Extração de Estratégias Operacionais Ótimas de Sistemas de Distribuição de Água Utilizando Algoritmos Genéticos Multiobjetivo e Aprendizado de Máquina – Aplicação ao Sistema de Macro-distribuição de Água de Goiânia

Ivaltemir Barros Carrijo
Saneamento de Goiás S.A. – SANEAGO – tinil@saneago.com.br

Luisa Fernanda Ribeiro Reis
Departamento de Hidráulica e Saneamento - EESC-USP – fernanda@sc.usp.br

Recebido: 08/03/05 – revisado: 12/04/05 – aceito: 01/09/05

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 aprendizado de máquina para extração de regras operacionais para o sistema.O modelo foi aplicado em uma parte do macro sistema distribuidor de água da cidade de Goiânia e os resultados demonstraram que podem ser produzidas estratégias operacionais satisfatórias para o sistema em substituição ao julgamento pessoal do operador..

Palavras-chave – *redes de distribuição, otimização multiobjetivo, algoritmos genéticos, aprendizado de máquina*

INTRODUÇÃO

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

Em geral a equipe de planejamento da operação define as regras (fixas ou variáveis) de controle dos sistemas, baseada em informações e experiências passadas e no conhecimento do estado atual do sistema. As regras são transmitidas ao controle do sistema que, por sua vez, retorna os resultados das operações efetuadas ao

planejamento, para avaliação do desempenho e para as alterações necessárias.

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 vantagens 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 atingir a operação ótima de sistemas de distribuição de água, essencialmente o sistema macro-adutor (esqueleto), composto por adutoras, elevatórias, reservatórios e estruturas de controle com 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.

Com o conhecimento do sistema de distribuição de água, através de um cadastro técnico e comercial georreferenciado, pretende-se simular hidraulicamente o sistema utilizando o simulador hidráulico EPANET 2 (Rossman, 2001), otimizá-lo operacionalmente através de AGs, fornecer regras operacionais através do aprendizado de máquina e, através de uma interface com o SCADA(Supervisory Control and Data Acquisition), fornecer ao operador do sistema a 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, desde a captação da água no manancial abastecedor até a distribuição ao consumidor final.

MODELO PROPOSTO

Quando a demanda total de um sistema de distribuição de água para abastecimento é menor que a capacidade total de distribuição há uma potencial redução nos custos de operação de bombas e válvulas. A busca por uma operação ótima do sistema, como sendo um problema de controle ótimo, que visa encontrar as decisões para operação de válvulas e bombas, dentro de um intervalo de

tempo pré-definido, deve possibilitar que as demandas sejam atendidas, as condições de pressões satisfeitas e o custo total de operação minimizado (Zessler e Shamir, 1989).

Na resolução de problemas de engenharia como esse, procura-se, inicialmente, representar o sistema real de maneira mais fidedigna possível, para que possa reproduzir o comportamento do mesmo. Os recursos computacionais atualmente disponíveis mostram que as simplificações, que há algumas décadas eram necessárias para resolução de problemas de engenharia em geral, vem se tornando cada vez mais dispensáveis. Entretanto, para a resolução do modelo matemático entendido como sendo o conjunto completo e consistente de equações que representam o sistema físico real, é necessária a inserção de parâmetros complementares obtidos, em geral, em campo, visando a obtenção de resultados numéricos significativos (Koelle, 1992).

Técnicas de otimização vêm sendo tradicionalmente usadas no dimensionamento de unidades de um sistema de distribuição de água, objetivando menores custos. 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.

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

A confiabilidade pode ser descrita através da minimização de dois tipos de falhas: mecânicas e hidráulicas. Falhas mecânicas são do tipo: rompimento de tubulações, falhas no bombeamento, falhas no controle de válvulas e perda de potência, dentre outras. São consideradas como falhas hidráulicas, as ocorrências de vazões e pressões inadequadas, para um ou mais pontos de demanda do sistema.

Usualmente, a política global empregada na operação de um sistema de abastecimento de água não é definida de maneira clara. A principal razão é a dificuldade em obter as informações, devido ao fato de que estas, geralmente, são baseadas apenas no conhecimento e experiência do operador. Outras razões são a variedade de mecanismos de controle e as frequentes mudanças na topologia do sistema de abastecimento de água (León et al., 2000).

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.

Como o sistema de macro distribuição de água é usualmente composto de estações de bombeamento e reservatórios, o procedimento básico de operação deve considerar o controle das vazões e cargas ou o número de bombas em operação e o nível de água disponível nos reservatórios, de forma a garantir o atendimento independente da variação diária das demandas. De acordo com Ko et al. (1997), o objetivo de uma efetiva operação de sistemas de abastecimento de água é minimizar os custos operacionais, mantendo a operação das bombas de forma estável, além de garantir uma alta confiabilidade do sistema, através da manutenção dos níveis de água nos reservatórios próximos dos máximos. A redução nos custos de energia elétrica pode ser obtida com uma maior utilização dos conjuntos elevatórios no período fora de pico de consumo de água, mas isto vai depender de uma complicada correlação entre demanda horária e tarifa de energia elétrica.

A otimização proposta neste trabalho leva em consideração dois objetivos: a minimização de custos operacionais decorrentes do consumo de energia elétrica nas estações de bombeamento 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).

Com o objetivo de reduzir os custos de operação relativos à utilização de estações de bombeamento, além de incrementar os benefícios hidráulicos do sistema, este trabalho propõe o desenvolvi-

mento de um modelo a ser utilizado para obtenção de regras de operação de unidades de um sistema real, com operação não automática, que satisfaça os objetivos pré-fixados. O modelo proposto é composto de três módulos distintos: um módulo de simulação hidráulica que utiliza os códigos do EPANET2, um módulo otimizador que utiliza algoritmos genéticos multiobjetivo e um módulo indutor de regras através de aprendizado de máquina.

Simulação Hidráulica

Os modelos de simulação de redes de distribuição de água consistem basicamente do equacionamento das vazões nos trechos de uma rede e das cargas hidráulicas nos nós entre trechos consecutivos. Podem ser aplicados em estudos de dimensionamento de redes, isoladamente, onde a solução é obtida por análises sucessivas, com alterações impostas pelos usuários, ou em combinação com modelos de otimização, onde a solução é obtida automaticamente.

A simulação hidráulica avalia as respostas do sistema em termos das suas variáveis de estado: pressão, vazão e níveis nos reservatórios, face às diversas conjunturas operacionais. É, portanto, uma ferramenta essencial à avaliação dos objetivos estabelecidos. Neste trabalho, o simulador utilizado se baseia no código original do EPANET2 (Rossman, 2001), acoplado ao C++ (Borland Builder C++ 5.0).

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 de 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 $f_1, f_2, \ldots . f_n$, é compor uma nova função objetivo F, obtida da soma ponderada das funções objetivo, ou :

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

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:

$$\left(\forall_i : f_i(x_1) \geq f_i(x_2)\right) e \left(\exists_j : f_j(x_1) > f_j(x_2)\right) \qquad (2)$$

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:

$$P = \left\{ x_i \in S \,\middle|\, \nexists x_j \in S : x_i \text{ é dominado por } x_j \right\} \qquad (3)$$

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, assim como apresentado em Carrijo (2004), 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 Jaszkiewicz (1999) 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). Maiores detalhes sobre o classificador SEE5 podem ser vistos em Carrijo (2004).

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 1), com suas características principais (tabelas 1 e 2).

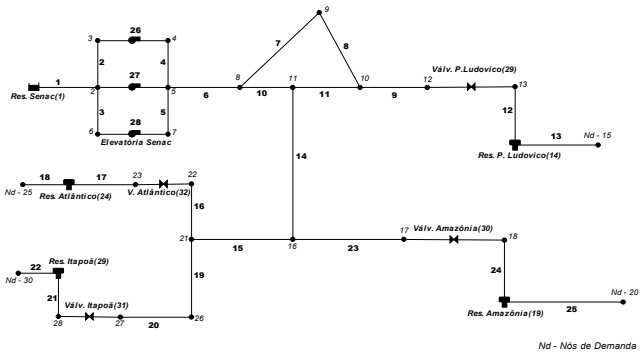

Figura 1 – Esquema de macro distribuição de Goiânia

Tabela 1 – Goiânia – Dados dos reservatórios

Reservatórios	Nív.Mín-Máx(m)	Elev.(m)
P.Ludovico(14)	1,5-6,0	858,0
Amazônia(19)	1,5-5,5	861,5
Atlântico(24)	1,5-7,0	836,5
Itapoã(29)	1,5-5,0	863,0

Tabela 2 – Dados das bombas da estação Senac

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

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.

Este trabalho propõe avaliar dois objetivos básicos. O objetivo econômico e o objetivo de benefícios hidráulicos 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.
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:

$$FO_1 = \sum_{t=1}^{24} \sum_{k=1}^{nb} Cu(t) \frac{Q(k,t) * H(k,t) * \gamma}{\eta(k,t)} + D * Tarifa \qquad (4)$$

onde:
k=número de bombas na estação elevatória;
t=hora do dia;
nb=número total de bombas na elevatória;
Cu(t)=custo unitário da terifa (R$kWh);
Q(i,t)=vazão bombeada (m^3/s);
H(i,t)=carga hidráulica (m);
η(i,t)=rendimento de conjunto (%);
D=fator de demanda máxima (kW);
Tarifa=tarifa correspondente ao fator de demanda máxima (R$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.

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(ψ_{bp}), de acordo com o seguinte equacionamento:

$$\psi_{bp(it)} = \left(\frac{P_{at(it)} - P_{mín}}{P_{req(t)} - P_{mín}} \right)^{1/2} \quad \text{se } P_{mín} \le P_{at(it)} \le P_{req(t)} \quad (5)$$

$$\psi_{bp(it)} = 0 \quad \text{se} \quad P_{at(it)} < P_{mín} \quad \text{ou} \quad P_{at(it)} > P_{req(t)} \quad (6)$$

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:

$$BH_{PN} = \sum_{t=1}^{24} \sum_{i=1}^{nn} \psi_{bp(i,t)} \quad (7)$$

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(ψ_{bn}), dado pela seguinte equação:

$$\psi_{bn(jt)} = \left(\frac{N_{at(jt)} - N_{mín(j)}}{N_{req(jt)} - N_{mín(j)}} \right)^{1/2} \quad \text{se} \quad N_{mín} \le N_{at(jt)} \le N_{req(jt)} \quad (8)$$

$$\psi_{bn(jt)} = 0 \quad \text{se} \quad N_{at(jt)} < N_{mín} \text{ ou } N_{at(jt)} > N_{req(jt)} \quad (9)$$

onde:
$N_{at(j,t)}$ = nível de água no reserv. J no tempo t (fornecido pelo simulador);
$N_{min\ (j)}$ = nível mínimo de água no reservatório j; e
$N_{req(j,t)}$ =nível de água requerido no tempo o t.

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

$$BH_{NR} = \sum_{t=1}^{24} \sum_{j=1}^{nr} \psi_{bn(j,t)} \quad (10)$$

Com relação à garantia de atendimento das demandas considerou-se o benefício (BH_{AD}) de acordo com a equação 11.

$$BH_{AD} = \sum_{t=1}^{24} \sum_{i=1}^{nn} \left(\frac{P_{at(i,t)} - P_{mín}}{P_{req(t)} - P_{mín}} \right)^{1/2} * \left(\frac{Q_{dem(i,t)}}{\sum_{i=1}^{nn} Q_{dem(i,t)}} \right) \quad (11)$$

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:

$$FO_2 = BH_{PN} + BH_{NR} + BH_{AD} \quad (12)$$

O problema formulado anteriormente é de otimização multiobjetivo. As funções objetivo custo (eq. 4) e benefício (eq. 12) 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.

Arquivo de dados para o sistema SEE5

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 sistema See5, 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.

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 as seguintes classes: 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 relativa à 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 foi adotada a mesma sistemática de definição de índices operacionais de Righetto (2002). Porém, o julgamento do especialista para definição da classe condição operacional, foi implementado através de um índice operacional que avalia a estratégia de funcionamento 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:

$$K_{BV}(i) = \sum_{k=1}^{nb}\sum_{t=1}^{24}\lambda_b(k,t) + \sum_{j=1}^{nv}\sum_{t=1}^{24}\lambda_v(j,t) \qquad (13)$$

Onde:

k, j = número de bombas e válvulas , respectivamente;
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
λ_b e λ_v = parâmetros de funcionamento das bombas e válvulas, respectivamente;

Estes parâmetros têm as seguintes definições:

$$\lambda_b^{(k,t)} \text{ ou } \lambda_v^{(j,t)} = 1 \text{ se } \lambda_b(k,t) = \lambda_b(k,t-1) \text{ ou } \lambda_v(j,t) = \lambda_v(j,t-1) \qquad (14)$$

$\lambda_b^{(k,t)}$ ou $\lambda_v^{(j,t)}$ = 0 em qualquer outra situação.

A classe condição operacional será ótima (o) se $K_{bv(i)} > \frac{1}{ns} * C$ e boa (b) se $K_{bv(i)} <= \frac{1}{ns} * C$. Onde $C = \sum_{i=1}^{ns} K_{bv(i)}$ e ns é o número total de soluções na frente Pareto.

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

Modelo computacional proposto

Um modelo computacional foi implementado compreendendo basicamente três módulos: primeiro o de avaliação hidráulica, utilizando códigos da biblioteca Toolkit-Epanet2 (Rossman, 2001), 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. A figura 2 apresenta um fluxograma do modelo proposto.

As principais características do sistema macro-adutor de Goiânia, utilizado como aplicação neste trabalho, foram apresentadas através da figura 1 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 aprendizado de máquina(AM) 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**. Para facilidade de compreensão, as soluções fornecidas pelo AM serão chamadas de AM1 (2494,3; 130,1) e AM2 (2650,4; 130,8).

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 as melhores soluções originadas a partir da aplicação do 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, 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.

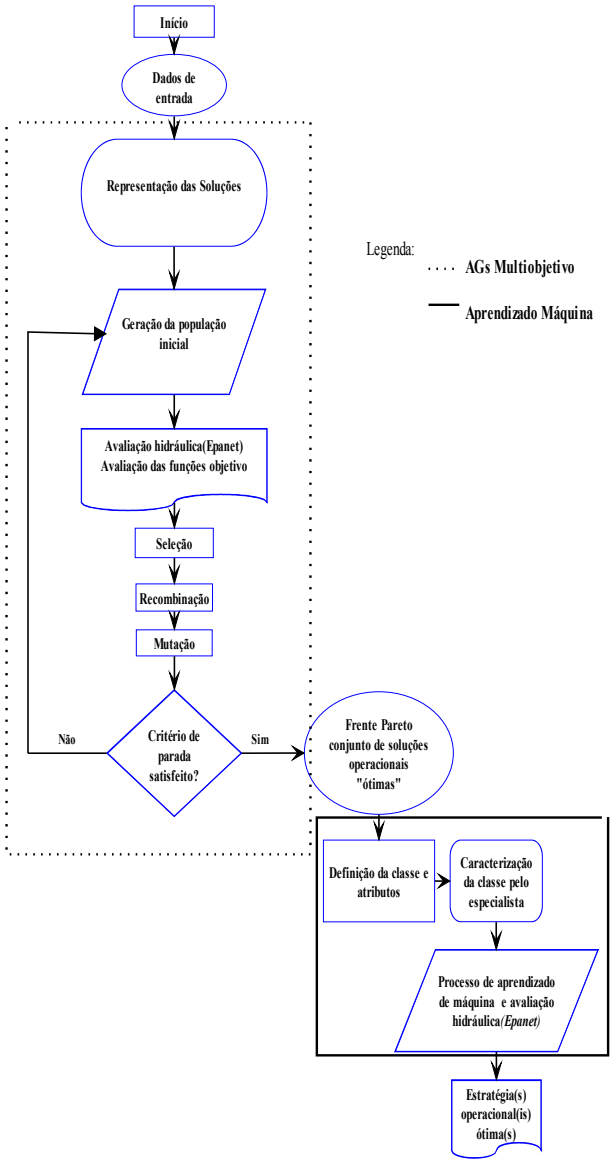

Figura 2 – Fluxogroma do modelo proposto

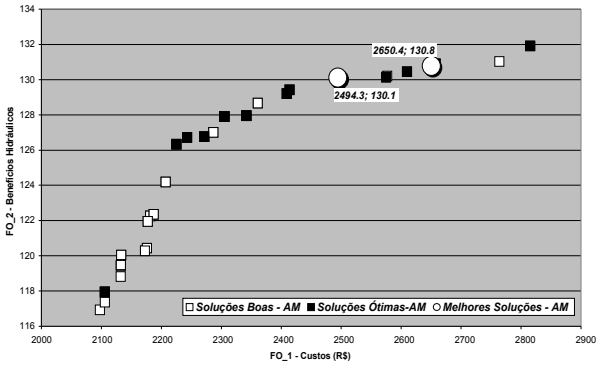

Figura 3 – Frente Pareto – soluções ótimas e boas – melhores soluções AM

Analisando-se 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 Amazônia apresenta uma variação de nível adequada(acima de 1,5m) em termos de uma operação eficiente sistema, enquanto no Itapoã o nível de água se mantém próximo do mínimo durante o período diurno de pico, porém garantindo o atendimento das demandas na área de abastecimento.

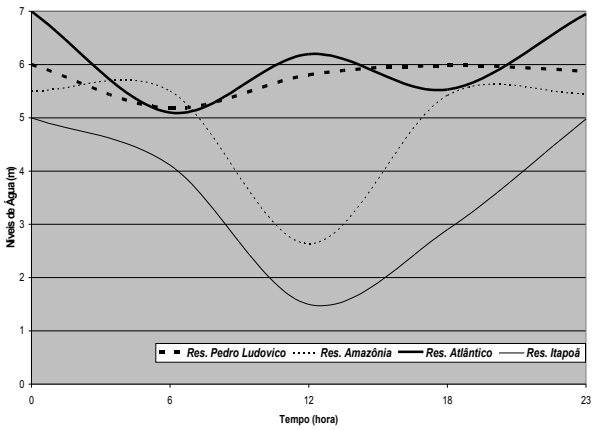

Figura 4 – Níveis de água nos reservatórios – Solução AM1

Comparando as curvas de variação do nível de água nos reservatórios de Goiânia apresentadas através da figura 4 pode-se perceber que, mesmo tendo trajetórias diferenciadas, com exceção do reservatório Itapoã, os níveis mínimos estão acima do mínimo admitido para cada reservatório (1,5m), o que garante o suprimento dos nós de demanda com pressões adequadas.

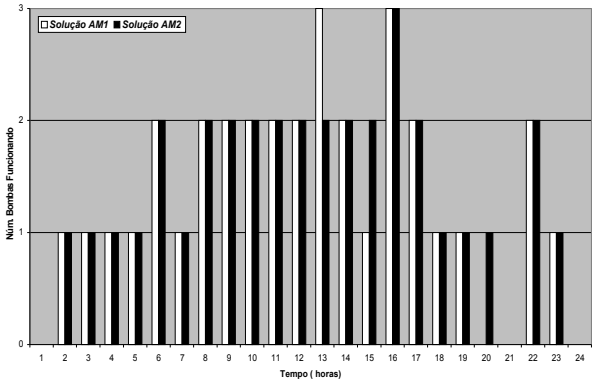

Figura 5 – Número de bombas em funcionamento – soluções AM1 e AM2

Em termos do funcionamento das bombas, os resultados apresentados através da figura 5 mostraram que é possível manter um número razoável de bombas desligadas ao longo do dia, principalmente no horário de pico (18 h às 21 h), em que o valor da tarifa é duas vezes maior que o valor fora deste horário. Neste período, foi possível reduzir em aproximadamente 60% o funcionamento das unidades. Este percentual representou uma redução de aproximadamente 40% no custo mensal de energia elétrica na elevatória Senac.

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

BAO, Y.; MAYS, L.W. (1990). *Model for Water Distribution System Reliability*. Journal of Hydraulic Engineering, v.116, n.9, p.1119-1137.

CARRIJO, I.B. (2004). *Extração de Regras Operacionais Ótimas de Sistemas de Distribuição de Água através de Algoritmos Genéticos Multiobjetivo e Aprendizado de Máquina.* São Carlos - S.P. 216p. Tese (Doutorado)-Escola de Engenharia de São Carlos da Universidade de São Paulo.

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.; AGRAWAL, S.; PRATAP, A.; MEYARIVANT, T. (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 Internacional 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.; NAFPLOITIS, N.; GOLDBERG, D. (1994). *A Niched Pareto Genetic Algorithm for Multi-Objective Optimization*.In:Proceedings of The First IEEE Conference on Evolutionary Computation, p. 82-87.

JASZKIEWICZ, A. (1999). *Multiobjective Methods Metaheuristic Library for C++.* http://www-idss.cs.put.poznan.pl/~jaszkiewicz/.

KO, S.K.; OH, M.H.; FONTANE, D.G. (1997). *Multiobjective Analysis of Service-Water-Transmission Systems*. Journal of Water Resources Planning and Management, v.123, n.2, March, p.79-83.

KOELLE, E. (1992). *Control Valves Inducing Oscillatory Flow in Hydraulic Networks*. In: The International Conference on Unsteady Flow and Fluid Transients, Durham, United Kingdom. Proceedings, v.1, 15p.

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.

LÉON, C.; MARTÍN, S.; ELENA, J.M. ; LUQUE, J. (2000). *Explore-Hybrid Expert System for Water Networks Management*. Journal of Water Resources Planning and Management, v.126, n.2, March, p. 65-74.

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. (2001). *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 5th 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.

ZESSLER, U.; SHAMIR, U. (1989). *Optimal Operation of Water Distribution Systems*. Journal of Water Resources Planning and Management, v.115, n.6, November, p. 735-752.

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).

ZITZLER, E.; LAUMANNS, 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.

Extraction of Optimum Operational Strategies for Water Distribution Systems Using Multiobjective Genetic Algorithm and Machine Learning – Used in the Goiania Water Macrodistribution System

ABSTRACT

The disordered growth of cities and the lack of funding for the basic sanitation system have rendered water supply systems complex and difficult to operate. Efficient operation of the system is an essential tool to lengthen its life as much as possible, ensuring perfect service to the consumers, besides keeping electric energy and maintenance costs within acceptable levels. It is essential to know the system for effective operation, since through the system, with tools such as hydraulic simulation models, optimization and definition of operational rules, it is possible to create ideal conditions to operate the system units, no longer depending exclusively on personal experience and keeping the model reliable. This study proposes to develop a computational model aiming at the optimum operational control of the drinking water macrodistribution system for the city of Goiania, and the results showed that satisfactory operational strategies can be produced for the system, replacing the personal judgment of the operation.

Key words – distribution systems, multiobjective optimization, genetic algorithms, machine learning.