

Document downloaded from:

<http://hdl.handle.net/10251/56182>

This paper must be cited as:

Mora Melia, D.; Iglesias Rey, PL.; Martínez-Solano, FJ.; Fuertes Miquel, VS. (2013). Design of water distribution networks using a pseudo-genetic algorithm and sensitivity of genetic operators. *Water Resources Management*. 27(12):4149-4162. doi:10.1007/s11269-013-0400-6.



The final publication is available at

<http://dx.doi.org/10.1007/s11269-013-0400-6>

Copyright Springer Verlag (Germany)

Additional Information

DESIGN OF WATER DISTRIBUTION NETWORKS USING A PSEUDO-GENETIC ALGORITHM AND SENSITIVITY OF GENETIC OPERATORS

D. Mora-Melia¹, P.L. Iglesias-Rey¹, F. J Martinez-Solano¹, V.S. Fuertes-Miquel¹

¹Departamento de Ingeniería Hidráulica y Medio Ambiente, Universitat Politècnica de València, España., E-mail: damome@doctor.upv.es; piglesia@upv.es; jmsolano@upv.es; vfuertes@upv.es

Abstract

Genetic algorithms (GA) are optimization techniques that are widely used in the design of water distribution networks. One of the main disadvantages of GA is positional bias, which degrades the quality of the solution. In this study, a modified pseudo-genetic algorithm (PGA) is presented. In a PGA, the coding of chromosomes is performed using integer coding; in a traditional GA, binary coding is utilized. Each decision variable is represented by only one gene. This variation entails a series of special characteristics in the definition of mutation and crossover operations.

Some benchmark networks have been used to test the suitability of a PGA for designing water distribution networks. More than 50,000 simulations were conducted with different sets of parameters. A statistical analysis of the obtained solutions was also performed. Through this analysis, more suitable values of mutation and crossover probabilities were discovered for each case. The results demonstrate the validity of the method.

Optimum solutions are not guaranteed in any heuristic method. Hence, the concept of a "good solution" is introduced. A good solution is a design solution that does not substantially exceed the optimal solution that is obtained from the simulations. This concept may be useful when the computational cost is critical. The main conclusion derived from this study is that a proper combination of population and crossover and mutation probabilities leads to a high probability that good solutions will be obtained.

Keywords

Genetic algorithms, design, water networks, optimization

Introduction

Traditionally, water distribution network design, upgrading, or rehabilitation is based on engineering judgment. However, in the last three decades, a significant amount of research has focused on the optimal design of water distribution networks. Initially, researchers used linear programming to optimize the design of a pipe network (Alperovits and Shamir 1977). Subsequent studies applied nonlinear programming to network design problems. Some examples include the optimization of looped pipe networks (Su et al. 1987) or the development of models that simulate pumps, tanks and multiple loading cases (Lansey and Mays 1989).

The application of heuristic optimization techniques facilitates searching beyond these local minimums, which generally extends the field search and, consequently, the capacity to obtain better solutions. Evolutionary algorithms are a set of optimization techniques that are based on natural principles of evolution. Some examples of these techniques include genetic algorithms (GA), particle swarm optimization (PSO), shuffled frog-leaping algorithms (SFLA), and memetic algorithms. Although some investigators have compared these techniques (Reca et al. 2008), no technique has been declared superior to all others.

Due to their versatility, evolutionary algorithms can be applied to all types of optimization processes. In the field of hydraulic engineering, some evolutionary optimization techniques, such as memetic algorithms (Baños et al. 2007), SFLA (Eusuff and Lansey 2003) or harmony search (Geem 2006), have been successfully applied to the design of water distribution networks. The flexibility of these methods has facilitated the development of hybrid models that combine features of different methodologies. An example of a hybrid model applied to the design of water networks is the GALP method (Cisty 2010), which combines genetic algorithms and linear programming.

Similarly, the heterogeneity of evolutionary algorithms can be used for the optimization of various problems related to water distribution networks. Some researchers have successfully applied genetic algorithms to water network

rehabilitation (Halhal et al. 1997), calibration of water distribution models (Balla and Lingereddy 2000), complex supply systems (Louati et al. 2011, Chung and Lansey 2008), and to control hydraulic pressure in a water distribution network (Nazif et al. 2010).

In this study, we use a modified genetic algorithm. In general, GA aim to find an optimal solution for a given design problem from the evolution of a random initial solution. This searching method, which is based on Darwin's theory of evolution, functions identically to the evolution of a population that experiences similar random actions to which they react in a biological evolution (mutations and genetic recombination). Based on established criteria, the individuals who are best adapted survive, and the individuals who are less adapted perish.

For a given network layout and demand, most of these methods consider the minimization of the cost of a pipe network as the objective. In the field of water engineering, previous studies by Goldberg and Kuo (1987), Savic and Walters (1997), Fujiwara and Khang (1990), Jin et al. (2007) and Tsai et al. (2008) reflect the importance of these algorithms in the optimal design of water distribution networks.

In this study, a method is developed for the optimal design of water distribution networks based on a GA and applied to different benchmark networks. The aim of the study is to minimize the necessary costs of investment for the implementation of a system, beginning with the topological layout and the demands and requirements of pressure in the nodes. The study also introduces the optimization of the different genetic operators, which influence the minimum solutions, such as the probabilities of mutation and crossover and the population size in which the algorithm will operate.

Methodology

All evolutionary methodologies share some principles: a data structure that stores the characteristics of a solution, certain operations that modify these solutions, an objective function, and a selection mechanism to ensure that only best solutions survive.

Traditionally, GA have included methods adapted for problems formulated in binary variables, but not suitable for other searching methods. Previous studies describe numerical encoding in GA (Reca and Martinez 2006; Vairavamoorthy and Ali 2000; Prasad and Park 2004). A pseudo-genetic algorithm (PGA), which is based on an integer coding of the solution, is introduced in the present study. A brief description of PGA follows.

In the PGA, every decision variable can store different values that are represented with alphanumerical variables. Each gene is identified with a decision variable through integer coding (Figure 1), which did not occur in the classic GA. PGA includes elements such as pipes, pumps, tanks, and valves.

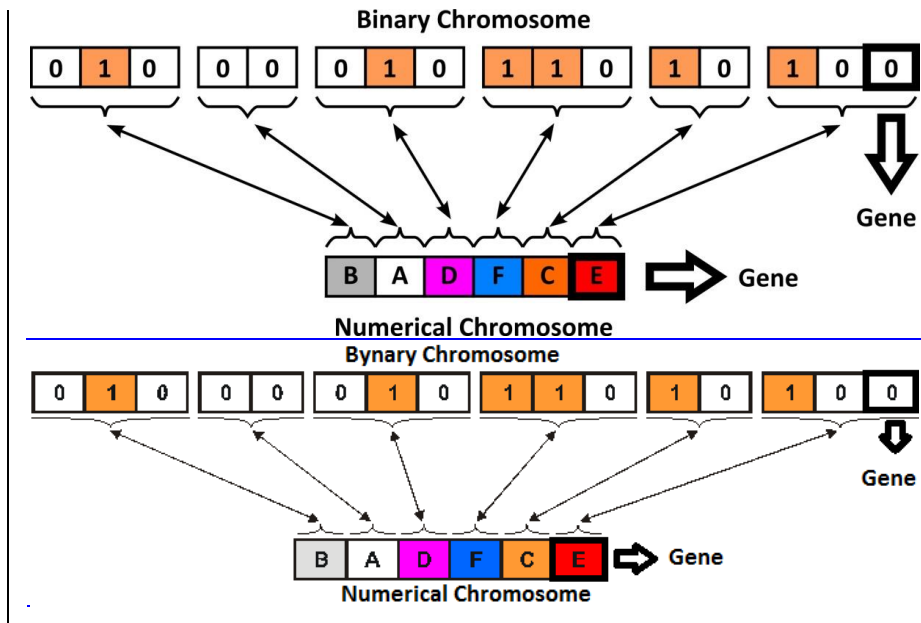


Fig 1. Binary and alphanumerical representation

To solve the optimization problem, it is necessary to establish a discrete set of solutions. A solution is represented by a chromosome. This is comprised of a series of genes. Each gene identifies the discrete value of a variable. In the PGA, a generic chromosome X^i consists of as many genes as decision variables retain (N_{DV}). This chromosome X^i is defined as a vector of numerical values.

$$X^i = \{X_1^i, X_2^i, \dots, X_{N_{DV}}^i\} \quad (1)$$

The aptitude is the ability of a chromosome to survive during the reproduction process. The aptitude of a generic chromosome X^i is identified by

the value of the objective function for the encoded solution. In the case of the proposed PGA, the N_{VD} decision variables are N_p pumps and N_D diameters. The objective function is defined as follows:

$$F(X^i) = \frac{\sum_{j=1}^{N_p} F_{e,j} H_{D,j}(X^i) Q_{D,j}(X^i) + F_a \sum_{j=1}^{N_D} C_j(X_j^i) \cdot L_j + \lambda_1 \cdot \sum_{s=1}^{N_S} \sum_{k=1}^{N_R} \delta_{k,s} \cdot (H_{\min,k} - H_{k,s}) + \lambda_2 \cdot \sum_{s=1}^{N_S} \sum_{k=1}^{N_R} \delta_{k,s} \cdot (V_{\max,k} - V_{k,s}) + \lambda_3 \cdot \sum_{s=1}^{N_S} \sum_{k=1}^{N_R} \delta_{k,s} \cdot (V_{\min,k} - V_{k,s})}{(2)}$$

Con formato: Centrado

In the first term of (1), $H_{D,j}(X^i)$ and $Q_{D,j}(X^i)$ represent the pressure head and the flow of each pump, respectively; $F_{e,j}$ is the energy cost in each of these pumps. The second term represents the capital cost of the pipes, where C_j is the associated unit cost of the decision variable contained in link j of chromosome i , and L_j is the length of pipe j .

There are N_R constraints that must be achieved by the solutions. These restrictions have been included as a penalty in the total cost of the solution. The constraints that must be satisfied are the minimum pressure height in each node ($H_{\min,k}$), maximum ($V_{\max,k}$) and minimum ($V_{\min,k}$) velocity in pipes. These constraints must be verified in N_S scenarios. The analyzed scenarios are the different loading conditions under which the water distribution system is tested.

To compute the penalties several binary variables ($\delta_{k,s}$) are used. These variables adopt the value one if the constraint is not satisfied, and zero otherwise.

The parameters λ_1 , λ_2 and λ_3 represent weight functions that establish penalties for not verifying the restrictions. In our model, pressure and velocity are considered hard constraints. For this reason, the values of λ_1 , λ_2 and λ_3 are large enough (10^7) to reject all solutions that violate the constraints. There is an existing controversy regarding the penalty terms. Although the use of large multipliers often prevents infeasible solutions, optimum solutions are located within the boundary of feasibility. However, infeasible solutions may also contain useful information that can be carried to the next generation (Vairavamoorthy and Ali 2000). Conversely, small variations of numerical values for certain parameters may lead to more expensive or infeasible solutions (Savic and Walters 1997; Reca and Martínez 2006). Therefore, steps should be taken to ensure that the solution will not violate any restrictions, such as by using large multipliers or adding a

multiplier that represents the existence of any type of penalty in the system (Iglesias et al. 2012).

The method of the proposed PGA tests the evolution of a random population through a parallelism that is similar to the laws of natural selection, which occurs in a classic GA. This is obtained through three basic processes: reproduction, crossover and mutation. The different stages of the optimization process are described below.

Genetic operators

Genetic operators are specific functions of a PGA that maintain genetic diversity in the population of the algorithm. The implementation of these genetic operators is different for each optimization model. The following lines define the genetic operators used in this study.

A ranking reproduction method was selected among all existing methods of reproduction (Wang, 1991; Savic and Walters, 1997). This method arranges the individuals of a population in increasing order according to their cost.

The probability assigned to each chromosome will fall between a maximum probability p_{\max} , associated with the smaller cost solution, and a minimum probability, associated with the greater cost solution. Both probabilities are defined as

$$p_{\max} = \frac{\beta}{N_C} \quad p_{\min} = \frac{2-\beta}{N_C} \quad (3)$$

where β is a constant, whose value is between 1.5 and 2, and N_C is the number of chromosomes.

The crossover process consists of randomly select the chromosomes of the intermediate population and modifying the different genes from a certain crossover gene randomly determined.

In a traditional GA, the random selection of a crossover gene can create the fraction of binary code that identifies a possible decision variable. This phenomenon is known as positional bias and degrades the quality of the solutions. One of the fundamental characteristics of a PGA is that the crossover operation does not produce this effect because each gene corresponds to a decision variable. For this reason, the PGA is less sensitive to this operator than a classic GA.

The mutation process is applied to the obtained population after the crossover and reproduction processes. The usefulness of the process is realized in the expansion of search areas within the space of feasible solutions. Every gene in each chromosome can eventually change if a randomly generated number falls under a certain probability.

PGA uses an alphanumeric representation based on Gray coding. One of the main advantages of Gray coding is that it prevents major changes in the decision variable by modifying a single gene. In PGA coding, mutation process consist of increasing or decreasing the selected gene only one step.

In a traditional GA, the mutation operator is considered a secondary priority regarding the use of the crossover probability as the main genetic operator. In practice, mutation rates between 0.1 and 1% are often recommended for binary representations. The crossover process in the PGA generates less alternatives than a classic GA. For this reason, the probability of mutation in the PGA is slightly higher (1 to 10%). The main operator to generate diversity in a PGA is mutation, not crossover. This is the greatest difference between a PGA and a traditional GA.

The mutation is not only a mechanism of generating diversity. The use of high values of mutation can reduce a genetic algorithm to a random search. It is always recommended to use other mechanisms of diversity generation, such as increasing the size of the population or guaranteeing the randomness of the initial population. The main problem with using large populations is that they require more convergence time in the algorithm. Therefore, it is necessary to reach a commitment solution depending on the approach of the problem.

The following sections analyze the capacity of the proposed PGA to determine the minimum design cost of different water distribution networks. The influence of different parameters on the final solution was analyzed. The analysis was divided into two phases. In the first stage, the best combination of the crossover and mutation probabilities was analyzed. In the second phase, the effect of population size on the performance of the algorithm was evaluated.

For each calculation, it was necessary to determine the pressure in the nodes and the flow in the pipes. These calculations were performed with the model

EPANET (Rossman, 2000). Simulations to obtain the parameters of the PGA were performed with a specific application developed by the authors.

Application examples

The evaluation of the proposed model was performed on the water distribution networks of Hanoi and New York. Both systems consist of large water networks with real layouts. As noted in the bibliography, some results were obtained with different design models, which allowed a comparison of the results of the different models.

The Hanoi network (Figure 2) was proposed by Fujiwara and Khang (1990). The network consists of a reservoir, 31 demand nodes and 34 pipes. The objective of the problem is to dimension all pipes of the network, considering that the minimum required pressure head at each node is 30 m.

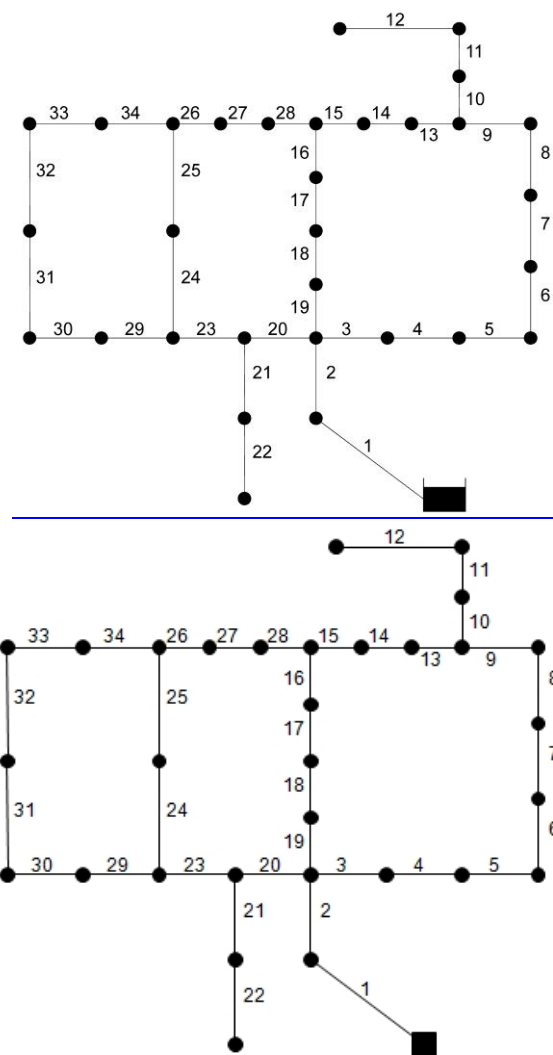


Fig 2. Hanoi network.

The second example (Figure 3) entails an analysis of the New York tunnel network (Goulter and Morgan 1985). The network consists of 20 demand nodes, 21 pipes and a reservoir. The objective of the problem is to add new pipes in parallel to existing ones because the pressure is inadequate for proper operation of the distribution network under current conditions.

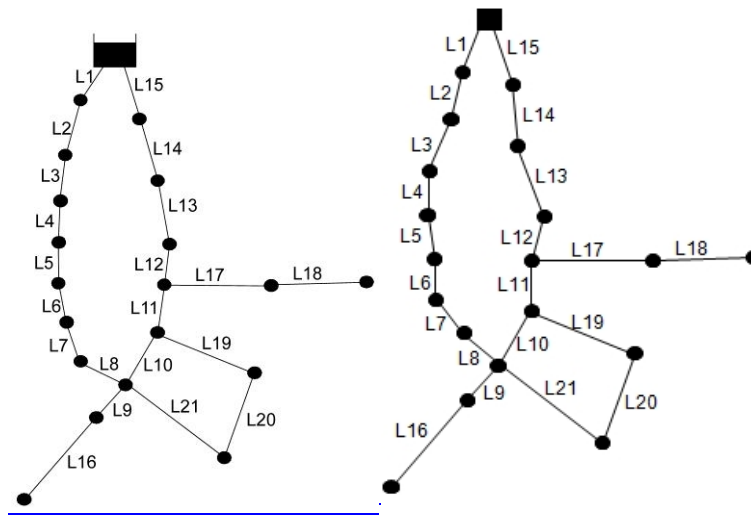


Fig 3. New York tunnels network.

One of the characteristics that contribute to the definition of an optimal solution of the network is the range of diameters employed. The original range from the bibliography in both cases was used (table 1).

Table 1. Original diameters for Hanoi and New York tunnel networks.

| Network | Hanoi tunnel network | New York tunnel network |
|--------------------------------|---|--|
| Available commercial diameters | 304.8; 406.4; 508; 609.6; 762; 1016 (mm) | 36; 48; 60; 72; 84; 96; 108; 120; 132; 144; 156; 168; 180; 192; 204 (inches) |
| Cost | 45.726; 70.4; 98.378; 129.333; 180.748; 278.28 (um/m) | 93.5; 134; 176; 221; 267; 316; 365; 417; 469; 522; 577; 632; 689; 746; 804 (um/feet) |
| Number of potential solutions | $6^{34} = 2.86 \cdot 10^{26}$ | $16^{21} = 1.93 \cdot 10^{25}$ |

Results

The best solution obtained by the APG has a cost of 6.081 million um. This combination of diameters results in all network nodes with pressures above 30. This result coincides with the current lowest cost of design obtained for this network, which satisfies all constraints of minimum pressure in the system. Table 2 displays the results obtained by different researchers, as well as the results obtained by the proposed method.

Table 2. Solutions to the Hanoi network design problem

| Pipe | <i>Bibliography Solutions</i> | | | | <i>Solution PGA</i> | |
|------|-------------------------------|-----------------------|-----------------------|----------------------|---------------------|--------------------|
| | Matías ⁽¹⁾ | Savic1 ⁽²⁾ | Savic2 ⁽³⁾ | Cunha ⁽⁴⁾ | Reca ⁽⁵⁾ | Sol ⁽⁶⁾ |
| 1 | 1016 | 1016 | 1016 | 1016 | 1016 | 1016 |
| 2 | 1016 | 1016 | 1016 | 1016 | 1016 | 1016 |
| 3 | 1016 | 1016 | 1016 | 1016 | 1016 | 1016 |
| 4 | 1016 | 1016 | 1016 | 1016 | 1016 | 1016 |
| 5 | 1016 | 1016 | 1016 | 1016 | 1016 | 1016 |
| 6 | 1016 | 1016 | 1016 | 1016 | 1016 | 1016 |
| 7 | 1016 | 1016 | 1016 | 1016 | 1016 | 1016 |
| 8 | 1016 | 1016 | 1016 | 1016 | 1016 | 1016 |
| 9 | 1016 | 762 | 1016 | 1016 | 1016 | 1016 |
| 10 | 762 | 762 | 762 | 762 | 762 | 762 |
| 11 | 609.6 | 762 | 609.6 | 609.6 | 609.6 | 609.6 |
| 12 | 609.6 | 609.6 | 609.6 | 609.6 | 609.6 | 609.6 |
| 13 | 508 | 406.4 | 508 | 508 | 508 | 508 |
| 14 | 406.4 | 406.4 | 406.4 | 406.4 | 406.4 | 406.4 |
| 15 | 304.8 | 304.8 | 304.8 | 304.8 | 304.8 | 304.8 |
| 16 | 304.8 | 406.4 | 304.8 | 304.8 | 304.8 | 304.8 |
| 17 | 406.4 | 508 | 406.4 | 406.4 | 406.4 | 406.4 |
| 18 | 609.6 | 609.6 | 508 | 508 | 609.6 | 609.6 |
| 19 | 609.6 | 609.6 | 508 | 508 | 508 | 508 |
| 20 | 1016 | 1016 | 1016 | 1016 | 1016 | 1016 |
| 21 | 508 | 508 | 508 | 508 | 508 | 508 |
| 22 | 304.8 | 304.8 | 304.8 | 304.8 | 304.8 | 304.8 |
| 23 | 1016 | 1016 | 1016 | 1016 | 1016 | 1016 |
| 24 | 762 | 762 | 762 | 762 | 762 | 762 |
| 25 | 762 | 762 | 762 | 762 | 762 | 762 |
| 26 | 508 | 508 | 508 | 508 | 508 | 508 |
| 27 | 304.8 | 304.8 | 304.8 | 304.8 | 304.8 | 304.8 |
| 28 | 304.8 | 304.8 | 304.8 | 304.8 | 304.8 | 304.8 |
| 29 | 406.4 | 406.4 | 406.4 | 406.4 | 406.4 | 406.4 |
| 30 | 304.8 | 406.4 | 406.4 | 304.8 | 304.8 | 304.8 |
| 31 | 304.8 | 304.8 | 304.8 | 304.8 | 304.8 | 304.8 |
| 32 | 406.4 | 304.8 | 304.8 | 406.4 | 406.4 | 406.4 |
| 33 | 406.4 | 406.4 | 406.4 | 406.4 | 406.4 | 406.4 |
| 34 | 609.6 | 508 | 508 | 609.6 | 609.6 | 609.6 |
| Cost | 6.093 | 6.187 | 6.073 | 6.056 | 6.081 | 6.081 |

(millions um)

(1) Solution obtained by Matías (2003).

(2) Solution obtained by Savic and Walters (1997).

(3) Solution obtained by Savic and Walters (1997). No pressure restrictions accomplished in EPANET2.

(4) Solution obtained with a heuristic method (Cunha and Sousa, 1999). No pressure restrictions accomplished in EPANET2

(5) Solution obtained by Reça and Martínez (2006).

(6) Best solution obtained with the proposed method.

In the case of the New York tunnel network, the best solution obtained has a cost of 38.642 million um. The diameters of each pipe are listed in Table 3.

A comparison of this solution with solutions obtained by other authors reveals that only one solution in the literature improves the results presented in this paper. Savic and Walters (1997) estimated an alternative solution with a cost of 37.130 million um. Although their solution does not meet the minimum pressure if solved by EPANET2, it complies with those constraints. The answer to this contradiction lies in the value of the Hazen-Williams coefficients using different versions of EPANET. The differences among various versions of EPANET affect flow rates and pressure in the nodes. The results by Savic were computed with the first version of EPANET; the difference in coefficients with respect to the second version prevents the solution from meeting the pressure constraints.

Table 3. Solutions to the New York tunnel design problem.

| <i>Pipe</i> | <i>Savic⁽¹⁾</i> | <i>Goulter & Morgan⁽²⁾</i> | <i>Solution PGA⁽³⁾</i> |
|--------------------|----------------------------|---|-----------------------------------|
| 1 | 0 | 0 | 0 |
| 2 | 0 | 0 | 0 |
| 3 | 0 | 0 | 0 |
| 4 | 0 | 0 | 0 |
| 5 | 0 | 0 | 0 |
| 6 | 0 | 0 | 0 |
| 7 | 108 | 144 | 144 |
| 8 | 0 | 0 | 0 |
| 9 | 0 | 0 | 0 |
| 10 | 0 | 0 | 0 |
| 11 | 0 | 0 | 0 |
| 12 | 0 | 0 | 0 |
| 13 | 0 | 0 | 0 |
| 14 | 0 | 0 | 0 |
| 15 | 0 | 0 | 0 |
| 16 | 96 | 96 | 96 |
| 17 | 96 | 96 | 96 |
| 18 | 84 | 84 | 84 |
| 19 | 72 | 60 | 72 |
| 20 | 0 | 0 | 0 |
| 21 | 72 | 84 | 72 |
| Cost (millions um) | 37.130 | 39.200 | 38.642 |

(1) Solution obtained by Savic and Walters (1997). No pressure restrictions accomplished in EPANET2.

(2) Solution obtained by Goulter and Morgan (1985).

(3) Best solution obtained with the proposed method.

Statistical Analysis

The characteristics of evolutionary algorithms do not guarantee with certainty that the optimal value of the system will be obtained. In addition, the obtained result sometimes contains certain variations. To analyze this randomness, it is necessary to statistically analyze the influence that the different parameters have on the proposed PGA in the solution of the analyzed network, with the purpose of optimizing them to increase the probability of obtaining the minimum. More than 50000 simulations were conducted in this analysis. A computer system in parallel with 24 computers to perform the simulations was used.

Influence of crossover and mutation probability

Previous studies have indicated that the population size must be large enough to ensure diversity in the solutions. This first study addresses fixed populations of 100 individuals, with the aim to optimize the crossover and mutation parameters. Another assessment was made from the best values obtained for both parameters, with the objective of studying the influence of population in the search for the best solution.

A large number of simulations are useful in a statistical analysis to filter the results based on any criteria. Sorting them according to the final value obtained increases the probability of finding a better set of parameters.

A histogram (Figure 4) that incorporates the accumulated probability of the obtained solutions was created to detect the more frequent solutions. Note that the histogram represents the results of all cases analyzed.

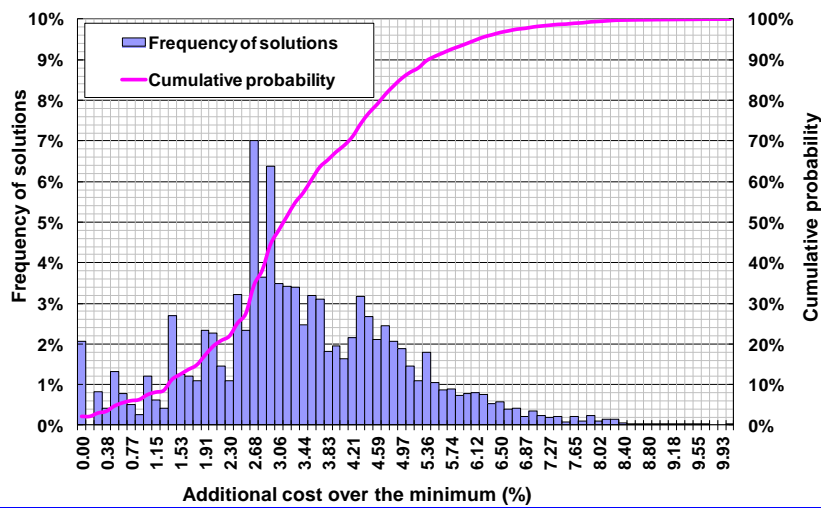
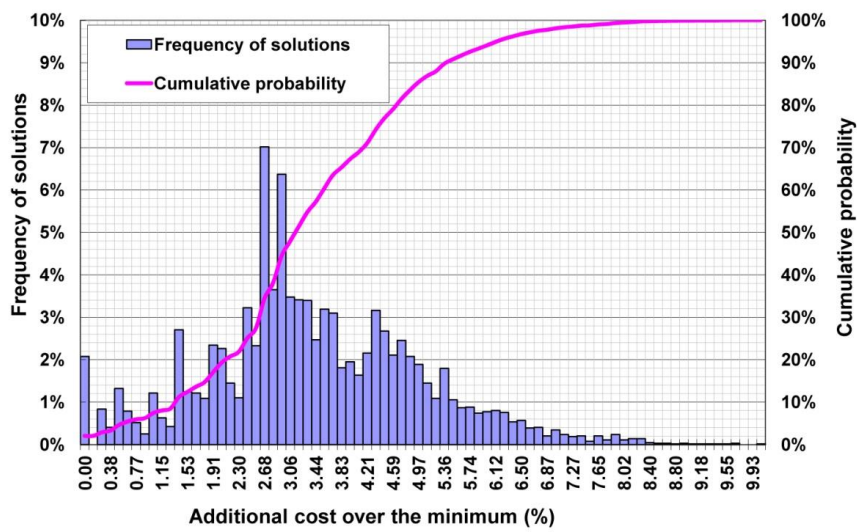


Fig. 4. Frequency and accumulated probability for the Hanoi network solutions.

To determine the influence of mutation and crossover probability, the solution corresponding to 6.081 million monetary units was adopted as optimal design value. The success rate of PGA was defined as the probability of obtaining the optimal design value with every set of parameters. The representation of this success rate is shown in Figure 5. Note that there are combinations of values of the mutation probability and crossover that may never generate an optimal value. The maximum success rate is obtained approximately for a mutation probability of 3–4% and an approximate crossover probability of 90%.

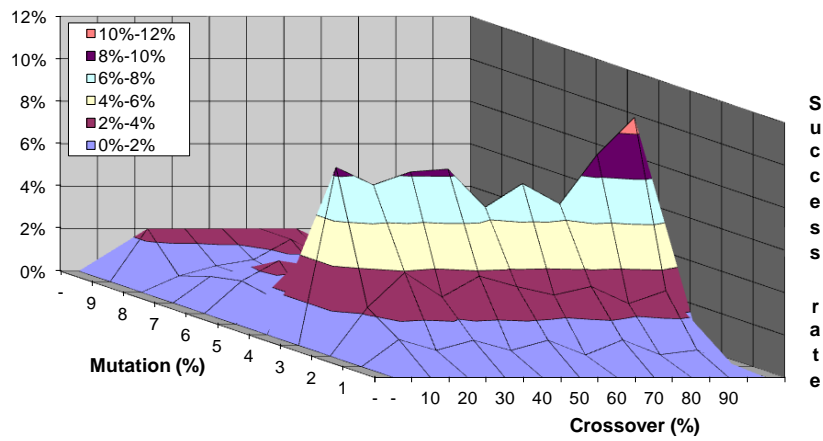
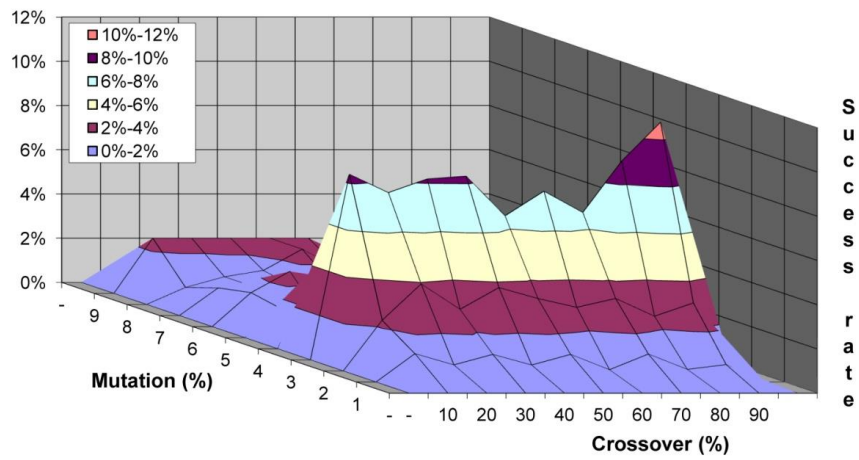


Fig. 5. Probability of obtaining the minimum solution about the mutation and crossover probability of the Hanoi network.

One of the general characteristics of the GA and the proposed PGA is the capacity to obtain not only one single optimal value but also to obtain a set of good solutions for the design problem.

A new concept is introduced. In this study, a "good solution" is defined as a combination of solutions whose cost exceeds the minimum cost until it reaches a maximum cost of 3%. Thus, Figure 6 indicates the probability of obtaining a "good solution" for each combination of mutation and crossover probabilities.

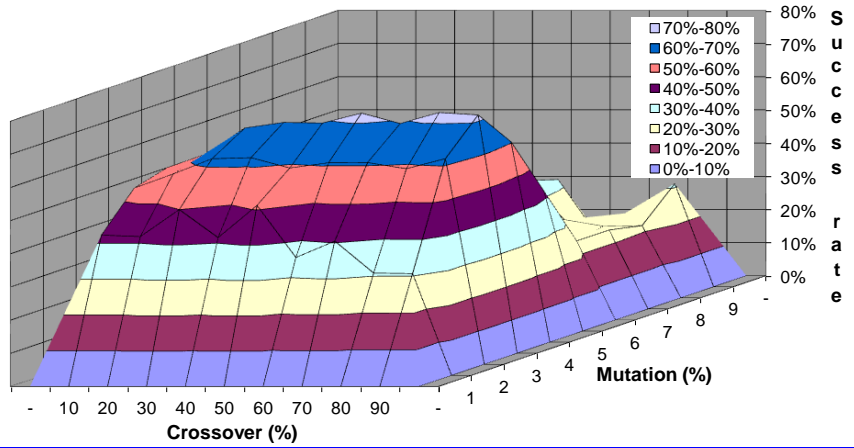
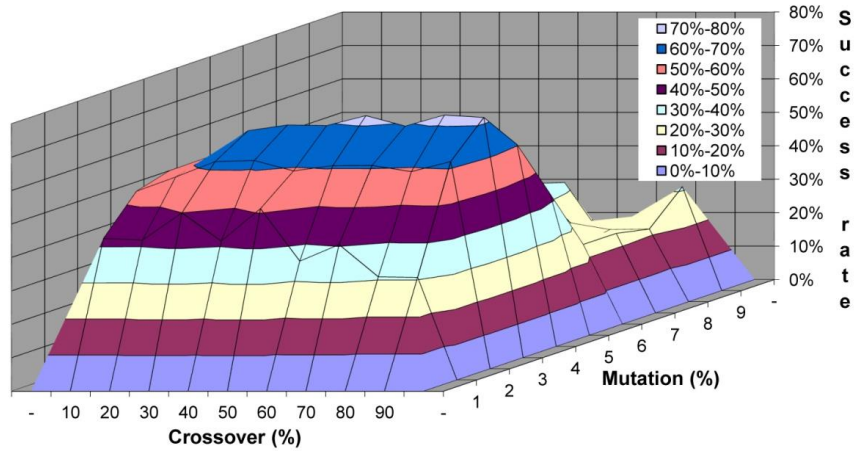


Fig 6. Probability of obtaining a good solution about the mutation and crossover probabilities in the Hanoi network.

Figure 6 lists the capacity of the method to not only obtain minimum values but also to frequently obtain values near the optimal defined value. The statistical analysis of the simulations verifies the robustness of the method. Thus, the effect of the crossover probability becomes smaller for the values of the best adapted mutation probabilities. This result is illustrated not only in the process of obtaining the minimum cost solution but also in the process of obtaining solutions near the optimal solution.

For the Hanoi network, the adjustment of crossover and mutation operators can enhance the probability of obtaining the minimum (6.081 million um) from 2 to 8%. Similarly, a good parameter setting increases the success rate of obtaining a good solution (<6.263 million um) to 68%.

The design of the New York tunnel network is a less challenging problem. The number of possible design solutions is less than the number of possible design solutions in the Hanoi network. In this case, the number of simulations exceeds 8000, which enables 100 repetitions for each of the possible parametric combinations. The results obtained are considerably better than the results obtained for the Hanoi network because the lowest cost solution (38.642 million um) was obtained in approximately 12% of simulations.

The analysis of genetic operators in the New York tunnel network confirmed the results obtained for the Hanoi network. Thus, the mutation operator is a key influence on the quality of the obtained design results. In this case, the optimal value of mutation is approximately 4–5%, whereas qualitatively the optimization process is slightly favored when the probability of crossover exceeds 50%.

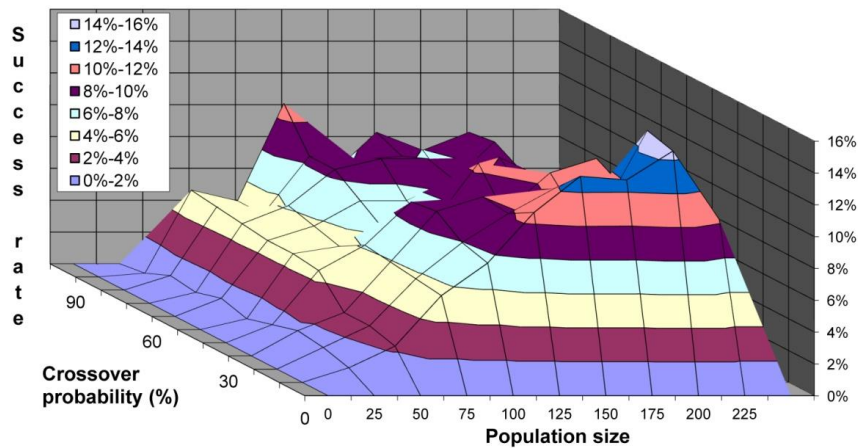
Under these conditions, the probability of obtaining the minimum design solution is approximately 27%. By extending the analysis to obtain good design solutions (<39.801 million um), the success rate increases to approximately 62%. The results show the usefulness of parameter setting because an improper choice of genetic operators drastically reduces the success rate of the optimization process.

Influence of population size

In the previous statistical analysis, the collected data demonstrate that certain combinations of crossover and mutation generate greater rates of success to obtain the optimal solution. The results are better in the analyzed networks for probabilities of mutation between 3 and 5% (a gene by chromosome). Conversely, we deduce from the analysis that the crossover probability does not exhibit a critical influence on the optimization process of the PGA when the population consists of 100 units.

The present analysis considers the crossover and mutation settings proposed in the previous section as optimal parameters of design, aiming the study in the optimization of the initial population. More than 25000 simulations were performed with populations ranging from 25 to 225 individuals.

We adopted the solution that corresponds to a cost of 6,081 thousands of monetary units as the optimal value. With this fixed value, it has been analyzed the probability that the PGA obtains the optimal solution for each combination of the population values and probabilities of crossover. Figure 7 displays the representation of this rate of success in obtaining the minimum.



Con formato: Centrado

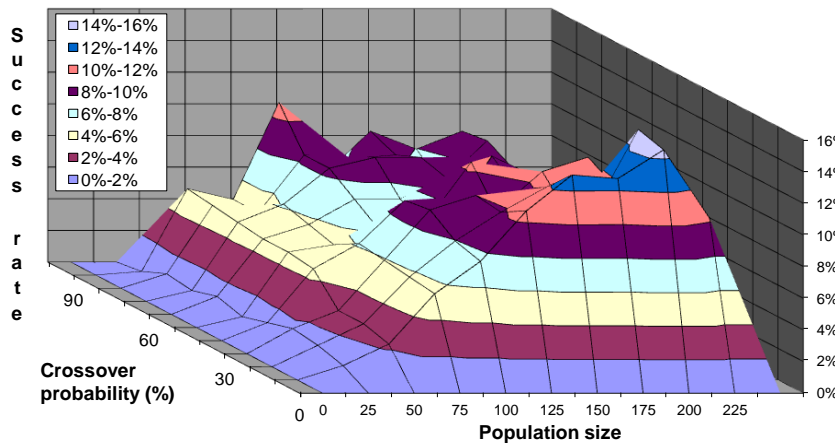


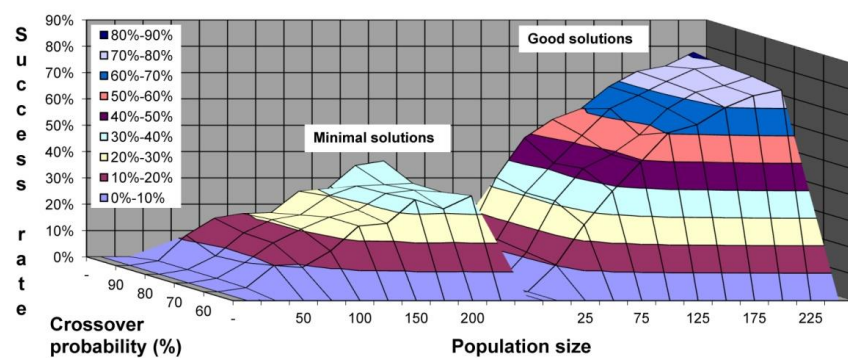
Fig 7. Probability of obtaining a minimal solution about population size, fixed mutation and crossover probability for the Hanoi network.

Con formato: Fuente: 10 pto

Figure 7 illustrates that as the initial population increases, the probability of obtaining the minimum value also increases. This increase becomes stable at a certain point. Initially, the best combination of values is populations of 200 with a probability of crossover of 10%. However, an evaluation will be necessary if this slight improvement is compensated with a reduction in calculation speed that

requires working with greater populations. Similarly, the algorithm is less effective for finding minimums for smaller populations of 50.

Next, the New York tunnel network was analyzed. Figure 8 represents the success rate of minimum/good solutions in terms of population size and crossover probability. The probability of mutation varies between 4 and 5%, which is equivalent to the optimal setting for this network.



Con formato: Centrado

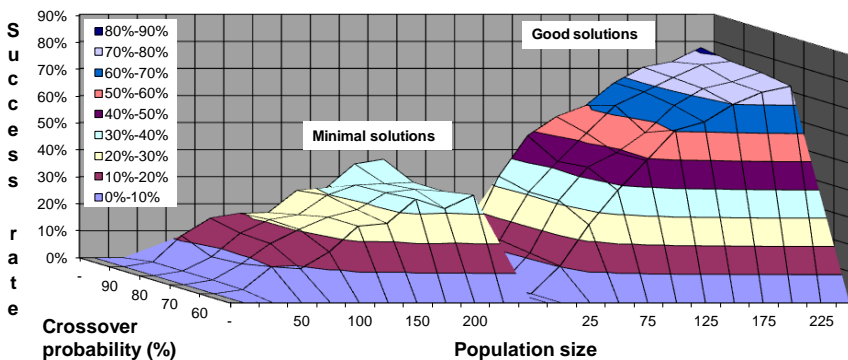


Fig 8. Probability of obtaining a minimal/good solution about population size, fixed mutation and crossover probability in the NY tunnel networks.

Figure 8 shows how the probability of finding better solutions increases with larger population sizes. In the range studied for the New York network, the results improve gradually without obtaining a maximum. Thus, for a population size of 225 individuals, the probability of obtaining the minimum cost solution for the tunnel network in New York is approximately 35–40%, whereas the probability of obtaining a good solution increases to 75–80%.

Conclusions

In this article a PGA method for design of water distribution networks has been presented. The use of a PGA-based method allows obtaining not only the minimum cost solution but also a set of different solutions very close to the optimal one. A “good solution” has been defined as a solution whose cost does not exceed the obtained minimum cost by 3%.

The setting of PGA parameters improves the success rate of the method. Statistical analysis helps to determine this success rate in obtaining minimum/good design solutions with a PGA. Although it is not possible to precisely quantify the improvement in any network, it is possible to obtain reference values for improving the performance of the PGA. Thus, based on the statistical analysis of the results in the proposed model, it is possible to emphasize the following conclusions:

- When the population of the PGA is fixed, it exhibits significant robustness in the values of the crossover probability. The mutation probability is a more sensitive parameter. It must fall between approximately 3 and 5% for the analyzed examples, which presumes to make the mutation of one gene by chromosome.
- The success rate increases using the best set of parameters. For Hanoi network, the success rate increases from 2 to 8%, whereas in New York network it increases from 12 to 27%. The optimum combination of parameters is similar in both networks. However, this combination is valid only in the networks of this study. It would be necessary to verify the hypotheses in other models.
- The results indicate that for a fixed population size, the influence of the crossover operator in the optimization process is less than the influence of the crossover operator in a traditional GA. In a PGA, the frequency of mutation is a key operator. The results obtained with the benchmark networks show that a mutation probability of approximately $1/N_{VD}$ is a good starting point.
- An increase in population size allows the availability of better solutions. But the greater the population size is, the longer the computational effort. It becomes necessary to reach a commitment solution. From a practical

point of view it can be better to obtain a set of good solutions instead the minimum one, reducing the computational effort.

The proposed model seems valid for the water distribution networks optimization. The adjustments of its parameters were verified by statistical analysis. The methodology can be extrapolated to other optimization problems using the obtained results as a starting point.

Acknowledgements

This work was supported by the project DPI2009-13674 (OPERAGUA) of the Dirección General de Investigación y Gestión del Plan Nacional de I+D+I del Ministerio de Ciencia e Innovación, Spain.

References

- Alperovits E, Shamir U (1977). Design of Optimal Water Distribution Systems. *Water Resour. Res.*, Vol 13(6):885-900.
- Balla M, Lingireddy S (2000). Distributed Genetic Algorithm Model on Network of Personal Computers. *J. Comput. Civ. Eng.*, Vol 14(3):199-205. doi: 10.1061/(ASCE)0887-3801(2000)14:3(199).
- Baños R, Gil C, Agulleiro JI, Reca J (2007). A Memetic Algorithm for Water Distribution Network Design. *Advances in Soft Computing*, Vol 39:279-289. doi:10.1007/978-3-540-70706-6_26
- Cisty M (2010). Hybrid Genetic Algorithm and Linear Programming Method for Least-Cost design of Water Distribution Systems. *Water Resour. Manage.* Vol 24(1):1-24. doi:10.1007/s1269-009-9434-1.
- Chung G, Lansey K (2008). Application of the Shuffled Frog Leaping Algorithm for the Optimization of a General Large-Scale in a WaterSupply System. *Water Resour. Manage.* Vol 23: 797–823. doi:10.1007/s11269-008-9300-6.
- Cunha MC, Sousa J (1999). Water distribution network design optimization: simulated annealing approach. *J. Water Resour. Plann. Manage.* Vol 125(4): 215–221. doi: 10.1061/(ASCE)0733-9496(1999)125:4(215).
- Eusuff M, Lansey K (2003). Optimization of Water Distribution Network Design Using the Shuffled Frog Leaping Algorithm.” *J. Water Resour. Plann. Manage.*, 129(3):210–225. doi: 10.1061/(ASCE)0733-9496(2003)129:3(210).
- Fujiwara O, Khang, DB (1990). A two phase decomposition method for optimal design of looped water distribution network. *Water Resour. Res.*, Vol 26(4):539-549. doi: 10.1029/WR026i004p00539.

- Geem Z.W (2006). Optimal cost design of water distribution networks using harmony search. *Eng. Optimiz*, Vol 38(3):259-277. doi: 10.1080/03052150500467430.
- Goldberg DE, Kuo, CH (1987). Genetic Algorithms in Pipeline Optimization. *J. Comput. Civil. Eng*, Vol. 1(2):128-141. doi: 10.1061/(ASCE)0887-3801(1987)1:2(128).
- Goulter IC, Morgan DR (1985). An integrated approach to the layout and design of water distribution systems. *Civil Eng. Syst*, Vol 2(2): pp 104–113. doi: 10.1080/02630258508970389.
- Halhal D, Walters GA, Ouazar D, Savic DA (1997). Water Network Rehabilitation with Structured Messy Genetic Algorithms. *J. Water Resour. Plann. Manage.* Vol 123(3):137-147. doi: 10.1061/(ASCE)0733-9496(1997)123:3(137).
- Iglesias-Rey PL, Martínez-Solano FJ, Mora-Meliá D, Ribelles-Aguilar JV (2012). The battle water networks II: combination of meta-heuristic techniques with the concept of setpoint function in water network optimization algorithms. In: *Proc. 14th Water Distribution Systems Analysis symposium (WDSA)*, Engineers Australia, Adelaide, Australia.
- Jin YX, Cheng HZ, Yan J, Zhang L (2007). New discrete method for particle swarm optimization and its application in transmission network expansion planning. *Electr. Pow. Syst. Res.* Vol 77 (3-4): pp 227-233. doi: 10.1016/j.epr.2006.02.016.
- Lansley KE, Mays LW (1989). Optimization model for design of water distribution systems. *Reliability analysis of water distribution systems*, L. R. Mays, ed., ASCE, Reston, Va.
- Louati M, Benabdallah S, Lebdi F, Milutin D (2011). Application of a Genetic Algorithm for the Optimization of a Complex Reservoir System in Tunisia. *Water Resour. Manage.* Vol 25(10):2387–2404. doi: 10.1007/s11269-011-9814-1.
- Matías A (2003). “Diseño de redes de distribución de agua contemplando la fiabilidad mediante Algoritmos Genéticos”. Ph.D. Thesis, Universidad Politécnica de Valencia, Valencia, Spain. 2003.
- Nazif S, Karamouz M, Tabesh M, Moridi, A (2010). Pressure Management Model for Urban Water Distribution Networks. *Water Resour. Manage.* Vol 24(3):437-458. doi: 10.1007/s11269-009-9454-x.
- Prasad DT, Park NS (2004). Multiobjective genetic algorithms for design of water distribution networks. *J. Water Resour. Plann. Manage.*, 130 (1):73-82. doi: 10.1061/(ASCE)0733-9496(2004)130:1(73).
- Reca J, Martínez J. (2006). Genetic algorithms for the design of looped irrigation water distribution networks. *Water Resour. Res.*, 42(5): , W05416, doi: 10.1029/2005WR004383.
- Reca J, Martínez J, Gil C, Baños R (2008). Application of Several Meta-Heuristic Techniques to the Optimization of Real Looped Water Distribution Networks. *Water Resour. Manage.* Vol 22(10):1367-1379. doi: 10.1007/s11269-007-9230-8.
- Rossman LA (2000). *EPANET 2.0 User’s manual*. EPA/600/R-00/057, 2000.
- Savic DA, Walters GA (1997). Genetic Algorithms for Least-Cost Design of water Distribution Systems. *J. Water Resour. Plann. Manage.* Vol 123(2):67-77. doi: 10.1061/(ASCE)0733-9496(1997)123:2(67).

Su YL, Mays LW, Duan N, Lansey K E (1987). Reliability based optimization model for water distribution systems. *J. Hydraul.Eng.* Vol 113(12):1539–1556. doi: 10.1061/(ASCE)0733-9429(1987)113:12(1539).

Tsai FTC, Katiyar V, Toy D, Goff RA (2008). Conjunctive Management of Large-Scale Pressurized Water Distribution and Groundwater Systems in Semi-Arid Area with Parallel Genetic Algorithm. *Water Resour. Manage.* Vol 23(8):1497-1517. doi: 10.1007/s11269-008-9338-5.

Vairavamorthy K, Ali M (2000). Optimal design of water distribution systems using genetic algorithms. *Comput. Aided Civ. Infrastruc. Eng.*, Vol 15(5):374-382. doi: 10.1111/0885-9507.00201.

Wang QJ (1991). The Genetic Algorithm and Its Application to Calibrating Conceptual Rainfall-Runoff Models. *Water Resour. Res.* Vol 27(9):2467-2471. doi: 10.1029/91WR01305.