ROBUST MULTI-OBJECTIVE OPTIMIZATION OF WATER DISTRIBUTION NETWORKS

Taishi Ohno, Hernán Aguirre, Kiyoshi Tanaka
Faculty of Engineering, Shinshu University, Wakasato, Nagano-shi, Japan
15tm209f@shinshu-u.ac.jp, ahernan@shinshu-u.ac.jp, ktanaka@shinshu-u.ac.jp

Abstract This work studies the incorporation of robustness to changes in water demand within the evolutionary search in order to enhance the design optimization of water distribution networks. An unconstrained multi-objective formulation of the problem is used together with the AεSεH algorithm, a multi- and many-objective evolutionary algorithm known to scale up well with larger population sizes and number of objectives. An effective incorporation of robustness within the evolutionary process will open the possibility to incorporate additional important objectives of water distribution networks that can be optimized simultaneously.

Keywords: Evolutionary algorithms, Multi- and many-objective algorithm AεSεH, Multi-objective design optimization, Water distribution networks.

1. Introduction

A water distribution network refers to the drinking water supply backbone of an urban area. Its main components are reservoirs, pipes, and nodes. The diameters of the pipe are chosen from a set of commercially available sizes and there is a certain demand of water associated to each node that must be fulfilled. The design optimization of water distribution networks aims to obtain optimal combinations of pipe diameters that minimize cost while keeping enough pressure at the nodes to satisfy the water demand. In addition, an optimal design must be reliable to pipe failures and robust to changes in water demand.

The design optimization of water distribution networks has been usually formulated as a single objective cost-minimization problem with a constraint on the minimum pressure to be available at the consumer nodes in order to satisfy the demand. In addition, reliability of the optimal solution to failures in pipes and robustness to changes in water demand have been usually verified after optimization.

Recently, design optimization of water distribution networks using multi-objective unconstrained formulations has been explored [5], where objectives related to cost, nodal pressure, and reliability to pipe failures are optimized simultaneously. The multi-objective formulation allows to generate a number of non-dominated solutions to provide the decision maker with insights on the trade-off between the objectives. Also, it incorporates the reliability to pipe failures within the optimization process rather than as a posterior step.

A comparison between the popular NSGA-II and SMS-EMOA on benchmark water distribution problems is presented in [7, 8] following the same multi-objective formulation used in [5]. Better results are reported for SMS-EMOA than for NSGA-II. This is attributed to the better scalability of the SMS-EMOA approach for problems with three or more objectives. In the same work, an attempt to consider within the evolutionary optimization process the robustness of the network to changes in water demand is pursued. However, for the budget of fitness evaluations and population size used in [7, 8], the conclusion was that solutions optimized for a single profile of water demand are more robust to changes in demand than solutions optimized simultaneously for multiple profiles. This is rather counterintuitive and unexpected.

In this work, our aim is to clarify whether robustness to changes in water demand can be effectively incorporated within the evolutionary process in order to enhance the design optimization of water distribution networks. We follow the same multi-objective formulation used in [5, 7, 8] incorporating the AεSεH algorithm, a multi- and many-objective evolutionary algorithm known to scale up well with larger population sizes and number of objectives. An effective incorporation of robustness within the evolutionary process will also open the possibility to incorporate additional important objectives of water distribution networks that can be optimized simultaneously.

2. Method

2.1 Optimization System

A water distribution network consists of pipes, nodes (pipe junctions), pumps, valves and storage tanks or reservoirs. For each node there is associated a water demand that must be satisfied. To simulate the water distribution network we use EPANET [9]. EPANET performs extended period simulation of hydraulic and water quality behavior within pressurized pipe networks. EPANET tracks the flow of water in each pipe,

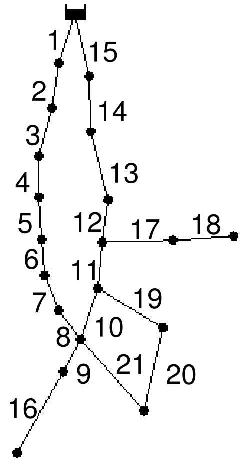


Figure 1: New York Water Distribution Network Model

Table 1: New York Water Supply Pipe Types [8]

Code	D (in)	Price ($/ft)	H-W
0	36	93.5	100
1	48	134	100
2	60	176	100
3	72	221	100
4	84	267	100
5	96	316	100
6	108	365	100
7	120	417	100
8	132	469	100
9	144	522	100
10	156	577	100
11	168	632	100
12	180	689	100
13	192	746	100
14	204	802	100

the pressure at each node, and the height of water in each tank. In can also track the concentration of a chemical species throughout the network. In addition, water age and source tracing can also be simulated.

In this work we assume a remediation problem, where there is a water distribution network in operation and all its pipes could be replaced by new ones. Thus, the coordinates where the pipes are laid down and the lengths of the pipes are already determined. Also, the available commercial diameters of the pipes and the baseline water demand for the nodes are known. We are interested in finding the diameter of the pipes that render robust configurations for the water distribution network when the nodes of the network are subjected to various demands of water.

An evolutionary algorithm is used to search for optimal combinations of pipes diameters. Thus, a solution encodes the diameters of all pipes of the network. For each solution explored by the evolutionary algorithm, we combine the information of the network with the diameters of the pipes provided by the solution to create the required input file for EPANET, run the simulator, and use its output to compute the fitness of the solution. As benchmark problem, we use the New York Water Distribution Network model, widely used in the literature and illustrated in Fig. 1.

In the following we explain the algorithm, genetic representation, fitness function, and operators used in this work.

2.2 AεSεH

In this work we use the Adaptive ε-Sampling and ε-Hood (AεSεH) [1, 2] algorithm to search optimal solutions. AεSεH is an elitist evolutionary multi- and many-objective optimizer that applies ε-dominance [6] principles both for survival selection and parent selection.

AεSεH follows the main steps of a population-based evolutionary algorithm, i.e., parent selection, offspring creation and survival selection, adjusting its operation depending on whether the population contains dominated solutions or not.

To perform survival selection, the current population and its offspring are combined and divided into non-dominated fronts using the non-dominated sorting procedure. If the number of non-dominated solutions in the first front is smaller than the population size, the sorted fronts of non-dominated solutions are copied one at a time to the next population until it is filled; if the last copied front overfills the population, the required number of solutions are chosen randomly from it to have the exact number specified by the population size. On the other hand, if the number of non-dominated solutions in the first front is larger than the population size, the first front is truncated to the size of the population using the *ε-sampling* procedure. *ε-sampling* randomly chooses solutions from the first front to include them in the surviving population, eliminating from the front those solutions that are ε-dominated by the chosen samples. As a result, solutions in the next population are spaced according to the $\mathbf{f}(\mathbf{x}) \mapsto^{\varepsilon_s} \mathbf{f}^{'}(\mathbf{x})$ mapping function and parameter ε_s used to compute ε-dominance between solutions.

For parent selection, the algorithm first uses a procedure called *ε-hood creation* to cluster solutions in objective space and then applies *ε-hood mating* to select parents. When all solutions in the population are non-dominated, *ε-hood creation* selects *randomly* an individual from the population and applies ε-dominance with mapping function $\mathbf{f}(\mathbf{x}) \mapsto^{\varepsilon_h} \mathbf{f}^{'}(\mathbf{x})$ and parameter ε_h. A neighborhood is formed by the selected solution and its ε_h-dominated solutions. Neighborhood creation is repeated until all solutions in the population have been assigned to a neighborhood. *ε-hood mating* sees the neighborhoods as elements of a list and visits them one at the time in a round-robin schedule. The first two parents are selected *randomly* from the first visited neighborhood in the list. The next two parents are selected randomly from the second neighborhood in the list, and so on. When the end of the list is reached, parent selection continues with the first neighborhood in the list. On the other hand, when dominated solutions are present in the population, *ε-hood creation* makes sure that the solution sampled to create the neighborhood

C_1	C_2	C_3	...	C_i	...	C_P

Figure 2: Genetic Representation

is a non-dominated solution and *ε-hood mating* uses *binary tournaments* based on dominance rank to select parents within the neighborhoods. Both epsilon parameters ε_s and ε_h used in survival selection and neighborhood creation, respectively, are dynamically adapted during the run of the algorithm.

This algorithm has been shown to work effectively on continuous and discrete multi- and many-objective optimization problems [1, 2, 3, 4]. Further details about the algorithm can be found in [1] and [2].

2.3 Genetic Representation

This paper considers n types of pipes, identified by a unique integer code $C = \{1, 2, ..., n\}$. Table 1 shows the codes of the type of pipes of the New York Network used in this work, together with their characteristics such as diameter, unit length price, and Hazen-Williams (H-W) roughness coefficient that depends on the material of the pipe. Fig. 2 illustrates the representation of the solutions used for optimization, where P is the total number of pipes of the water distribution network, C_i is the code of the i-th pipe in the network, i.e $C_i \in C$. Thus, in the New York Problem, the size of the search space is $15^{21} = 4.99 \times 10^{26}$.

2.4 Fitness Function

There are several criteria for evaluating a water distribution network in the literature. Here, we use the total cost of the pipes of the water distribution network, demand supply ratio, and system entropy [7] .

Total cost of the pipes f_1 is expressed as the sum needed for inserting new pipes in the network, priced in segments of unit length with a certain diameter.

$$f_1 = \sum_{i \in P} price\,(C_i) \cdot L_i, \tag{1}$$

where $price\,(C_i)$ is the unit length price of the new pipe i of code C_i (the prices are determined by the diameter of the pipe and are included in the table of available commercial diameters, which is supplied with the test problem), and L_i is the length of the pipe i.

The demand supply ratio is a measure of reliability of the distribution network that computes to what extent the water demand at the various internal nodes in the network is met. To compute the demand supply ratio, first a functional relationship between pressure and water demand at a node is established. This determines the portion of the water demand at a node that could be satisfied and is expressed as follows

$$Q_j^{met} = \begin{cases} 0 & p_j \leq p^{zero} \\ Q_j \sqrt{\frac{p_j - p^{zero}}{p^{req} - p^{zero}}} & p^{zero} \leq p_j \leq p^{req} \\ Q_j & p_j > p^{req}, \end{cases} \tag{2}$$

where j is the node index, N is the total number of nodes, Q_j^{met} is portion of the water demand met at node j, Q_j is the nodal demand (water demand drawn from internal node j), p_j is the effective pressure remaining at internal node j, p^{zero} is the pressure corresponding to zero nodal water demand satisfaction, and p^{req} is the pressure required to satisfy the nodal water demand completely.

The average demand supply ratio f_2, normalized by the nodal demands, is expressed by

$$f_2 = \frac{\sum_{j \in N} Q_j^{met}}{\sum_{j \in N} Q_j}. \tag{3}$$

The system entropy is also a function of reliability of the network. The water demand at a given internal node is ideally met using multiple different paths to that node. The required flow should be distributed over these routes as evenly as possible. This way, should a segment of the network fail, alternative routes exist that could still supply a reasonable part of the demand. The system entropy f_3 is expressed as follows:

$$s_j = \sum_{i:dest_i = j} \frac{-|q_i|}{in_j} \ln \frac{|q_i|}{in_j}, \tag{4}$$

$$S = \sum_{j \in N} \frac{in_j}{IN} s_j + \sum_{j \in N} \frac{-in_j}{IN} \ln \frac{in_j}{IN}, \tag{5}$$

$$f_3 = exp(S), \tag{6}$$

where s_j is the entropy at node j, S is the entropy of the entire network, q_i is the water flow in pipe i, $dest_i$ is the destination of the water flow in pipe i, in_j is the water flow transported towards node j by the incoming pipes i connected to it, and IN is the sum of all incoming flows in the network. Note that system entropy f_3 is expressed as the exponential of S.

These functions are optimized simultaneously. The investment cost of installing water pipes f_1 is to be minimized, whereas the demand supply ratio f_2 and the system entropy f_3 are to be maximized.

2.5 Crossover and Mutation Operators

In this work, we use two point crossover with crossover rate P_c. Mutation is applied to all variables with mutation rate P_m. If a variable is chosen, mutation either increases or decreases with equal probability the code of the pipe by one. That is, mutation changes the current diameter of the pipe to the next higher/lower available diameter.

3. Methods of Evaluation for Evolution and Robustness

In this work we search robust optimal configurations of water distribution networks using two different methods to evaluate solutions during evolution. The first method (Method I) evolves solutions evaluating them using the baseline profile for water demand in the nodes of the network given by the benchmark problem. Thus, the fitness vector of a solutions $\mathbf{x}$ is

$$\mathbf{f}(\mathbf{x}) = (f_1(\mathbf{x}), f_2(\mathbf{x}), f_3(\mathbf{x})). \tag{7}$$

The second method (Method II) establishes in advance a set E of N_E different profiles of water demand for the nodes of the network; during evolution solutions are evaluated in all N_E profiles of the set E and fitness is computed as the average of these evaluations. Thus, the fitness vector of a solutions $\mathbf{x}$ is

$$\mathbf{f}(\mathbf{x}) = \frac{1}{N_E}\sum_{i=1}^{N_E}\mathbf{f}^{(i)}(\mathbf{x}), \tag{8}$$

where $\mathbf{f}^{(i)}(\mathbf{x}) = (f_1^{(i)}(\mathbf{x}), f_2^{(i)}(\mathbf{x}), f_3^{(i)}(\mathbf{x}))$ is the fitness vector of the solution evaluated with the i-th profile of the set E.

The approximation of the Pareto optimal set ($\mathcal{A}_{POS}$) generated by an algorithm consists of the non-dominated solutions computed from all solutions explored during the search. We use another set R of N_R profiles of water demand for the nodes in order to assess the robustness of the evolved solutions and compare the two methods fairly. Thus, the approximations $\mathcal{A}_{POS}$ found by the algorithms are evaluated again with the profiles in R by

$$\mathbf{f}(\mathbf{x}) = \frac{1}{N_R}\sum_{i=1}^{N_R}\mathbf{f}^{(i)}(\mathbf{x}), \tag{9}$$

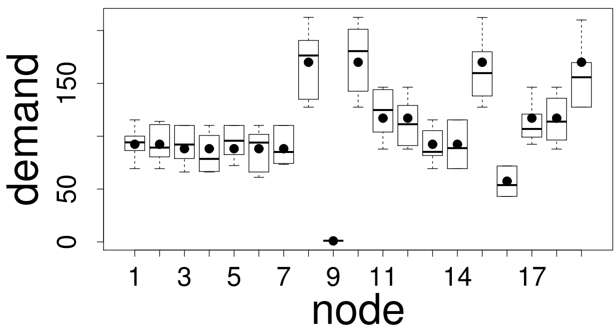


Figure 3: Profiles of water demand to measure robustness of evolved solutions

where $\mathbf{f}^{(i)}(\mathbf{x}) = (f_1^{(i)}(\mathbf{x}), f_2^{(i)}(\mathbf{x}), f_3^{(i)}(\mathbf{x}))$ is the fitness vector of the solution evaluated with the i-th profile of the set R. The set of non-dominated solutions is obtained for each method with this new fitness value and call the new approximation set as $\mathcal{A}_{POS}^{R}$. Both methods of evolutions are compared computing the hypervolume of their approximations $\mathcal{A}_{POS}^{R}$.

The profiles of water demand for sets E and R are created by randomly changing the baseline demand of the nodes by

$$\hat{d}_j = d_j(1 + 0.25 \times U(0,1)), \tag{10}$$

where, d_j is the baseline demand of node j, and $U(0,1)$ is random number sampled from normal distribution with expectation 0 and variance 1. Fig. 3 shows boxplots of the demand profiles used for the nodes of the network to evaluate robustness of the evolved solutions. The dot inside each boxplot shows the baseline demand.

4. Simulation Results

4.1 Experimental Setup

As mentioned above, in this work we use as benchmark problem the New York Model of Water Distribution that consists of a reservoir, 19 nodes, and 21 pipes. For the evolutionary algorithm we use the following parameters. Crossover rate was set to 1.0, mutation rate to $1/|P|$, population size to 300, and number of function evaluations is se to 300,000. The number of water demand profiles to evaluate solutions in Method I is 1 (the baseline profile) and the number of water demand profiles in Method II is $N_E = 10$. Thus, to keep the same number of fitness evaluations in both methods, in Method I the number of generations is set to 1,000 whereas in Method II the number of generations is 100. The number of profiles to evaluate robustness of the approximation sets is $N_R = 200$. Unless stated otherwise, we report results of ten independent runs of the algorithms.

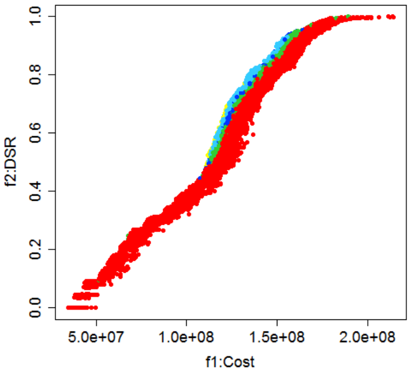
4.2 Method I: Evolution with Single Profile for Water Demand

The three-objectives fitness vector of the approximation of the Pareto optimal set $\mathcal{A}_{POS}$ obtained by Method I for one of the runs of the algorithm is projected to two-objectives planes as shown in Fig. 4. Here, solutions are colored by the value of the third not displayed objective. The colors codes are yellow, light blue, dark blue, green, and red in order of objective value from low to high. The range between the maximum and minimum objective values is divided evenly into five subranges and assigned to the colors code.

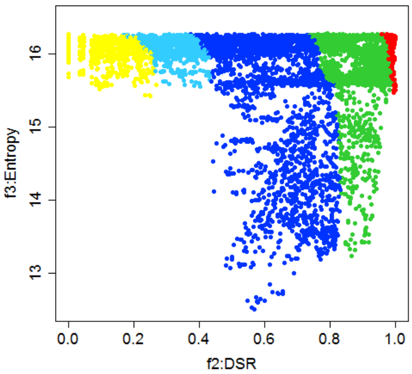
By displaying the fitness of solutions evenly divided by color we can verify the trade-offs in objective space between solutions in the Pareto optimal set. As an example, we focus on Fig. 4 (b). Recall that the demand supply ratio DSR f_2 and entropy f_3 are maximized, whereas cost f_1 is minimized. Fig. 4 (b) plots f_2 and f_3, coloring by f_1. Note that as DSR f_2 increases (DSR gets better) cost f_1 changes from yellow to light blue, $\cdots$, to red (cost gets higher). Hence, there is a clear trade-off between DSR and cost; improving DSR makes the network more expensive and vice versa. It is also interesting to note that high values of entropy can be achieved for both low and high cost networks.

4.3 Method II: Evolution with Multiple Profiles for Water Demand

Similar to Method I, Fig. 5 shows the three-objectives fitness vector of the approximation $\mathcal{A}_{POS}$ by Method II projected to two-objectives planes. Fig. 5 looks similar to Fig. 4 and the trade-off between objective functions can be verified, particularly between cost and DSR. However, note that in Fig. 5 only networks with high entropy remain. That is, solutions with low entropy have been eliminated from the Pareto optimal set by the method that evaluates solutions with multiple profiles of water demand during the evolutionary process.

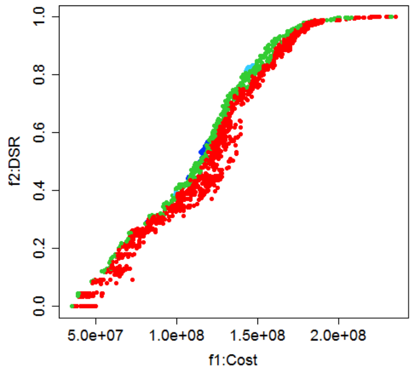


(a) Cost - DSR, colored by Entropy

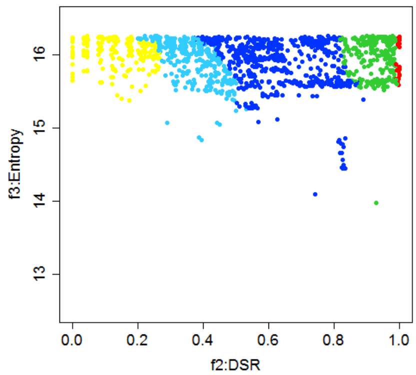


(b) DSR - Entropy, colored by Cost

Figure 4: Method I



(a) Cost - DSR, colored by Entropy



(b) DSR - Entropy, colored by Cost

Figure 5: Method II

4.4 Analysis Based on Hypervolume

In this section we compare the hypervolume HV of the approximations $\mathcal{A}^{R}_{POS}$ obtained by both methods. The fitness values given by Eq.(9) are used to compute the HV. Table 2 reports the average HV in ten independent runs of the algorithms. The standard deviation of HV is presented in parenthesis. The table also shows the average standard deviation of the objective values for each fitness function and the average cardinality of the approximations for both methods. Note that the HV value by Method II is higher than by Method I. This shows that overall solutions with better convergence and diversity in objective space are found by Method II, suggesting that by using Method II is posible to find designs of the water distribution network that can cope better with changes in the demand of water.

Table 2: Hypervolume and standard deviation of the approximations $\mathcal{A}^R_{POS}$

	Method I	Method II
HV	0.59022 (0.026)	0.62216 (0.028)
stdev f_1	0.21	0.21
stdev f_2	0.29	0.28
stdev f_3	0.21	0.19
Number of POS	6289.9 (336.99)	1407.1 (96.68)

Remember that the number of fitness evaluations is the same in both methods, however the number o generations is different. This indicates that robust evolution can be performed in this problem using a reduced number of generations evaluating simultaneously a number of water demand profiles rather than evolving solutions for a larger number of generations evaluating just one profile. The oposite conclusion was reached in [7, 8], where solutions by the method that evolves solutions evaluating just one demand profile showed higher robustness to changes in the demand of water. This is because [7, 8] used fewer fitness evaluations, Method II evolved for too few generations and could not converge.

Method II achieves solutions with f_3 higher than Method I, as can be seen in Fig. 5 compared to Fig. 4. Also, the standard deviation of f_3 by Method II is smaller than Method I as shown in Table 2. As mentioned above, solutions with low f_3 are eliminated by Method II. This suggests that changes in demand affect system entropy.

It is also worth noting from Table 2 that Method II finds fewer solutions than Method I. One reason for this is that Method I runs for a larger number of generations and explores many more solutions than Method II.

5. Conclusion

This work has shown that robustness to changes in water demand can be effectively incorporated within the evolutionary process in order to enhance the multi-objective design optimization of water distribution networks. We followed an unconstrained multi-objective problem formulation where three objectives were optimized simultaneously: cost of the network, nodal pressure based on demand supply ratio, and reliability to pipe failures based on a measure of entropy. Robustness to changes in the demand of water was pursued and two approaches were compared. One approach to robustness was to guide evolution based on a single profile of water demand, whereas the other approach used multiple profiles. We showed that using the multi- and many-objective evolutionary AεSεH algorithm the approach that evolves solutions based on multiple profiles attained significantly more robust networks. This clarifies previous counterintuitive and misleading findings regarding robustness incorporated within NSGA-II and SMS-EMOA reported elsewhere and opens the possibility to incorporate additional important objectives of water distribution networks to be optimized simultaneously. In this work we used the hypervolume to compare the robustness of the optimal solutions. In the future we would like to explore other indicators to study robustness and pursue many-objective formulations for design optimization of water distribution networks with and without constraints.

References

[1] H. Aguirre, A. Oyama, and K. Tanaka. Adaptive ε-sampling and ε-hood for evolutionary many-objective optimization. *Lecture Notes in Computer Science*, 7811:322–336, 2013.

[2] H. Aguirre, Y. Yazawa, A. Oyama, and K. Tanaka. Extending AεEεH from many-objective to multi-objective optimization. *Lecture Notes in Computer Science*, 8886:239–250, 2014.

[3] H. Aguirre, S. Zapotecas, A. Liefooghe, S. Verel, and K. Tanaka. Approaches for Many-objective Optimization: Analysis and Comparison on MNK-landscapes. *Lecture Notes in Computer Science*, 9554, 2015.

[4] R. Armas, H. Aguirre, S. Zapotecas-Martínez, and K. Tanaka, Traffic Signal Optimization: Minimizing Travel Time and Fuel Consumption. *Lecture Notes in Computer Science*, 9554, 2015.

[5] K. Formiga, F. Chaudhry, P. Cheung, and L. Reis. Optimal Design of Water Distribution System by Multiobjective Evolutionary Methods. *Lecture Notes Computer Science*, 2632:677-691, 2003.

[6] M. Laumanns, L. Thiele, K. Deb, and E. Zitzler. Combining convergence and diversity in evolutionary multiobjective optimization. *Evolutionary Computation*, 10(3):263–282, 2002.

[7] E. Reehuis. Multiobjective Robust Optimization of Water Distribution Networks. Master's Thesis Leiden University, Netherlands, 2010.

[8] E. Reehuis, J. Kruisselbrink, A. Deutz, T. Back, and M. Emmerich. Multiobjective Optimization of Water Distribution Networks Using SMS-EMOA. *Proceedings of the International Conference on Evolutionary and Deterministic Methods for Design, Optimization and Control with Applications to Industrial and Societal Problems*, pages 269–279, 2011.

[9] L. Rossman. EPANET 2 Users Manual. Technical Report EPA/600/R-00/057, Water Supply and Water Resources Division, National Risk Management Research Laboratory, U.S. EPA, Cincinnati, OH, USA, 2000.