A Real Case of Multi-Period Water Distribution Network Design solved by a Hybrid Simulated Annealing

Carlos Bermudez[1], Hugo Alfonso[1], Gabriela Minetti[1], and Carolina Salto[1,2]

[1] Facultad de Ingeniería, Universidad Nacional de La Pampa
[2] CONICET, Argentina
bermudezc,alfonsoh,minettig,saltoc@ing.unlpam.edu.ar

Abstract. In this work, we propose an optimization solver based on a hybrid Simulated Annealing (HSA) to optimize the water distribution network design. The problem formulation includes multi-period restrictions with time-varying demand patterns. The HSA search process is affected by the Markov Chain Length (MCL), making modifications in the network design. For that reason, we studied the HSA behaviour by considering static and dynamic methods to compute the MCL. We test the algorithms by using networks reported in the state-of-the-art and also a real and new median size network that arises from a regional necessity. The experimentation suggests the use of a dynamic method, which exhibits the balance between efficacy and efficiency.

Keywords: Water Distribution Network Design, Optimization, Simulated Annealing, Markov chain length

1 Introduction

The water distribution network design represents a major challenge because the distribution system has many components to be considered (pipes of various sizes for carrying water, valves for controlling the flow, service connections to the individual homes, and distribution reservoirs for storing the water to be fed into the distribution pipes). The solution concerning the layout, design, and operation of the network of pipes should result from efficient planning and management procedures. This problem is known as Water Distribution Network Design (WDND) and requires handling a big number of variables and constraints, in consequence, it is classified as NP-hard [1]. The water distribution network optimization aims to find the optimal pipe diameters in the network for a given layout and demand requirements. The optimal pipe sizes are selected in the final network, by satisfying the conservation of mass and energy and also the constraints.

Currently, the WDND formulations include the extension to a multi-period setting, i.e., time-varying demand patterns, which are more realistic and complex problem formulations. Some works expressed the design problem as a multi-objective optimization problem and applied a multi-objective evolutionary algorithm [2]. A genetic algorithm was developed to solve six small networks [3], which considered the velocity constraint on the water flowing through the pipes. In [4] also regarded this constraint, but the authors used mathematical programming on bigger, closer-to-reality networks. Other metaheuristics were used to tackle more complex WDND formulations [5–7]. In this line, an Iterative Local Search [8] (ILS) considered that every demand node has 24 hrs water demand pattern and included a new constraint related to the limit on the maximal velocity of water through the pipes.

The size of the cities is constantly growing, and our city, General Pico, is not an exception. Consequently, our community needs to optimize an independent WDND of a new neighborhood of 5 km^2, subject to multi-period demands, hydraulic restrictions, among others. CORPICO is the organization in charge of distributing this essential element and presents a real requirement for an optimization system to determine the best engineering solution in meeting established design criteria while at the same time minimizing capital costs. The objective of our research is to develop an optimal solution based on state-of-the-art optimization techniques, which intends to support the decision making to design, plan, and management of complex water systems.

In a previous work [9], we tackled the WDND considering time-varying demand patterns and the maximum water velocity constraint as formulated in [8]. The optimizer adopted is a hybrid Simulated Annealing (HSA)[10], which optimizes the pipe diameters by using a local search technique based on GRASP. In [9], we analyzed the HSA behavior by considering the variants for two control parameters: the cooling scheme and the initial temperature. Following this research line, in this work we propose to study the third control parameter, the Markov Chain Length (MCL), and its variants, applying the obtained algorithmic approaches to a real case. The computation of its appropriate length is a difficult task to apply in practice. Consequently, we consider three different strategies to set the MCL to understand their effectiveness in terms of solution quality and execution time. We test the performance of our HSA with the HydroGen network instances [11]. Moreover, we describe and give the features of a real medium size distribution network, which is used to test the capabilities of the HSA model developed.

- The main contributions of this work are the following:
- Analysis of how the MCL influences the HSA performance when solving the WDND problem.
- Resolution of a real distribution network from our community.
- Comparison of the developed HSA solver with state-of-the-art techniques.

The remainder of this article is structured as follows. Section 2 introduces the problem definition. Section 3 explains our algorithmic proposal, HSA, to solve the WDND optimization problem and the HSA's configurations. Section 4 describes the experimental analysis and the methodology used. Then, sections 5 and 6 present the result analysis of the variants and the comparison with an ILS [8] from the literature, respectively. Finally, Section 7 summarizes our conclusions and sketches out our future work.

2 Multi-Period Water Distribution Network Design

The objective of the WDND problem is to minimize the Total Investment Cost (TIC) in a water distribution network design. The problem can be characterized as: simple-objective, multi-period, and gravity-fed. Two restrictions are considered: the limit of water speed in each pipe and the demand pattern that varies in time. The network can be modeled by a connected graph, which is described by a set of nodes $N = \{n_1, n_2, ...\}$, a set of pipes $P = \{p_1, p_2, ...\}$, a set of loops $L = \{l_1, l_2, ...\}$, and a set of commercially available pipe types $T = \{t_1, t_2, ...\}$. The TIC value is obtained by the formula shown in Equation 1, where IC_t is the cost of a pipe p of type t, L_p is the length of the tube, and $x_{p,t}$ is the binary decision variable that determines whether the tube p is of type t or not. The objective function is limited by: physical laws of mass and energy conservation, minimum pressure demand in the nodes, and the maximum speed in the pipes, for each time $\tau \in \mathcal{T}$.

$$\min TIC = \sum_{p \in P} \sum_{t \in T} L_p IC_t x_{p,t} \tag{1}$$

2.1 A Real WDND problem

The principal motivation of this research is to get involved in solving community problems, in particular the water distribution. To provide some context, the water access problem in the province of La Pampa is a priority treatment for being a scarce natural resource. CORPICO[3], the supplier of this essential service in General Pico, has to design an independent drinking water distribution network for a new neighborhood of 5 km^2, minimizing the network cost through the proper selection of the pipe dimensions according to consumption and the physical laws of this type of problem. The water distribution network should cover an area called "Barrio Quintas Sur" identified as Zone 2 or Z2 (shaded region of Fig 1). Initially, the network has an extension of 1.65 km^2 foreseeing for the next ten years an adjacent extension of 3.4 km^2, identified as the zones Z1, Z3, and Z4. The network designed for Z2 is independent but requires taking into consideration the demand of the other zones to become an extra supply network or to receive water from them (bypass).

3 HSA for the Multi-Period WDND Problem

In this work, we continue with the study of the HSA as WDND solver, initiated in [9]. In particular, HSA consists of adapting and hybridizing the SA algorithm [12] to solve the Multi-Period WDND optimization problem. The first design issue is fixed regarding the problem to be solved. For this reason, the

[3] CORPICO is the Regional Cooperative for Electricity, Works and other Services in the city of General Pico, province of La Pampa, Argentina.

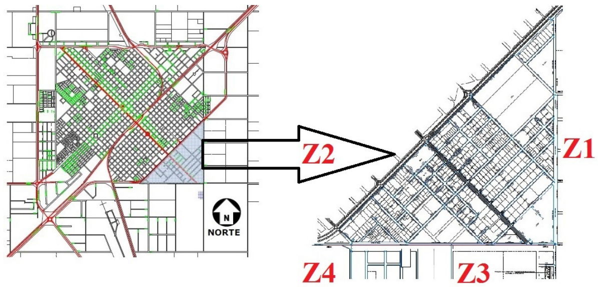


Fig. 1. Map laying water networks in General Pico.

Algorithm 1 HSA to solve the WDND Optimization Problem

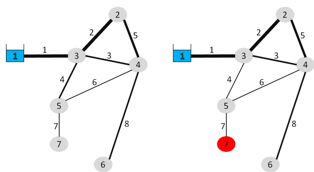
```
k = 0;
initTemp(T)
initialize(S_0);
TIC_0 = evaluate(S_0);
repeat
   repeat
      k = k + 1;
      S_1 = MP-GRASP_LS(S_0);
      TIC_1=evaluate(S_1);
      if TIC_1 < TIC_0 then
         S_0 = S_1; TIC_0 = TIC_1
      end if
      S_2 = perturbation_operator(S_0);
      TIC_2 = evaluate(S_2);
      if (TIC_2 < TIC_0) or (exp^((TIC_2 - TIC_0)/T) > random(0,1)) then
         S_0 = S_2; TIC_0 = TIC_2
      end if
   until (k mod MCL == 0)
   update(T);
until stop criterion is met
return S_0;
```

search space, solution codification, cost (evaluation) function, perturbation operator, and local search are directly related to the WDND problem. Meanwhile, the second design issue contemplates the algorithmic search strategy and includes the definition of the following control parameters: temperature, their annealing schedule, and the Markov chain length [13]. In the previous work, we analyzed different configurations of the two first HSA control parameters. In this work, we investigate the third one and the way to calculate the Markov chain length. The next sections describe the HSA algorithm and Algorithm 1 shows the HSA pseudocode.

3.1 HSA design issue based on problem definition

A solution for the WDND problem is a network, as shown in Figure 2 (a) and (b), which is represented by a vector. Each vector element is the diameter selected for the pipe it represents. Table 1 shows the vectors that correspond to the candidate

ISBN 978-987-4417-90-9



(a) (b)

Pipe ID	Length (m)
1	31
2	20
3	35
4	37
5	24
6	50
7	12
8	65

(c)

Diam. (mm)	Roughness unitless	Cost
60	130	22
80	130	35
100	130	38
150	130	50
200	130	61

(d)

Fig. 2. Different solutions or network designs. (a) Solution 1; (b) Solution 2; (c) Pipe lengths; (d) Available pipe types with their corresponding costs.

Table 1. Different solutions or network designs in vector representation.

Solution	Pipe ID	1	2	3	4	5	6	7	8	Feasibility
	Length (m)	**31**	**20**	**35**	**37**	**24**	**50**	**12**	**65**	**TIC**
1	diam. (mm)	150	150	80	80	100	60	60	80	feasible
	cost	1550	1000	1225	1295	912	1100	264	2275	9621
2	diam. (mm)	150	150	80	60	100	60	60	80	infeasible
	cost	1550	1000	1225	814	912	1100	264	2275	9140

solutions in Figure 2(a) and (b). The TIC for each solution is calculated by the Equation 1, using the input data from tables (c) and (d) of Figure 2. The first solution is hydraulically feasible (satisfying all constraints mentioned in Section 2) and the second one is infeasible (violating the minimum pressure constraint in node 7). HSA uses the EPANET 2.0 toolkit [14] to solve the hydraulic equations, since this hydraulic solver is applied in most existing works.

Following with the HSA search components related to this problem, HSA generates a feasible initial solution S_0 (line 3) applying both `HighCost` and `Lowcost` mechanisms proposed in [8]. Furthermore, the HSA is hybridized with MP-GRASP local search [8], which intensifies the search into the current region of the solution space (line 8). Then a greedy selection mechanism is performed (lines 10-12). The last search component associated with the WDND problem is the perturbation operator, used to obtain a feasible neighbor (line 13) and explore other areas of the search space. It randomly changes some pipe diameters.

3.2 HSA design issues specific to the search

The algorithmic search strategy compares the solution S_2 generated by the perturbation operator with the current one, S_0. If S_2 is worse than S_0, S_2 can be accepted under the Boltzmann probability (line 15, second condition). In this way, at high temperatures (T) the exploration of the search space is strengthened. In contrast, at low temperatures the algorithm only exploits a promising region of the solution space, intensifying the search. In order to update T, a cooling schedule is used (line 19) and it is applied after a certain number of iterations (k) given by the Markov chain length (line 18). Finally, HSA ends the search when the total evaluation number is reached or $T = 0$.

ISBN 978-987-4417-90-9

The choice of the right initial temperature has an important role in the HSA performance to find good solutions. A classical and intuitive method is used in this work that is introduced by Kirkpatrick et al. [12]. The scheme to control the cooling process is also crucial, so that the system gradually cools from a higher temperature, ultimately freezing to a global minimum state. Many attempts have been made to derive or suggest good schedules [13], being the most known cooling processes: the proportional [12], exponential [12], and logarithmic [15] schemes. A fourth cooling processes, named Random scheme, was successfully introduced and tested in order to solve this problem in [10].

The Markov chain length is the number of required transitions (moves) to reach the equilibrium state at each temperature, T_k . This number can be either static or adaptive. In the first case, the number of movements is calculated before the search starts. The static approach, named as MCLs, assumes that each temperature T_k is held constant for a sufficient and fixed number of iterations. In this work, each T_k is held constant for 30 iterations, a common number used in the scientific community.

For the adaptive cases, the Markov chain length depends on the characteristics of the search. For instance, Cardoso et al. [16] consider that the equilibrium state is not necessarily attained at each level of the temperature. Consequently, the cooling schedule is applied as soon as an improved candidate (neighbor) solution is generated. In this way, the computational effort can be drastically reduced without compromising the solution quality. This approach is referred as MCLa1. Another adaptive approach is proposed by Ali et al. [17], named as MCLa2, which uses both the worst and the best solutions found in the Markov chain (inner loop) to compute the next MCL. This strategy allows to increase the number of function evaluations at a given temperature if the difference between the worst and the best solutions increases, but if an improved solution is found the MCL remains unchanged.

4 Experiments with HSA's Variants

In our experiments, we use the best HSA's variant found in [9] to solve the multi-period WDND problem, named HSA$_{Rand100}$, which uses the random cooling scheme and 100 as initial temperature. This variant, renamed HSA in what follows, is executed under the three approaches to compute the MCL, as a consequence, three new HSA configurations arise. The stop condition of these algorithmic approaches is to reach 1,500,000 evaluations of the objective function. We performed 30 independent runs of each instance because of the stochastic nature of the algorithms, in order to gather meaningful experimental data and apply statistical confidence metrics to validate our results and conclusions. Before performing the statistical tests, we first checked whether the data followed a normal distribution by applying the Shapiro-Wilks test. Where the data was distributed normally, we later applied an ANOVA test. Otherwise, we used the Kruskal–Wallis (KW) test. This statistical study allows us to assess whether or

Table 2. Pipe types and their corresponding costs for the GP-Z2-2020 network.

Number	Diam. (mm)	Roughness	Cost	Number	Diam. (mm)	Roughness	Cost
1	63	110	2.85	5	315	110	69.10
2	90	110	5.90	6	400	110	110.89
3	110	110	8.79	7	450	110	140.15
4	125	110	11.00	8	630	110	273.28

not there were meaningful differences between the compared algorithms with $\alpha = 0.01$.

HydroGen Networks. The HydroGen networks [11] arise from 10 different distribution networks, named as HG-MP-i with $i \in \{1, 10\}$. The demand nodes are divided into five categories (domestic, industrial, energy, public services, and commercial demand nodes), each one with a corresponding base load and demand pattern[4]. In this way, each HG-MP-i network consists of five different instances, totaling 50 instances. The combinations of 16 pipe types and the number of pipes of each instance determines the network complexity. In this way, four different categories are obtained, to know HG-MP1-3, HG-MP4-6, HG-MP7-9, and HG-MP10, regarding the number of pipes of each instance.

GP-Z2-2020: a Real Network. The GP-Z2-2020 network, which arises from CORPICO's requirements, is composed of 222 domestic demand nodes and only one water reservoir. Moreover, this zone is connected with three other ones by means of some peripherical nodes which have different demand patterns, as explained in Section 2.1. Table 2 summarizes available pipe diameters, their corresponding roughness, and their unit costs (expressed in US dollars). The area is residential with a demand according to the current distribution of the customers in 584 plots, but considering a development pattern over a timespan of 30 year. The daily pattern demand corresponds to the summer period (based on the model demand of historical records) having a maximum resolution of one hour. The total number of possible combinations of design for a set of 8 commercial pipe types and 282 pipes is 8^{282} which is difficult to test them; this shows the importance of optimization.

5 Analysis of the HydroGen Instance Results

In this section, we summarize and analyze the results of using the three new HSA's configurations to solve the 50 Hydrogen instances grouped by their corresponding distribution network. First, we analyze the behavior of these configurations by considering the results shown in the Table 3, taking the different MCL approaches into account. The columns 2-4 show the average of the best TIC values found by the three HSA's configurations for the 50 instances grouped by their corresponding distribution network. The minimal average TIC values found by each group are boldfaced. Last column summarizes the p-values obtained by the KW test. We follow with the best known TIC values found by

[4] The base loads can be found in the EPANET input files of the instances

ISBN 978-987-4417-90-9

Table 3. Averages of the best TIC values found by each HSA's configurations.

Network	MCLs	MCLa1	MCLa2	KW
HG-MP-1	**335723**	336984	337809	0,87
HG-MP-2	298430	298652	**297823**	0,85
HG-MP-3	**384210**	384665	384687	0,92
HG-MP-4	683780	**682361**	685424	0,80
HG-MP-5	719569	717268	**711906**	0,40
HG-MP-6	739923	**735738**	738108	0,93
HG-MP-7	807367	**792435**	801907	0,27
HG-MP-8	843657	841653	**841240**	0,13
HG-MP-9	823783	817325	**816485**	0,00
HG-MP-10	787399	**771461**	784385	0,08

Table 4. The best known TIC values found by our proposals and ILS.

Network	MCLs	MLCa1	MLCa2	ILS
HG-MP-1	**298000**	**298000**	**298000**	**298000**
HG-MP-2	245330	245330	245330	**245000**
HG-MP-3	310899	310706	**310493**	318000
HG-MP-4	592048	**590837**	592036	598000
HG-MP-5	**631000**	**631000**	**631000**	**631000**
HG-MP-6	617821	**609752**	614917	618000
HG-MP-7	648372	644568	**639932**	653000
HG-MP-8	795996	792436	**790037**	807000
HG-MP-9	716944	715863	**712450**	725000
HG-MP-10	730916	**712847**	727818	724000

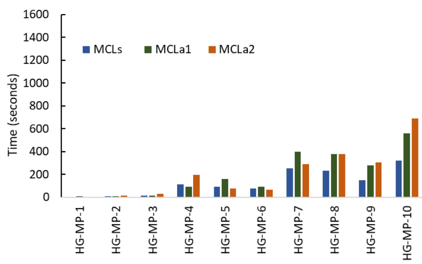


Fig. 3. Minimum execution times consumed by each HSA's configuration.

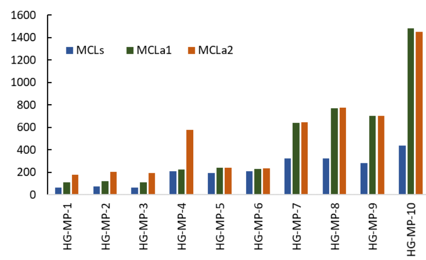


Fig. 4. Total execution times consumed by each HSA's configuration.

our proposals shown in Table 4 (the minimal values are boldfaced). The performance of the HSA's configurations are compared with the ILS proposed in [8]. This metaheuristic is chosen from literature since its authors also used the HydroGen instances to test it. To ensure a fair comparison, both algorithms use the same stop criterion that is set in 1,500,000 Epanet calls. Finally, we analyze the computational effort for each HSA proposed, considering the time to find the best solution (see Figure 3) and the total execution time (see Figure 4) of the search measured in seconds, grouped by distribution network

Analyzing the results from the quality point of view for 90% of the instances, the three approaches behave statistically similar (p-value $> \alpha$). The HSA with the adaptive options find the best TIC value in 80% of the cases. For the HG-MP-9 network, the HSA's behavior is significantly different, being MCLa2 the configuration for HSA that allows to find the lowest TICs values.

Focusing on computational effort analysis, for all configurations the HSA run times grow as the instance complexity increases. HSA with MCLa2 is the slowest configuration, whereas HSA with MCLs reduce significantly the total execution time (approximately 30-50%). However when the best time values to reach the best solution are analyzed, these differences are noticeably narrowed, minimizing the gap between these three approaches.

These results clearly state that the MCLa1 configuration exhibits a good tradeoff between the solution quality and the required execution time. This

ISBN 978-987-4417-90-9

Table 5. Results of the GP-Z2-2020 Network.

Metrics	MCLs	MCLa1	MCLa2	ILS
Minimal TIC	366684.00	358717.00	**347596.00**	355756.00
Average TIC	420949.53	408075.93	**401570.20**	429948.97
Average Best Time	**171.78**	564.26	555,44	169.55
Average Total Time	**186.46**	627.97	615.52	183.83

adaptive approach interrupts the Markov chain when the candidate solution is better than the current one, reducing the computational effort in comparison with MCLa2. Furthermore, in comparison with the literature our proposals are better than ILS in 80% of the networks and equal in the remaining ones.

6 Analysis of the GP-Z2-2020 Network Results

This section is devoted to the analysis of the results of the HSA when solving the real instance (GP-Z2-2020 network). Table 5 presents the results of the three HSA's configurations and the corresponding to the ILS algorithm [8] for different metrics: best and average TIC values (rows 1 and 2), the average time to find the best solution (row 3), and the average total time (row 4), expressed in seconds. The minimal values found by each metric are boldfaced.

HSA using MCLa2 finds minimal TIC values. This is an expected result since similar observations were made in the previous analysis taking into account networks of comparable complexity. An analogous situation is observed regarding the execution times. Furthermore, the three HSA solvers find better average TIC values than ILS, being HSA with MCLs the option that consumes similar execution times than ILS.

7 Conclusions

In this article, we analyze the influence of a relevant control parameter of the HSA solver proposed to solve the WDND optimization problem. The study includes a static (MCLs) and two adaptive (MCLa1 and MCLa2) configurations. We test the HSA's performance with three MCL configurations by using 50 Hydrogen networks. Furthermore, we introduced a new WDND instance based on a real case, which was solved successfully by this solver.

The analysis of the results obtained in the experimentation allows us to conclude the following. The adaptive configurations find better solutions than the static one by adapting the MCL to the search context. These adaptations require an extra computational effort to calculate the MCL during the search. The HSA with MCLa1 presents a good tradeoff between solution quality and effort since this adaptive method interrupts the Markov chain if a better solution than the current one is found. When our proposals are contrasted against ILS [8], the HSA's configurations outperform or equal ILS in every WDND network.

A challenging extension of this work will be to implement and test our HSA using big-data distributed frameworks to deal with larger dimension WDND networks.

Acknowledgments. The authors acknowledge the support of Universidad Nacional de La Pampa (Project FI-CD-151/15 and POIRe-03-2019) and the Incentive Program from MINCyT. The last author is also funded by CONICET.

References

1. D. Yates, A. Templeman, and T. Boffey, "The computational complexity of the problem of determining least capital cost designs for water supply networks," *Engineering Optimization*, vol. 7, no. 2, pp. 143–155, 1984.
2. R. Farmani, G. A. Walters, and D. A. Savic, "Trade-off between total cost and reliability for anytown water distribution network," *Journal of Water Resources Planning and Management*, vol. 131, no. 3, pp. 161–171, 2005.
3. I. Gupta, A. Gupta, and P. Khanna, "Genetic algorithm for optimization of water distribution systems," *Environmental Modelling & Software*, vol. 14, no. 5, pp. 437–446, 1999.
4. C. Bragalli, C. D'Ambrosio, J. Lee, A. Lodi, and P. Toth, "On the optimal design of water distribution networks: a practical MINLP approach," *Optimization and Engineering*, vol. 13, no. 2, pp. 219–246, 2012.
5. R. Uma, "Optimal design of water distribution network using differential evolution," *International Journal of Science and Research (IJSR)*, vol. 5, no. 11, pp. 1515–1520, 2016.
6. R. Mansouri and M. Mohamadizadeh, "Optimal design of water distribution system using central force optimization and differential evolution," vol. 7, no. 3. [Online]. Available: http://ijoce.iust.ac.ir/article-1-310-en.html
7. A. Sedki and D. Ouazar, "Hybrid particle swarm optimization and differential evolution for optimal design of water distribution systems," *Adv. Eng. Informatics*, vol. 26, pp. 582–591, 2012.
8. A. De Corte and K. Sörensen, "An iterated local search algorithm for water distribution network design optimization," *Network*, vol. 67, no. 3, pp. 187–198, 2016.
9. C. Bermudez, C. Salto, and G. Minetti, "Designing a multi-period water distribution network with a hybrid simulated annealing," in *XLVIII JAIIO: XX Simposio Argentino de Inteligencia Artificial (ASAI 2019)*, 2019, pp. 39–52.
10. C. Bermudez, G. Minetti, and C. Salto, "SA to optimize the multi-period water distribution network design," in *XXIX Congreso Argentino de Ciencias de la Computación (CACIC 2018)*, 2018, pp. 12–21.
11. A. De Corte and K. Sörensen, "Hydrogen," available online: http://antor.uantwerpen.be/hydrogen (accessed on 27 June 2018).
12. S. Kirkpatrick, C. G. Jr, and M. Vecchi, "Optimization by simulated annealing," *Science*, no. 220, pp. 671–680, 1983.
13. E.-G. Talbi, *Metaheuristics: From Design to Implementation.* Wiley, 2009.
14. L. Rossman, *The EPANET Programmer's Toolkit for Analysis of Water Distribution Systems*, 1999.
15. B. Hajek, "Cooling schedules for optimal annealing," *Mathematics of Operations Research*, vol. 13, no. 2, pp. 311–329, 1988.
16. M. Cardoso, R. Salcedo, and S. de Azevedo, "Nonequilibrium simulated annealing: A faster approach to combinatorial minimization," *Ind. Eng. Chem. Res*, vol. 33, pp. 1908–1918, 1994.
17. M. Ali, A. Törn, and S. Viitanen, "A direct search variant of the simulated annealing algorithm for optimization involving continuous variables," *Computers & Operations Research*, vol. 29, no. 1, pp. 87 – 102, 2002.