Dynamic Behavior of Contaminants in the Water Distribution Network of Cuernavaca Mexico, a Real Application of Multiobjective Distributed Reinforcement Learning

No Author Given

No Institute Given

Abstract. Water systems often allow efficient water uses via water reuse and/or recirculation. The design of the network layout connecting water-using processes is a complex problem which involves several criteria to optimize. The use of the water pinch approach to define which of the effluents from unitary operations are most convenient to reuse is a good alternative used by some practitioners. Previously papers have presented an approach to minimize the freshwater consumption and infrastructure cost, which had been tested with real data from the Cuernavaca city water distribution network with good results [14, 15]. One of the challenges identified from previous work, was the necessity to incorporate the dynamic behavior of distribution systems. In this paper the response of the optimization model to changes in the mass charges of contaminants effluents from unitary operations is presented. The test scenario is the distribution system of Cuernavaca México.

1 Introduction

Water pinch technology (WPT) evolved out of the broader concept of process integration of materials and energy and the minimization of emissions and wastes in chemical processes. WPT can be seen as a type of mass-exchange integration involving water-using operations, that enables practicing engineers to answer important questions when retrofitting existing facilities and designing new water-using networks. There are three basic tasks in WPT: a) identification of the minimum freshwater consumption and wastewater generation in water-using operations (analysis), b) water-using network design to comply with the flow rate targets for freshwater and wastewater through water reuse, regeneration, and recycling (synthesis), and c) modification of an existing water-using network to maximize water reuse and minimize wastewater generation through effective process changes (retrofit).

Most WPT problems are formulated as non linear highly restricted programming problems [1, 9, 10]. Important efforts have aimed to make the mathematical models more robust and applicable to real world problems [2, 5, 8]. Other efforts have aimed to apply WPT technology to other fields such as design and retrofit of urban distribution systems [3, 14, 15].

In general, WPT traditionally minimizes freshwater flow rate entering a system, using mass balance and the concentrations of contaminants at the inlet and outlet in all water-using operations as restrictions. Because of the diverse types of water-using operations, treatment effectiveness and cost, and types of contaminants, the criteria for efficient use of water is inherently non linear with multiple and conflicting objectives [2, 8, 9]. Some of the criteria that can easily be identified are: equipment cost minimization, maximization of reliability (amount of contaminant captured at treatment plants) and minimization of wastewater production.

In [15] an optimization bi-objective model was presented and tested with real data from the Cuernavaca city water distribution network considering static behavior of the mass charges from unitary operations. The model was solved using reduced gradient over a linear combination of the two objective functions and *MDQL* (Multiobjective Distributed Q-Learning). Two objectives were considered in [15], freshwater minimization entering the system and infrastructure cost. Proposed optimization model was applied over three test cases: 1) benchmark problem with four unitary operations and one contaminant; 2) real world industrial problem with ten unitary operations and four contaminants proposed in literature [10], 3) the Cuernavaca city water distribution network retrofit problem with six unitary operation two contaminants and three different freshwater sources. In this work the application of an heuristic approach used to solve Markov decision processes *MDQL* were presented, comparison of results were performed over those obtained for the same problems but applying an aggregated approach solved using reduced gradient method. The objective of the cited work was to demonstrate the capabilities of *MDQL* in the solution of highly constrained optimization problems with real data.

In many real world applications, the values of the variables governing the problem can change over time, displacing the optimum and creating what is known as a non-stationary problem. The goal for this type of problems is to maintain an optimal condition in the face of varying conditions of the environment [20]. The search of the optimum then becomes a continuous process. According to the speed of movement of the optimum, it may be necessary to give the task to an automaton [20], but if the position of the optimum in a dynamic process is shifting very rapidly, the way in which the search process follows the extrema takes on a greater significance for the overall quality. In these cases iterative methods such as dynamic programming or stepwise optimization of Bellman are more adequate [20].

MDQL has strong affinity with the characteristics of dynamic programming, it satisfies Bellman's optimum principle [12].

The current study builds on prior work by considering the dynamic behavior of distribution systems in the optimization process. In this paper mass charges or loads of contaminants change in time. The capabilities of *MDQL* to track the moving optimum is evaluated graphically with the evaluation of Pareto solutions obtained after convergence of the algorithm.

2 Mathematical formulation

The mathematical model describing an industrial water demanding process considers two main components: a) the available freshwater sources to satisfy demands, and b) the water-using operations described by loads of contaminants and concentration levels. An example of two sources and two operations is sketched in Figure 1. This figure represents with rectangles the two unitary operations (O_i), and with solid lines on the left side of the operations their corresponding freshwater demands (f_i). Wastewater from operations are represented with dashed lines on the right side of operations. The rest of the connections represent all the potential links between unitary operations (water reuse), leaks, and treatment plants. The direction arrow heads at the end of lines indicate the direction of flux.

The design task is to find the network configuration that minimizes the overall demand for freshwater, $\sum f_i$, (and consequently reduce the wastewater volume $\sum W_i$) compatible with minimum investment cost. In order to complete the design task, the optimization problem is stated in terms of low freshwater consumption, a suitable network topology for water reuse, $X_{i,j}$, and a low investment cost.

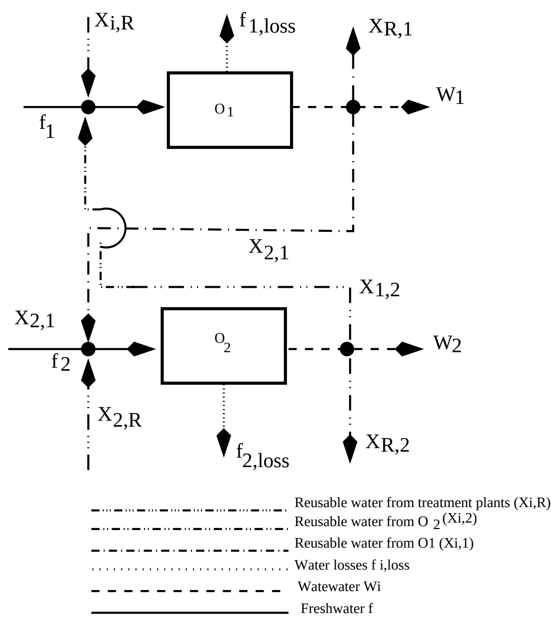

Fig. 1. Block diagram of a water-using system with two sources and two operations.

Unitary operations of demanded water are defined through their contaminant loads, required flow rates, and allowable minimal and maximal contaminant concentrations at influxes and discharges.

The objective functions for freshwater consumption minimization and for infrastructure minimization are represented by Equations 1 and 2.

$$MinZ_1 = F_1 = \sum_j cst_j + TPC, \tag{1}$$

$$MinZ_2 = F_2 = \sum_i f_i \tag{2}$$

Where: F_1 is the total cost of the distribution network considering the connection of freshwater sources to unitary operations receiving water directly, and the connection for reusing water between unitary operations. The total distribution network cost is composed by the sum of the partial costs, cst_j, of the pipe segments used for connecting freshwater sources to unitary operations and unitary operations to unitary operations, and TPC, the treatment plant construction cost that applies only for new treatment infrastructure. In F_1 we are not considering maintenance and rehabilitation cost because the information required to estimate this type of cost is not sufficiently validated for local conditions of Cuernavaca. But also because the infrastructure cost is much greater than maintenance and rehabilitation cost, and for the economic rule that the infrastructure designed will be used for a payable service and the cost of the service is estimated, most of the times, considering production costs which consider rehabilitation and maintenance cost.

F_2, is the total freshwater demanded by the system, obtained by the partial demands of freshwater from each of the unitary operations in the system. Partial demands from unitary operations, say operation O_i, are represented as f_i. That is f_i is the partial freshwater demand of operation O_i.

2.1 Infrastructure cost

Evaluation of the first objective function, F_1, depends only on the pipe segment costs in the network. These costs are represented as cst_j, and depend on three variables (see Equation 3): a) pipe length, L_j; b) cost per unit length, PC_j; which depends on the pipe diameter required to transport the demanded flow of water, D_j; and c) a cost factor, CF_j, related to pipe materials required to resist corrosive effects of contaminants.

$$cst_j = L_j \times PC_j \times CF_j \tag{3}$$

As previously mentioned, PC_j depends on the minimum pipe diameter, $D_j = f(Q_j)$, required to transport the water flow through the pipe. The minimum diameter, D_{min_j}, is obtained applying Equation 4; deduced from the definition of flow ($Q = velocity/area$) considering maximum velocities of water in pipes of $2.5m/s$. D_{min_j} is approximated to the closest upper commercial diameter.

Table 1 shows diameters and cost per unit length for commercial pipes considered in this work. The data in Table 1 is only demonstrative and can be substituted with real data from local markets.

$$D_{min} = 0.714\sqrt{Q} \tag{4}$$

where: D_{min} is the minimal pipe diameter in mm required to transport flow rate Q ; $Q \in \{f_i, X_{i,j}, W_i\} \forall i, j$ and is given in m^3/s.

Table 1. Cost per unit length for commercial diameter pipes.

Diameter (mm)	PC $\$/m$	**Diameter** (mm)	PC $\$/m$
99	4.8	500	40.9
150	5.0	610	42.6
200	8.9	762	45.9
250	12.9	838	54.6
300	17.7	$1,016$	69.9
350	23.6	$1,118$	83
400	25.6	$1,219$	94
450	34.1	$1,372$	110

In a similar manner, the factor CF_j is related to the capacity of the pipe segments to resist corrosive effects due to the presence of contaminants in water flows. Values for the CF_j factor are included in Table 2, calculated considering local prices in Mexico for non corrosive pipes.

Table 2. Cost factors for pipes resistant to abrasive effects of contaminants.

Contaminant concentration (mg/l)	CF
$0 \leq c \leq 50$	1.25
$50 < c \leq 100$	1.5
$100 < c \leq 150$	2.0
$150 < c \leq 200$	3.0
$200 < c \leq 500$	5.0
$500 < c$	10.0

Finally the treatment plant construction cost considered in this work is $10\$/l$, that is the construction cost in monetary units per liter of treatment capacity for the plant or plants.

2.2 Freshwater demand

To guarantee steady state conditions in the system, it is necessary to restrict the objective functions by the mass balance between unitary operations, and by the maximum and minimum allowed contaminant concentrations on the influxes and discharges of operations [10].

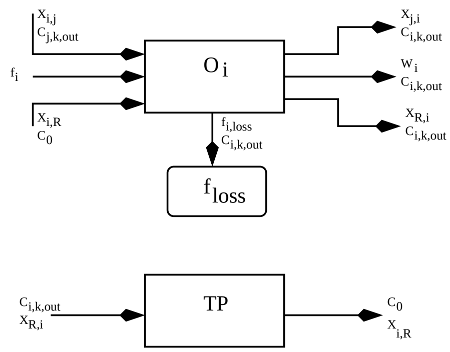

Fig. 2. General structure for mass balance.

The flow-rate required in each unitary operation is related to the mass load of contaminants ($\Delta m_{i,k,tot}$) discharged by operations. This is described in Equation 5.

$$f_i = max_c \frac{\Delta m_{i,k,tot}}{c_{i,k,out}^{max} - c_{i,k,in}^{max}} \tag{5}$$

where f_i is the freshwater flow rate for operation O_i; $\Delta m_{i,k,tot}$ is the total mass transfer for each contaminant, k, to the water used at operation O_i (this term is also known as the contaminant mass charge [3] and is expressed in kg/h); $c_{i,k,out}^{max}$ and $c_{i,k,in}^{max}$ are the maximum allowed concentration of contaminant k on the discharge and influx of operation O_i, in mg/l respectively.

The optimization model depends on the mass balance between all inlets and all outlets of water to the operation O_i. According to Figure 2, the expression for the mass balance has the form shown in Equation 6.

$$f_i + \sum_{j \neq i} X_{i,j} + X_{i,R} - f_{i,loss} - W_i - \sum_{j \neq i} X_{j,i} - X_{R,i} = 0 \tag{6}$$

where, $X_{i,j}$ is the reusable water flow rate from other operations, say O_j, in operation O_i; $X_{i,R}$ is the treated water from the wastewater treatment plants that can be used in operation O_i; $f_{i,loss}$ is the portion considered as water loss in the operation or water consumption by the operation; W_i is the wastewater flow rate from operation O_i; $X_{j,i}$ is the reusable water flow rate from operation O_i in operations O_j; and $X_{R,i}$ is the portion of the discharged water from operation O_i that receives treatment. All flow-rates are represented in m^3/h. TP in Figure 2 represents a treatment plant.

Different contaminants k can be considered in the optimization model. This consideration requires the definition of constraints to restrict the concentration of contaminants at the inlets and outlets of operations, in order to guarantee that water influxes will not affect the operation performance, and to avoid the violation of environmental or operation norms. The satisfaction of these constraints will determine the quantities of fresh and reused water to supply to operations. The contaminant concentration constraint at the influx of the i^{th} operation, $c_{i,k,in}$ is defined by Equation 7.

$$c_{i,k,in} = \frac{\sum_{j \neq i} X_{i,j} c_{j,k,out} + c_{k,0} X_{i,R} - f_{i,loss} c_{i,k,in}^{max}}{\sum_{j \neq i} X_{i,j} + f_i + X_{i,R} - f_{i,loss}} \leq c_{i,k,in}^{max} \tag{7}$$

where, $c_{j,k,out}$ is the concentration of contaminant, k, at the discharge of operation O_j, $c_{k,0}$ is the concentration of contaminant k in the treated water, $c_{i,k,in}^{max}$ is the maximum allowable concentration of contaminant k at the influx of operation O_i. Concentrations are expressed in mg/l.

The same way, contaminant concentration constraint at the outlet of j^{th} operation, $c_{j,k,out}$ is defined by Equation 8.

$$c_{j,k,out} = c_{i,k,in} + \frac{\Delta m_{i,k,tot}}{\sum_{j \neq i} X_{i,j} + f_i + X_{i,R} - f_{i,loss}} \leq c_{i,k,out}^{max} \tag{8}$$

Finally, non negativity constraints are established according to the following equations.

$$\begin{aligned} X_{i,j} &\geq 0, \\ f_i &\geq 0, \\ L_j \times PC_j \times CF_j &\geq 0. \end{aligned}$$

3 Multiple Objective Distributed Q-Learning(MDQL)

Taking advantage of some of the characteristics of evolutionary approaches, optimization problems can be solved considering the search processes of a Markov decision problem. Similar ideas have been previously used with the Ant Colony Optimization Meta-heuristic [6, 7]. The main difference between ant Colony Meta-heuristic and *MDQL* is the way value functions are updated, being the actualization rule used in *MDQL* based on Markov decision precesses theory (see [18] and [13]).

MDQL considers a group of agents searching a terminal state, s_t, in an environment formed by a set of states, S. The set of states, or environment is constructed with the division of the parameter space into a fixed number of parts, considering that all the decision variables can be discretized into a finite number of divisions. Each division is considered as a state, as illustrated in Figure 3. An environment with these characteristics allows the agents to propose values for each one of the decision variables in the problem.

For each state, $s \in S$, a set of actions, A_s, is established, see Figure 3. All state-action pairs have an associated value function, $Q(s,a)$, indicating the goodness of taking action a in state s, for reaching a terminal state $s_t \in S$ (complete a task).

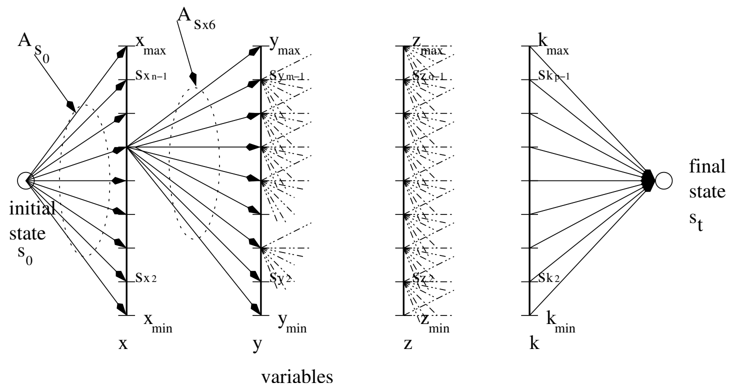

Fig. 3. Variable space division for MDQL.

The search mechanism for an agent in MDQL operates when an agent located in a state selects an action based on its value function, $Q(s,a)$. Most of the times the agent selects the best evaluated action (the action with the higher estimated value for $Q(s,a)$), and sometimes a random action is selected with a probability $\epsilon \approx 0$. Action value functions are updated depending on how useful an action can be for an agent to reach a terminal state. This behavior is adjusted with the help of a reward value, $r \in \Re$, and the value function for the best evaluated action in the future state reached by the agent after the execution of the selected action, $Q(s',a')$. This update rule is expressed in Equation 9. Each action moves the agent to a state of the next consecutive variable, i.e. assigns a value in the discretized space of the next consecutive variable. Figure 4 shows two traces of two different agents. Each of the two traces represents a solution to the optimization problem, that is a set of values for the parameters of the problem. A trace is formed by a value for each of the decision variables in the problem, which are considered as states in the environment. Actions in states correspond to the transportation to any of the states or partition in the next decision variable.

$$Q(s,a) \leftarrow Q(s,a) + \alpha \left[r + \gamma \max_{a' \in A'_s} Q(s',a') - Q(s,a) \right] \tag{9}$$

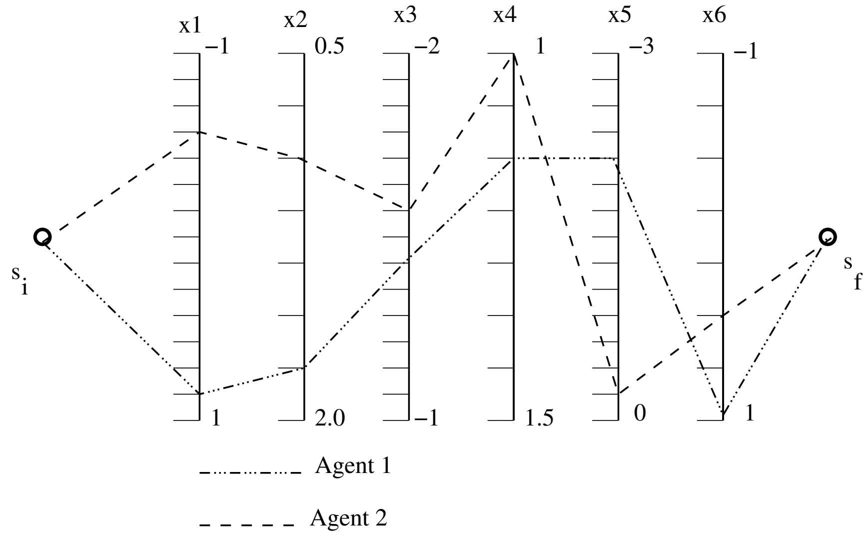

Fig. 4. An example of a path taken by two agents in the MDQL implementation.

where $Q(s,a)$ is the value function for the action, $(0 \leq \alpha \leq 1)$ is the learning step, [1],$(0 \leq \gamma \leq 1)$ is a discount parameter [2] r is an arbitrary reward value, $r \in \Re$, γ is a discount factor, s' and a' are the next state and the best evaluated action for s' respectively.

As an agent explores the state space, the $Q(s,a)$ estimates improve gradually, and, eventually, each $\max_{a' \in A'_s} Q(s',a')$ approaches: $E\left\{\sum_{n=1}^{\infty} \gamma^{n-1} r_{t+n}\right\}$ [18]. Here r_t is the reward received at time t due the action chosen at time $t-1$. Watkins and Dayan [21] have shown that this Q-learning algorithm converges to an optimal decision policy for a finite Markov decision process.

In *MDQL* there is a group of agents, instead of a single agent, interacting with the environment described above, and since the task for the agents is the

[1] In dynamic programming, action-value methods there is the need to estimate actions values as a sample averages of observed rewards. The obvious implementation is to maintain, for each action a, a record of all the rewards, r, that have followed the selection of that action $Q_t(s,a) = \frac{r_1+r_2+\ldots+r_{k_a}}{k_a}$, where $r_1,\ldots,r_{k_a}$ are all the rewards received following all selections of action a prior to play t. So it is easy to devise incremental update formulas for computing averages with small, constant computation required to process each new reward. For some action, let Q_k denote the average of the first k rewards (not to be confused with $Q_k(s,a)$, the average for action a at the kth *play*). Given this average and a $(k+1)$st reward, r_{k+1}. then the average of all $k+1$ rewards can be computed by $Q_{k+1} = Q_k + \alpha[r_{k+1} - Q_k]$, being $\alpha = \frac{1}{k+1}$

[2] The discount parameter determines the present value for future rewards: a reward received k time step in the future is worth only γ^{k-1} times of what it would value as bound as the reward sequence r_k is bounded. If $\gamma = 0$, the agent is "myopic" in being concerned only with maximizing immediate rewards.

construction of the Pareto set, the original *Q-Learning* [21] algorithm must be adapted. The main adaptations considered in *MDQL* are listed below.

- Decision variables in the environment have a predefined order, the agents move in the decision variables space obeying this order, so the definition of the values for the decision variables is made in the same order by all the agents. Each agent assigns a value for one decision variable at a time.
- When all the agents finish (set values for all the decision variables), all solutions are evaluated using the Pareto dominance criterion. Environments for non dominated solutions and solutions that violate any constraint remain in memory to be used in future episodes.
- Agents are randomly assigned to the environments in memory.
- Action values are updated in two stages. The first is made when agents make a transition using a 'map' of the environment. Maps are constructed making a copy of the environments in memory, and are used by agents to show to the rest of the agents the experience acquired during the search process [13]. This experience is represented by the actualization of the action value functions in the 'map' using the *Q-learning* rule of Equation 9. At the end of an episode and after the evaluation of solutions, non dominated solutions receive a positive reward and solutions violating any constraint receive a negative reward, which is used to update the original value functions in the environment where they were found (second stage). After the update procedure, all 'maps' are destroyed and a new episode initiates. More details of *MDQL* algorithm can be found in [11] and [4].

To verify the *MDQL* capability to solve the bi-objective optimization problem presented in section 2 two instances of the problem, proposed in [2] and and instance of the water distribution system of Cuernavaca problem considering static contaminant discharges were used. Obtained results and analysis can be consulted with detail in [15].

The main objective in this paper is to show the response of *MDQL* to dynamic changes of contaminant discharges for the water distribution system of Cuernavaca.

4 Water distribution system of Cuernavaca

There are three different types of sources of freshwater in the city, according to the National Water Commission (NWC): 42 water springs supplying $1,409\ l/s$, 328 deep wells with a contribution of $1,503.58\ l/s$, and water wheels contributing with $751.50\ l/s$[3].

Water users are classified into five categories according to the water works user census. A brief description of the kind of exploitation given to water by

[3] It is important to note that the net extractions and run offs from the sources reported are greater because they also supply freshwater to other other towns close to Cuernavaca.

each category is given below, accompanied with their freshwater demand taken from [17]. In order to be consistent with the nomenclature previously used, every category is considered as an unitary operation.

Self service (SS): Users that have its own source to satisfy any kind of needs including human consumption.

Industrial (I): Users exploiting water to operate only industrial processes in which there are no human needs to satisfy.

Agriculture (A): Covers all the users exploiting freshwater only for irrigation. The main crops cultivated in the region of Cuernavaca are rice, corn, grass and rose trees.

Services (S): Users with high consumption rates, such as hotels, schools, restaurants, supermarkets, etc.

Urban & Public (UP): Most of the domestic users in the city, including small schools, stores, public offices and small workshops.

Multiple (M): Users not classified in any of the previous categories with an activity that can be classified as a service, but with less consumption rate.

It is relevant to note that part of the demanded water is consumed by the operation itself, another part cannot be registered and is considered as a loss caused by leaks occurring along the distribution systems. The rest is declared as wastewater and is supposedly discharged with the effluents to the receiving water bodies. For Cuernavaca city this body is the Apatlaco river. It is estimated that the water consumption and the flow lost in leaks is about 43.41% of the water demanded.

Two contaminants indexes are considered, in connection with the contaminants threw by the operations to the effluents, 5 day biochemical oxygen demand (BOD_5) and total suspended solids (TSS). These indexes are used in the general water quality index, according to the NOM-001-ECOL-1996 standard, which is the Mexican official standard for wastewater discharges. Wastewater treatment plants treat 339.15 l/s to BOD_5 and TTS mean concentration of 50 mg/l according to the data reported in the literature [3].

Values for both water quality indexes, $c_{i,k,out}^{max}$, were established using information from studies that evaluated the degree of contamination in the Apatlaco river [16]. For both contaminants, the concentration in the freshwater supplied to the system is considered to be zero, see Table 3.

Figures 5 and 6 show the variation of contaminant charges of the two more contaminant operations, agriculture and urban and public. The values represented in both Figures were obtained extracting samples of water and analyzing its contaminant concentration. Each of the presented values in Figures were obtained averaging 30 samples, that is, a sample were extracted every three hours during 30 days. Contaminant charges from the rest four operations are considered as static because the relative small amount of mass they throw. So, Figures 5 and 6 can be considered variation laws for contaminants discharged by monitored users. These laws are used to verify the behavior of Pareto fronts constructed by *MDQL*.

Table 3. Inflow and outflow limit concentration and the min (1) and max (2) contaminant mass charge for all current operations in the city of Cuernavaca.

Operation O	BOD_5				TSS			
	$c^{max}_{i,A,in}$	$c^{max}_{i,A,out}$	$(\Delta m_{i,tot})_{1,2}$		$c^{max}_{i,B,in}$	$c^{max}_{i,B,out}$	$(\Delta m_{i,tot})_{1,2}$	
	mg/l	mg/l	kg/h		mg/l	mg/l	kg/h	
UP	0.00	220.00	500	$1,767.74$	0.00	220.00	450	$1,403.07$
S	0.00	220.00	3	9.53	0.00	200.00	2.4	7.56
A	50.00	350.00	150	449.57	50.00	300.00	150	449.57
M	0.00	220.00	0	1.32	0.00	220.00	0	1.05
I	0.00	874.00	20	85.57	0.00	371.00	5	36.32
SS	0.00	220.00	0	0.60	0.00	240.00	0	0.73

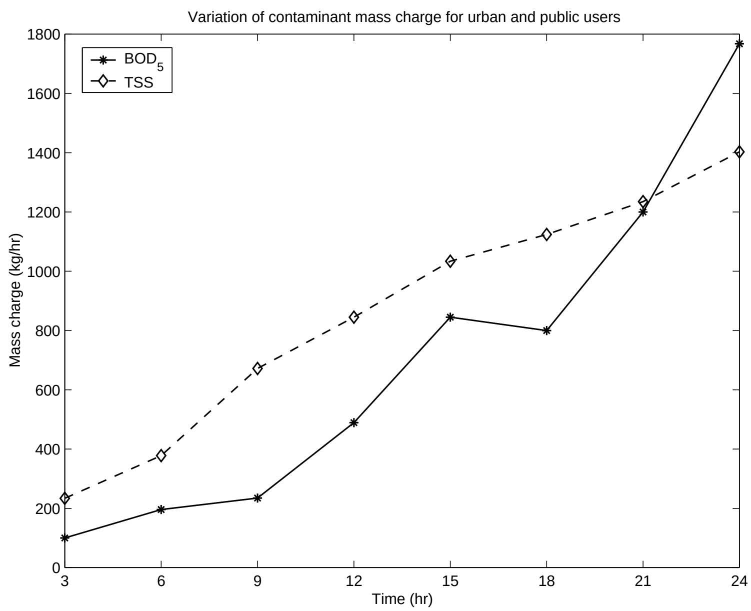

Fig. 5. Variation of mass charge for urban and public users

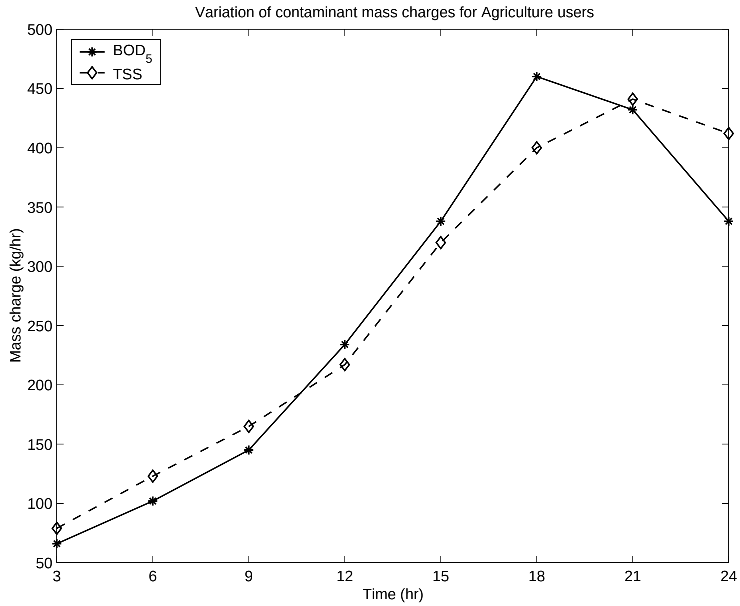

Fig. 6. Variation of mass charge for agriculture users

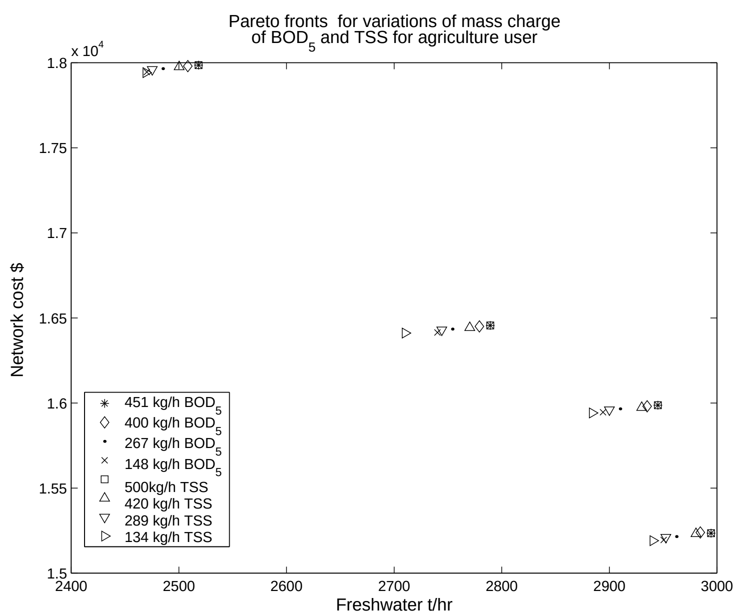

Fig. 7. Pareto fronts obtained for the variations in mass charge for BOD_5 and TSS for agriculture users

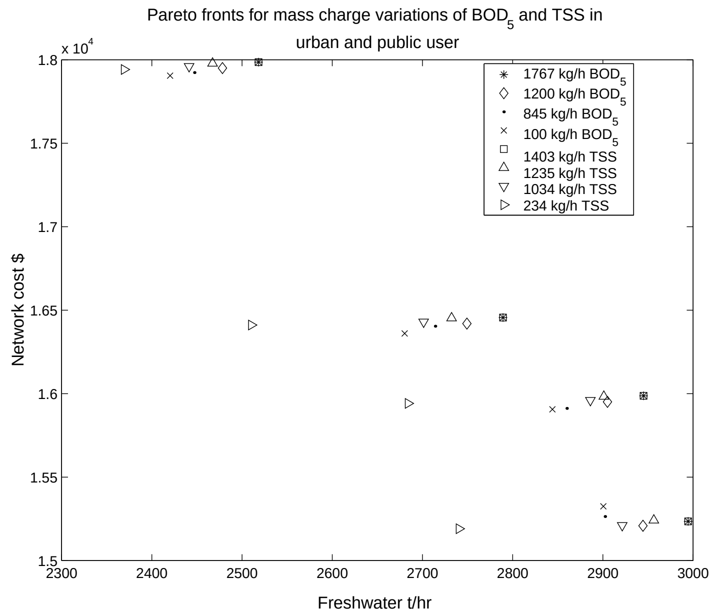

Fig. 8. Pareto fronts obtained for the variations in mass charge for BOD_5 and TSS for urban and public users

5 Results

The Cuernavaca city water distribution network solution considers the reduction of water losses from 43 to 25% and the treatment of waste water using the current operational capacity, that is $339.65 l/s$.

The *MDQL* operation parameters used for all test cases were: $\alpha = 0.1$, $\gamma = 0.9$, $\epsilon = 0.01$ and $r = 1$ for non dominated solutions and $r = -1$ for solutions violating constraints. Previous values for the operation parameters in *MDQL* are in some sense typical and were originally suggested in [19]. Some work related with the sensitivity of the algorithm to these parameters is presented in [12] using benchmark evaluation functions. The conclusion of the previous work indicates that the best combination of values for the operation parameters is to consider $\alpha \approx 0$, $\gamma \approx 1$ and $\epsilon \approx 0$.

Discharges from unitary operations to effluents were evaluated considering four different values of the mass charge of each of the two contaminants. Contaminant mass charges were taken from the variation curves presented in Figures 5 and 6. Four decrements for each of the two contaminants in the two most contaminant operations or users were evaluated, going from the higher to the lower mass charges. The values of the mass charge considered are included in the legend of Figures 7 and 8.

Obtained results are presented in Figures 7 and 8 including the best approximation to the Pareto front from ten executions of the algorithm. Criterion used to solve the dynamic optimization problem presented considers that agents in the algorithm start from the Pareto solutions obtained. That is, *MDQL* starts with a deterministic environment constructed with fixed values for the value functions for the first parameter of mass charge in the variation laws; when convergence is reached and a Pareto set is obtained, a new cycle is started, changing contaminant mass discharge to the next value in the variation laws. Agents start searching (adapting solutions) from the existing environments which correspond to the previous solutions obtained for the previous contaminant mass discharge. Searching for new solutions, from the last Pareto set, given the new values of contaminants discharges, significantly reduces convergence times.

The leftmost graphic includes the results obtained with the variation mass charges of BOD_5 and TSS by the agriculture user. It can be seen that *MDQL* get four solutions in the Pareto front for each different value of the mass charge, behavior that indicates that there are no more solutions in the Pareto front. The rightmost graphic includes the Pareto fronts obtained for each value of the mass charge for the two contaminants discharged by the urban and public user. The same behavior is appreciated for this operation.

There is not established criteria to determine if a solution is good on not in problems for which there is no evidence of the location of optimal solutions. For the problem presented in this paper, and considering that the problem is formulated with real data about the behavior of contaminants in the effluents from unitary operations, and there the aim is to identify layouts between unitary operations that minimize the governing criteria satisfying imposed restrictions, an alternative is to consult human experts. The solutions generated by *MDQL* were given to an expert in process engineering. The expert, analyzed the solutions and qualified them in terms of its operational correctness.

From Figures 7 and 8 it is possible to conclude that freshwater minimization criterion is more sensible to the variation of contaminant mass discharges, while cost remains almost on the same values. It is important to note that *MDQL* finds the same number of solutions for all values of contaminant mass discharges, which is a consistence characteristic of the algorithm and the mathematical model. In order to verify that there are not other Pareto optimal solutions not identified *MDQL*, exhaustive search was made and no other solutions were found.

MDQL performance was evaluated calculating the number of solutions in the Pareto set (Pareto count) along the execution. Figure 9 shows values for Pareto count for the first 200 function evaluations. It can be appreciated that *MDQL* identifies all solutions in the Pareto set in a relative short number of evaluations. To evaluate the behavior of *MDQL* with more function evaluations, Figure 10 is included. No changes on the number of solutions can be detected, also solutions remains the same. This behavior can be considered as a consistence characteristic, as mentioned previously.

Adaptation capability of *MDQL* to changes in the contaminant mass discharges was also evaluated. Figure 11 shows the comparison of Pareto count performance metric when a change in the mass discharge occurs. The graph presented in Figure 11 was constructed considering the first value for the BOD_5

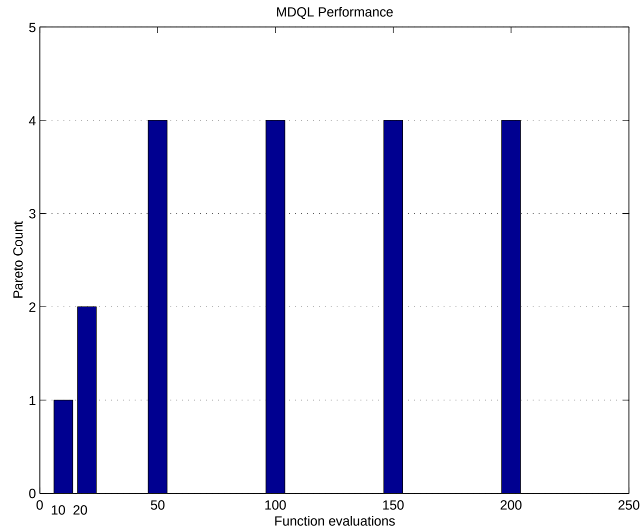

Fig. 9. *MDQL* Performance on the solution of the Cuernavaca city water distribution network for the first 200 function evaluations

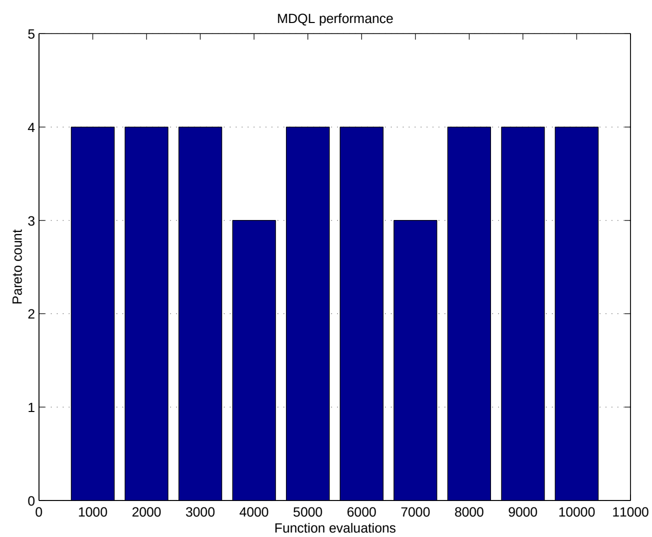

Fig. 10. *MDQL* Performance on the solution of the Cuernavaca city water distribution network after 1000 function evaluations

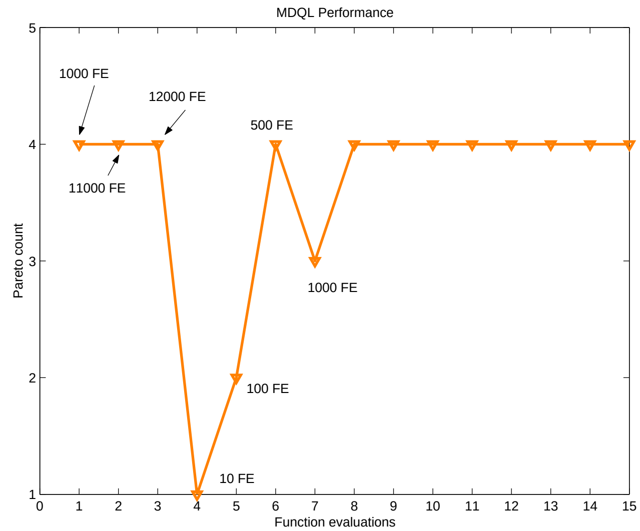

Fig. 11. Adaptation behavior of *MDQL* for a change on mass discharge rate

contaminant discharged by the agriculture user $(451t/h)$. Change in the mass of contaminant discharged was made just after 12000 function evaluations were completed. Pareto count results indicate that after 10 evaluations of objective functions, only 1 Pareto optimal solution were found; at 100 evaluations 2 Pareto optimal solutions; and at 500, the four solutions in the Pareto set were obtained by *MDQL*. After 500 evaluations, the number of solutions, and solutions remain the same.

Previous behavior indicate that adaptation mechanism is not as efficient to adapt Pareto solutions as *MDQL* is to find initial solutions, that is, *MDQL* shows better performance, less function evaluations to reach the Pareto set, starting with initial values of mass discharges than adapting solutions when changes in mass discharges occurs, at least for this problem.

6 Conclusions and future work

In this work we presented a multi-objective optimization problem for water distribution systems using water pinch technology criteria, we evaluated the multi-objective optimization model, and we verified the capability of *MDQL* to solve complex real problems with highly restricted non convex spaces. *MDQL* capability to solve dynamic changes in the decision variables was also tested with the variation of contaminant charges from operations.

The water pinch optimization model considers more than one criteria. The model considers the reuse of wastewater from operations, wastewater treatment, consumption flow-rates and leaks in the system. With the reduction of freshwater demands it is possible to guarantee that the quality of the water served to the different users do not violate ecological and sanitary norms. The bi-objective optimization model operates considering mass balances between operations, freshwater sources, wastewater treatment plants, and wastewater disposal effluents. Contaminants loads from operations to water flows are restricted by environmental and operational constraints, and their variation in time is implemented in the model, resulting in a highly non linear dynamic optimization model.

Solutions to water pinch problems, represent important technical challenges that are only partially solved by the industry. The results presented here represent an example of how real applications can be solved with the participation of multidisciplinary teams involving researches from different communities, as in this case.

As future work we are considering implementing constraints to select more efficiently different processes. For example, if wastewater treatment technology is selected in terms of the type of contaminants, the mass remotion could be made more effective and the system more efficient if the proper process is selected and optimized in terms of cost and efficiency. Another important aspect to implement is the cost function, which needs to be extended in order to quantify operation costs, reuse costs, and other economic factors affecting the operation of a system with these characteristics.

There is also an interest to increment the precision of Pareto solutions obtained with *MDQL*, which at this moment depends on partition size adopted for the parameter space. We believe that creating an adaptive mechanism for the size of partitions can be helpful in the approximation to more precise solutions.

References

1. Alcocer, V. and Arreguín, F., Minimización de agua de primer uso en procesos industriales a través de técnicas de optimización. *Memorias del XX Congreso Latinoamericano de Hidráulica*, La Habana, Cuba, september 2002, (in spanish).
2. Alva-Argaez, A *An Automated design tool for water and wastewater minimization*, PhD Thesis, Department of Process Integration University of Manchester IST, UK 1999.
3. Arreguín, C., F., Alcocer, Y. V. Modelación sistémica del uso eficiente del agua. *Revista Ingeniería Hidráulica en México*. Vol. XIX, No. 3, july, 2004.
4. Coello, C. and Mariano, C., (2002) *Multiple criteria optimization. state of the art annotated bibliographic surveys*, Edited by Matthias Ehrgott, Xavier Gambibelux, Kluwer Academic Publishers, Chapter 6.*Evolutionary Algorithms and Multiobjective Optimization*, pp. 277-331.
5. Coetzer, D., Stanley, C. and Kumana, J. Systemic Reduction of Industrial Water Use and Wastewater Generation, In *Proc. of AIChE National Meeting*, Houston, March 1997.
6. Dorigo, M. (1992) *Optimization, Learning and Natural Algorithms*. Ph.D.Thesis, Politecnico di Milano, Italy, in Italian.
7. Dorigo M., V. Maniezzo and A. Colorni (1996). The Ant System: Optimization by a Colony of Cooperating Agents. *IEEE Transactions on Systems, Man, and Cybernetics-Part B*, 26(1):29-41

8. Galan, B., and Grossmann, I.E., Optimal Design of Distributed Wastewater Treatment Networks, *Ind. Eng. Chem. Res.*, 37:4036, 1998.
9. Kuo, W.C. and Smith, R. Efluent treatment system design, *Chem Eng. Sci.*, 52:4273, 1997.
10. Mann, G.J. and Liu, Y.A. *Industrial Water reuse and wastewater minimization* McGrawHill Eds., 1999.
11. Mariano, C. and Morales, E., (2000) A new approach for the solution of multiple objective optimization problems based on reinforcement learning, *Lecture Notes in Artificial Intelligence, Proceedings of the Mexican International Conference on Artificial Intelligence*, pp. 212-223, Acapulco, Mex, April, 2000.
12. Mariano, C., (2001)*Reinforcement learning in multiobjective optimization* PhD Thesis in computer Science, Instituto Tecnológico y de Estudios Superiores de Monterrey, Campus Cuernavaca, March, 2002, Cuernavaca, Mor., México.
13. Mariano, C. and Morales, E., (2001) A new updating strategy for reinforcement learning based on Q-learning, In P. Flach and L. de Raeldt (eds.), *Springer Verlag, Lecture Notes in Artificial Intelligence,vol 2167: 12th European conference on Machine Learning*, pages 324-335, 2001.
14. Mariano, C., Alcocer, V., and Morales E. (2005) Diseño de sistemas hidráulicos bajo criterios de optimización de puntos de pliegue y múltiples criterios, *Ingeniería hidráulica en México*, vol XX, no. 3, pp. 31-42, july-september, 2005.
15. Mariano, C., Alcocer, V., and Morales, E. (2007) Multi-objective optimization of water using systems, *Europen Journal on Operational Research*, volumen 181-3, september (2007), pp. 1691-1707.
16. National Water Commission, *Clasification study in the Apatlaco river*, Sanitation and Water Quality Department from the Morelos state office, 1996 (In Spanish).
17. National Water Commission, *Hydro-geological update of the aquifer of Cuernavaca, state of Morelos* 1999, CNA 2000, (In Spanish).
18. Putterman, M. *Markov Decision Processes Discrete Stochastic Dynamic Programming*, Wiley Series in Probability and Mathematical Statistic, (1994)
19. Sutton R., and Barto G. *Reinforcement Learning An Introduction*, MIT Press, (1998).
20. Schwefel H., *Evolution and Optimum Seeking.* John Wiley and Sons, New York, USA, 1995.
21. Watkins, C., Dayan, P., (1992) Q-Learning, *Machine Learning*, 3:279-292, 1992.