_b(k, t-1)$  ou  $\lambda_v(j, t) = \lambda_v(j, t-1)$

$\lambda_b$  ou  $\lambda_v = 0$  em qualquer outra situação

A classe condição operacional será ótima (o) se  $K_{bv(i)} > \frac{1}{ns} * C$  e boa (b) se

$K_{bv(i)} \leq \frac{1}{ns} * C$ . Onde  $C = \sum_{i=1}^{ns} K_{bv(i)}$  e  $ns$  é o número total de soluções na frente Pareto.

Este trabalho utiliza o classificador denominado conjuntos de regras (*rulesets*) para extração de regras operacionais a partir de um conjunto de exemplos (casos) fornecidos pelo modelo de otimização através das soluções ótimas pertencentes à frente Pareto. Para tanto é utilizado o indutor *SEE5*, que é a mais recente versão do indutor *C4.5* descrito por Quinlan (1993).

Para utilização do indutor *SEE5* são necessários dois arquivos de entrada distintos. O **arquivo.data**, contendo os exemplos (definidos através dos valores de seus atributos) a serem analisados e o **arquivo.names** com a definição dos atributos e classes.

### Modelo computacional proposto

Um modelo computacional foi implementado compreendendo basicamente três módulos: primeiro o de avaliação hidráulica, utilizando códigos da biblioteca *Toolkit-Epanet2* (ROSSMAN, 2002), um de implementação de AGs multiobjetivo, utilizando códigos fornecidos pela biblioteca *Multiple Objective Metaheuristics Library in C++- MOMHLib*, e um último de indução para criação do classificador de conjunto de regras a partir do indutor *SEE5*.

As principais características do sistema macro-adutor de Goiânia, utilizado como aplicação neste trabalho, foram apresentadas através da figura 2 e das tabelas 1 e 2.

Os parâmetros finais utilizados nos módulos de simulação hidráulica e otimização foram:

- níveis iniciais máximos nos reservatórios de distribuição (zero hora);
- aproveitamento total dos volumes de reservação até o nível de água 1,5m;
- recombinação uniforme ;
- mutação não-uniforme;
- tamanho da população – 300;
- número máximo de gerações – 5000;
- probabilidade de recombinação – 0,90;
- probabilidade de mutação – 0,006.

## RESULTADOS

Utilizando os parâmetros descritos anteriormente, foi gerada a frente Pareto apresentada na figura 3. As soluções dessa frente representam as melhores estratégias operacionais para o sistema de macro distribuição de água de Goiânia.

Com o objetivo de obter um conjunto menor de soluções para a operação do sistema, a partir das estratégias da frente Pareto, foi implementado o classificador *SEE5* e obtiveram-se duas estratégias finais, denominadas no gráfico de melhores soluções - AM. Estas estratégias foram produzidas após a aplicação do algoritmo de aprendizado de máquina no arquivo de saída do otimizador e utilização do índice de condição operacional ( $K_{pv}$ ). Como descrito anteriormente, para aplicação do aprendizado de máquina, foi definida uma classe denominada condição operacional. Esta classe pode assumir duas condições: **boa e ótima**.

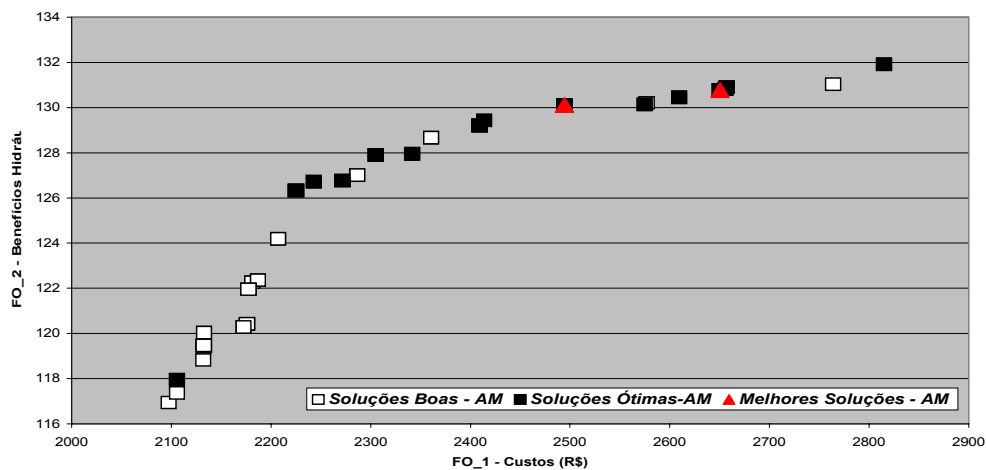


Figura 3 – Goiânia – frente Pareto – soluções AM

As figuras 4 e 5 apresentam, respectivamente, os níveis de água nos quatro reservatórios e o funcionamento das bombas do sistema de macro adução de Goiânia, para uma das melhores soluções originadas a partir da aplicação do algoritmo de aprendizado de máquina.

Em termos de avaliação global do sistema, ou seja, interligação entre os quatro reservatórios para um possível aproveitamento de “sobras” de reservação, percebe-se que, pela localização das unidades, que esta estratégia não é representativa para a melhoria da qualidade da operação. O reservatório que apresenta dificuldades no atendimento diário da demanda é o itapoã. Esta unidade está localizada numa cota, 863,00 m, superior aos reservatórios cujos volumes admitem “sobras”. Portanto, este aproveitamento somente seria possível através da implantação de mais uma estação elevatória intermediária.

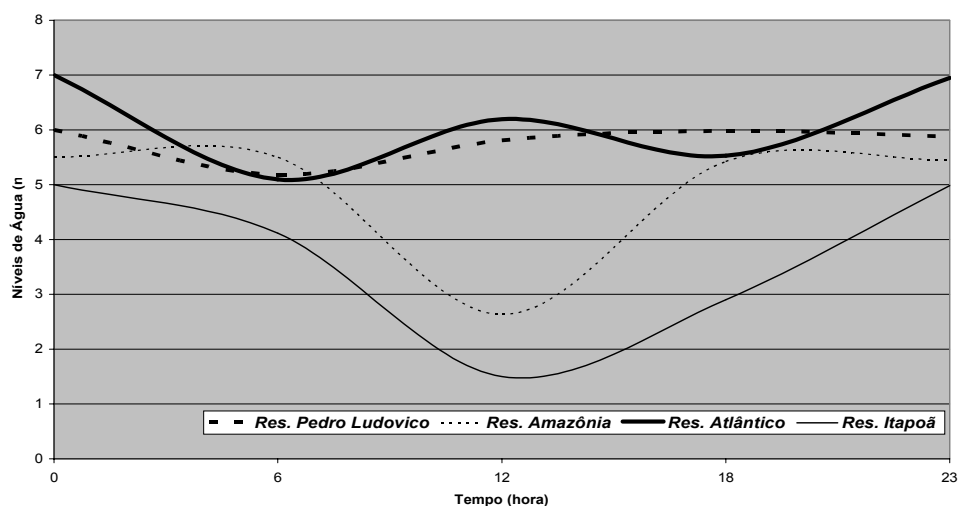


Figura 4 – Goiânia – níveis de água nos reservatórios – solução AM

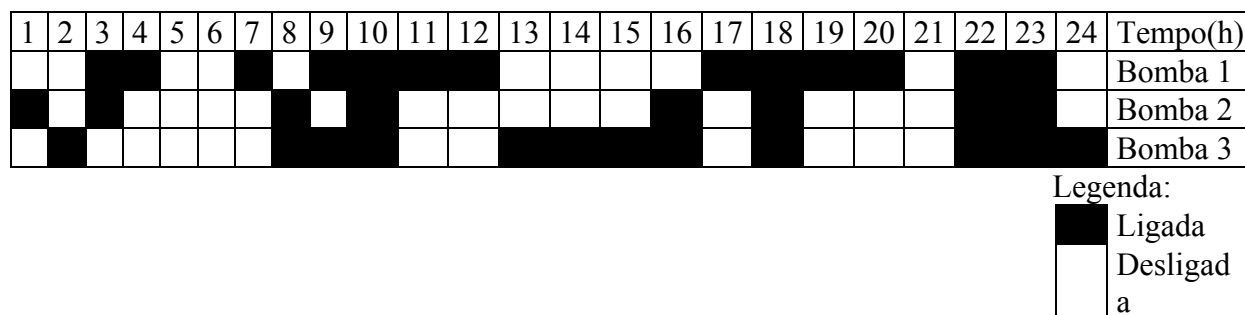


Figura 5 – Goiânia – funcionamento das bombas – solução AM

Ao analisarmos os gráficos da figura 4 percebe-se que estes aspectos relatados anteriormente são evidenciados pelos resultados. Os reservatórios atlântico e pedro ludovico apresentam níveis de água próximos aos máximos durante quase todo o período de análise. O amazonas apresenta uma variação de nível adequada em termos de operação normal do sistema, enquanto no itapoã o nível de água se mantém próximo do mínimo durante grande período do dia, porém garantindo o atendimento das demandas na área de abastecimento.

Ressalta-se que os resultados apresentados, em termos de níveis de água nos reservatórios, são extremamente factíveis, demonstrando a validade do modelo desenvolvido.

Em termos de funcionamento das bombas, os resultados apresentados através da figura 5 mostram que é possível manter um número razoável de bombas desligadas ao longo do dia, principalmente no horário de pico (18 h às 21 h) onde o valor da tarifa é duas vezes maior que fora deste horário. Neste período foi possível reduzir em aproximadamente 60% o funcionamento das unidades. Isto demonstra a viabilidade do modelo em termos de minimização dos custos relativos ao consumo de energia elétrica.

## CONCLUSÕES

O maior problema dos sistemas de abastecimento público de água é a natureza multiobjetivo da operação aliada às incertezas na previsão de demandas e custos. Em geral considera-se, como principal objetivo, a diminuição dos custos de produção e operação, porém com a garantia de um serviço adequado priorizando a confiabilidade em termos de atendimento qualitativo e quantitativo do consumidor.

A maioria dos modelos desenvolvidos para sistemas de distribuição de água, consideram apenas um objetivo, o custo de operação e manutenção. Isto acontece pela complexidade computacional envolvida quando são considerados outros objetivos como confiabilidade hidráulica e mecânica, flexibilidade e qualidade da água.

Algoritmos multiobjetivo (MOGAs) têm sido eficientes na otimização da operação de sistemas de água. A desvantagem dos MOGAs é que, enquanto eles são eficientes em encontrar a região da solução ótima, são ineficientes na identificação do ponto ótimo dentro dessa região.

Este trabalho apresentou o desenvolvimento de um modelo baseado em algoritmos genéticos multiobjetivo e o aprendizado de máquina. Objetivou-se otimizar a operação e extrair estratégias operacionais factíveis, que garantissem o adequado funcionamento do sistema de macro distribuição de água potável da cidade de Goiânia. O problema foi formulado considerando dois objetivos distintos e conflitantes, a minimização dos custos de energia elétrica na estação elevatória e a maximização dos benefícios hidráulicos.

Um conjunto de soluções denominadas Pareto ótimas, no espaço de busca dos objetivos (custo e benefício hidráulico), foi gerado a partir da aplicação do algoritmo *multobjetivo strength Pareto evolutionary algorithm (SPEA)*. Este conjunto representa as melhores estratégias operacionais para o sistema de distribuição de água de Goiânia. Porém, a frente Pareto fornece um

grande número de soluções, o que dificulta a escolha da mais adequada. É importante fornecer um conjunto menor de regras operacionais, evitando que diante das dificuldades de escolha da mais factível o operador adote uma solução baseada apenas em sua experiência. Da mesma forma, é interessante fornecer mais de uma estratégia para que fatores externos, não considerados no modelo, possam eventualmente ser incorporados, aumentando a flexibilidade de ação do operador sem comprometer a qualidade da operação.

O algoritmo de aprendizado de máquina foi aplicado no conjunto Pareto com o objetivo de encontrar um número menor de soluções porém, com a garantia do fornecimento das estratégias com o melhor desempenho em termos de confiabilidade operacional de bombas e válvulas.

Os resultados mostraram a eficácia do modelo para os objetivos propostos. O algoritmo genético multiobjetivo *SPEA* forneceu uma frente Pareto com as melhores soluções em termos de custo e benefício hidráulico do sistema, assim como o aprendizado de máquina apresentou duas estratégias operacionais com ótimas performances operacionais de bombas e válvulas.

## REFERÊNCIAS BIBLIOGRÁFICAS

- BAO, Y.; MAYS, L.W. (1990). *Model for Water Distribution System Reliability*. Journal of Hydraulic Engineering, v.116, n.9, p.1119-1137.
- CORNE, D.; KNOWLES, J.; OATES, M. (2000). *The Pareto Envelope-Based Selection Algorithm for Multi-Objective Optimization*. In Proceedings of The Sixth International Conference on Parallel Problem Solving from Nature VI (PPSN-VI), p. 839-848.
- DEB, K. et al. (2002). *A Fast and Elitist Multiobjective Genetic Algorithm: NSGA II*. IEEE Transactions on Evolutionary Computation. 6(2), April, p. 182-197.
- FONSECA, C.M.; FLEMING, P.J. (1993). *Genetic Algorithms for Multiobjective Optimization: Formulation, Discussion and Generalization*. In: Fifth International Conference, San Mateo, 8p.
- GARGANO, R.; PIANESE, D. (2000). *Reliability as Tool for Hydraulic Network Planning*. Journal of Hydraulic Engineering, v.126, n.5, May, p.354-364.
- HORN, J.; NAEPLOITIS, N.; GOLDBERG, D. (1994). *A Niche Pareto Genetic Algorithm for Multi-Objective Optimization*. In: Proceedings of The First IEEE Conference on Evolutionary Computation, p. 82-87.
- KNOWLES, J.D.; CORNE, D.W. (2000). *Approximating The Non-Dominated Front Using The Pareto Archived Evolution strategy*. Evolutionary Computation Journal, 8(2), p. 149-172.
- QUINLAN, J.R. (1993). *C4.5: Programs for Machine Learning*. Morgan Kaufmann Publishers. San Francisco, CA.
- RIGHETTO, A.M. (2002). *Operação Ótima de Sistema Urbano de Distribuição de Água*. In: Seminário-Planejamento, Projeto e Operação de Redes de Abastecimento de Água. O Estado da Arte e Questões Avançadas. João Pessoa. CD-Rom. 16p.

- ROSSMAN, L.A. (2002). *EPANET2-Users Manual*. U.S. Environmental Protection Agency. Ohio,USA.
- RUDOLPH, G. (1998). *On a Multi-Objective Evolutionary Algorithm and its Convergence to The Pareto Set*. In:Proceedings of The 5<sup>th</sup> IEEE Conference on Evolutionary Computation, p. 511-516.
- SRINIVAS, N; DEB, K (1994). *Multi-Objective Function Optimization Using Non-Dominated Sorting Genetic Algorithms*. Evolutionary Computation Journal, 2(3), p. 221-248.
- TANYIMBOH, T.T.; TABESH, M.; BURROWS, R. (2001). *Appraisal of Source Head Methods for Calculating Reliability of Water Distribution Networks*. Journal of Water Resources Planning and Management, v.127, n.4, July-August, p. 206-213.
- WEISS, S.M.; KULIKOWSKI, C.A. (1991). *Computer Systems that Learn*. San Mateo, California, USA. Morgan Kaufmann.
- XU, C.; GOULTER, I.C. (1999). *Reliability-Based Optimal Design of Water Distribution Networks*. Journal of Water Resources Planning and Management, v.125, n.6, November, p. 352-362.
- ZAHED FILHO, K. (1990). *Previsão de Demanda de Consumo em Tempo Real no Desenvolvimento Operacional de Sistemas de Distribuição de Água*. São Paulo. 135p. Tese(Doutorado). Escola Politécnica da Universidade de São Paulo.
- ZITZLER, E.; THIELE, L. (1998). *An Evolutionary Algorithm for Multi-Objective Optimization: The Strength Pareto Approach*. Technical Report 43, Zurich, Switzerland: Computer Engineering and Networks Laboratory(TIK), Swiss Federal Institute of Technology(ETH).
- ZITLER, E.; LAUMANN, M.; THIELE, L. (2001). *SPEA2: Improving The Strength Pareto Evolutionary Algorithm*. Technical Report 103. Computer Engineering and Networks Laboratory (TIK). Swiss Federal Institute of Technology (ETH). Zurich, Switzerland, May, 19p.