# All-in-one model for designing optimal water distribution pipe networks

**Dagnachew Aklog**[1] **and Yoshihiko Hosoi**[2]

[1]Center for International Affairs, Tottori University, Koyama Minami, Tottori, Japan
[2]Graduate School Engineering, Tottori University, Koyama Minami, Tottori, Japan

*Correspondence to:* Dagnachew Aklog (aklog@ciatu.tottori-u.ac.jp)

**Abstract.** This paper discusses the development of an easy-to-use, all-in-one model for designing optimal water distribution networks. The model combines different optimization techniques into a single package in which a user can easily choose what optimizer to use and compare the results of different optimizers to gain confidence in the performances of the models. At present, three optimization techniques are included in the model: linear programming (LP), genetic algorithm (GA) and a heuristic one-by-one reduction method (OBORM) that was previously developed by the authors. The optimizers were tested on a number of benchmark problems and performed very well in terms of finding optimal or near-optimal solutions with a reasonable computation effort. The results indicate that the model effectively addresses the issues of complexity and limited performance trust associated with previous models and can thus be used for practical purposes.

## 1 Introduction

The conventional approach to designing water distribution pipe networks is trial and error. A designer first assigns reasonable values to the design variables and analyzes the system using a simulation model, such as EPANET, to check whether the system requirements are satisfied. Based on the analysis results, the designer makes some changes in the design and analyzes the system again; the process is repeated until a "satisfactory" solution is found. This method of design, however, does not guarantee a low-cost system, let alone an optimal one. In addition, its application to large distribution systems tends to be exceedingly tiresome and time consuming. An alternative to this method of design is the optimization approach, in which the network design is formulated as an optimization problem to be solved using appropriate methods.

The optimization approach to designing pipe networks has been the focus of several studies in the past, and numerous mathematical models have been developed using different optimization techniques, including classical linear programming (Calhoun, 1971, etc.) and nonlinear programming (Lansey and Mays, 1989; etc.), as well as several stochastic

search techniques, such as simulated annealing (Cunha and Sousa, 1999), ant colony optimization algorithms (Maier et al., 2003) and genetic algorithms (GAs), which are by far the most widely used (Simpson et al., 1994; Savic and Walters, 1997; and many others). More recently, mixed-integer nonlinear programming has also been used (Bragalli et al., 2012).

Despite such significant research efforts and the development of numerous models, the application of the models for practical purposes has been limited, mainly because of the complexities associated with the use of the models and inadequate trust in their performances. In this study, bearing in mind that pipe network optimization models would be useful in developing countries where there is a need to construct new systems and expand existing ones to cope with high population growth and rapid urbanization, we have attempted to address the issues of complexity and limited trust. To that end, we have developed an easy-to-use, all-in-one model by combining different optimization techniques into a single package in which a user can easily choose what optimizer to use and compare the results of different optimizers to gain confidence in their performances. At present, three optimiza-
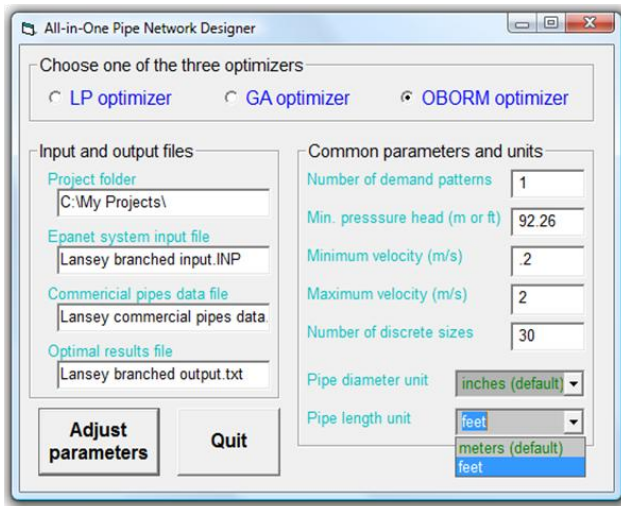
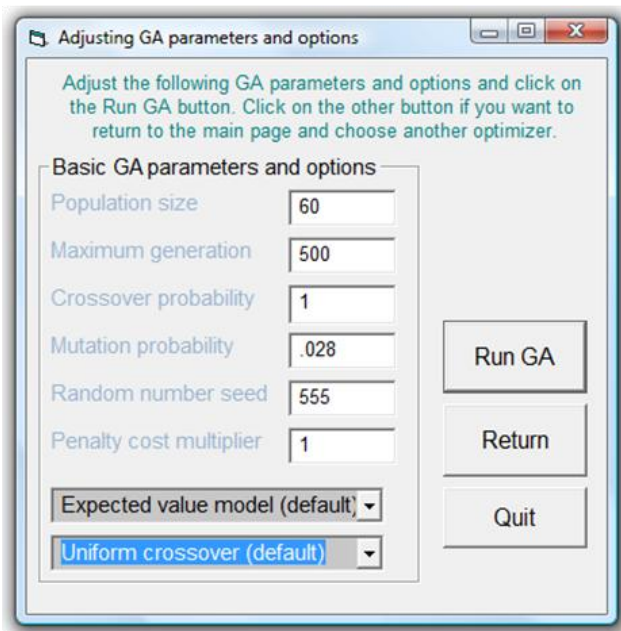**Figure 1.** The user interface of the all-in-one model.



**Figure 2.** The dialogue box for GA parameter adjustment.

tion techniques are included in the model: linear programming (LP optimizer), genetic algorithm (GA optimizer) and a heuristic one-by-one reduction method (OBORM optimizer). An overview of the model, how each optimizer works and the performance of the optimizers on test problems is presented in the subsequent sections.

## 2   Overview of the model

The model is organized with special attention to simplicity and ease of use. Figure 1 shows the user interface of the model, which is developed in Microsoft Visual Basic (VB)

6.0. The first step for users is to create an EPANET input file of the system to be designed and save it in the project folder. Here, EPANET's input file format is adopted because EPANET is believed to be the most widely used distribution system modeling software. The next step is to create a text file containing commercial pipe data, which is a list of the available discrete sizes and their unit costs, and save it in the same folder. If the system has to be designed under multiple loading conditions, a demand pattern data file should also be created and saved in the same folder. After specifying the name of the output file and setting the common parameters and pipe diameter and length units, the user simply needs to choose and run one of the three optimizers by checking the radio button. If the first radio button is checked, the LP optimizer will be ready to be run without any parameter adjustment. Both the GA and OBORM optimizers have a few parameters to be adjusted. Figure 2 shows the dialogue box for GA parameter adjustment.

## 3   The three optimizers

The three optimizers were previously developed by the authors in FORTRAN and have recently been compiled as separate DLL functions and added to the VB project as modules. As a precondition for using the model on a Windows platform, the optimizer DLL functions should be copied to the Windows system folder. It should be clear that while the LP optimizer provides optimal solutions for a branched networks design, there is no guarantee that the solutions obtained using the GA and OBORM optimizers will be optimal. However, in most cases they are "near-optimal" and satisfactory.

### 3.1   LP optimizer for branched networks

Since pipes are only available in discrete sizes, pipe costs and friction losses are both linearly related to pipe length, and pipe flows do not depend on pipe sizes, and the branched network design can be easily formulated as an LP problem by partitioning a pipe section into several segments of different diameters and then finding the optimal lengths of each segment with the aim of minimizing the total pipe cost of the network while satisfying the constraints.

The LP optimizer solves the nonstandard LP problem formulated this way using a two-phase method, which applies the well-known simplex method in two phases. As depicted in Fig. 3, after obtaining the necessary information through the user interface, the LP optimizer first reads input data, which includes network pipes and nodes as well as commercial pipe data. Then it identifies the end nodes in the network and generates the flow paths to each end node. The velocity of flow in each pipe and each discrete size is then calculated to exclude sizes that violate velocity constraints. The generation of first-tableau matrix elements is a crucial step that requires a careful look into the formulation of the branched net-

work LP problem. Once this matrix is generated properly, the problem can be easily solved using the two-phase method.

The application of LP in pipe network design is not new, but the improvement in this study is that routine tasks, such as generating the flow paths and matrix elements of the first tableau as well as interpreting the optimal results, are all automated in order to make the model easy to use, hence saving effort and time for the user.

## 3.2 GA optimizer

The GA optimizer basically applies the standard GA procedure shown in Fig. 4. First, an initial population is generated randomly. The individuals in the population are strings that are usually made up of binary bits (0 and 1), and the decimal value of each string represents a trial solution to the problem. The initial population is transformed into a new population by using the three basic GA operators: selection, crossover and mutation. The new population then undergoes a similar transformation and the process is repeated until a preset maximum generation number is reached. Of the three GA operators, selection is the most important. To select individuals from the current population for further reproduction, the survival probability (SP) of each string should be calculated. This requires analyzing the network represented by each trial solution, adding penalty costs to the objective functions when the pressure head and other constraints are violated, calculating fitness and performing scaling to avoid premature convergence.

In addition to these basic steps, heuristic improvement is introduced in the GA optimizer to address the slow convergence issue that is associated with standard GA. The improvement is based on the hypothesis that the probability of obtaining better individuals (strings) increases with an increase in the average performance of the population. The average population performance is supposed to, and generally does, increase from generation to generation. However, on some "bad" occasions, it deteriorates, as does the performance of the best string in a given generation, which contributes to the slow convergence. In the heuristic improvement, the strings are ranked based on their objective values, and a deteriorating string is replaced with a string of the same rank from the previous population. This avoids the occasional deterioration of the population while maintaining its variability or randomness.

## 3.3 OBORM optimizer

The OBORM, a simplified flowchart of which is shown in Fig. 5, is a simple yet efficient heuristic method previously developed by the authors, mainly for designing looped networks. It can also be applied to solving other nonlinear combinatorial optimization problems in which the decision variables have finite discrete solution spaces. When applied to pipe network design, the OBORM initially sets the sizes of
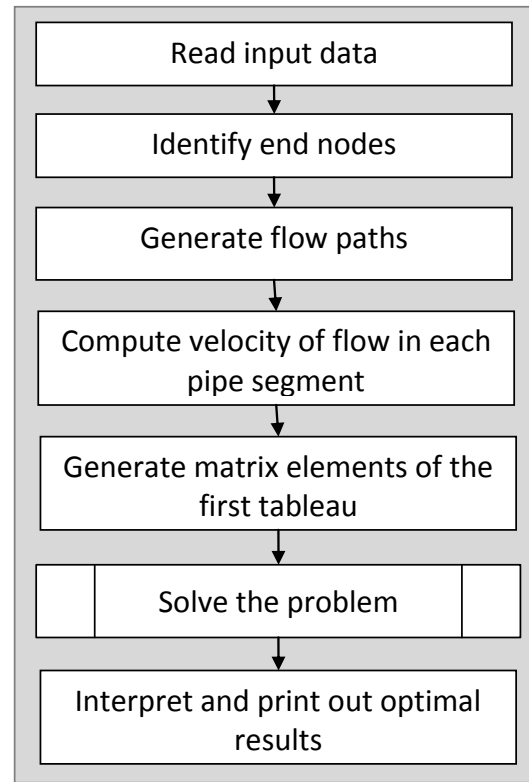
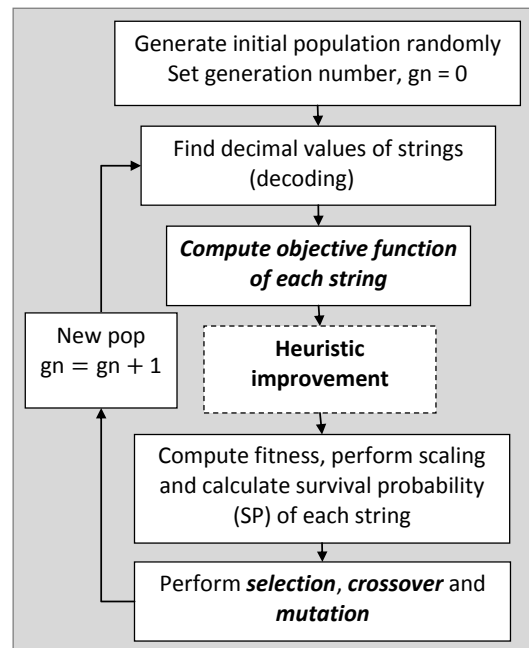**Figure 3.** A simplified flowchart of the LP optimizer.
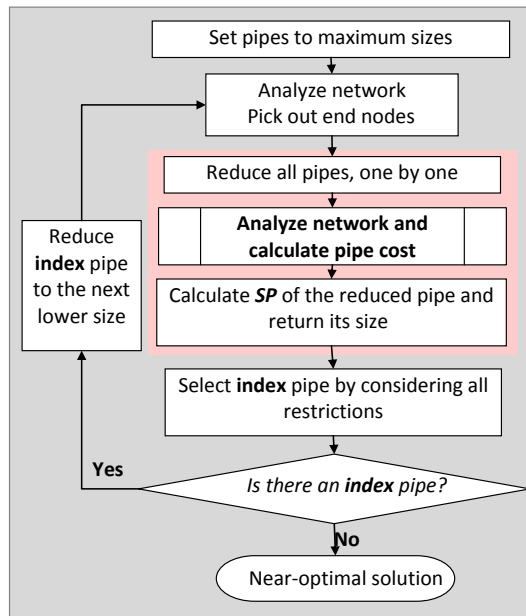
**Figure 4.** The basic steps in standard GA.

**Figure 5.** A simplified flowchart of the OBORM.

**Table 1.** The layout of a hypothetical branched network.

| | Node data | | | Pipe data | |
|---|---|---|---|---|---|
| Node No. | Elevation (ft[1]) | Demand (gpm[2]) | | Pipe No. | Length (ft) |
| 1 | 20 | 500 | | 1 | 100 |
| 2 | 50 | 200 | | 2 | 12 000 |
| 3 | 50 | 200 | | 3 | 6000 |
| 4 | 50 | 200 | | 4 | 9000 |
| 5 | 50 | 500 | | 5 | 6000 |
| 6 | 50 | 500 | | 6 | 12 000 |
| 7 | 50 | 500 | | 7 | 6000 |
| 8 | 50 | 1000 | | 8 | 6000 |
| 9 | 50 | 500 | | 9 | 6000 |
| 10 | 50 | 500 | | 10 | 12 000 |
| 11 | 120 | 200 | | 11 | |
| 12 | 120 | 200 | | 12 | 6000 |
| 13 | 80 | 200 | | 13 | 6000 |
| 14 | 120 | 200 | | 14 | 6000 |
| 15 | 120 | 800 | | 15 | 6000 |
| 16 | 120 | 200 | | 16 | 6000 |
| 17 | 220.2 | −6400 | | | |

[1] 1 ft = 0.304794 m. [2] 1 gpm = 0.2271 m$^3$ h$^{-1}$.
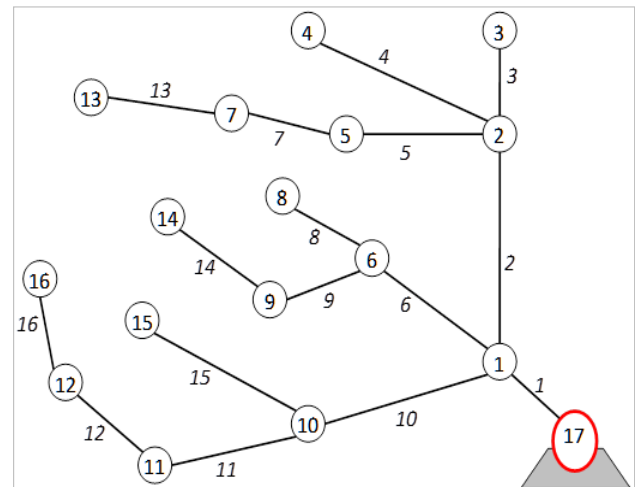


**Figure 6.** The layout of a hypothetical branched network.

all the pipes to certain maximum values. At this stage, both the pipe cost and the pressure head at the end nodes have maximum values. Then the pipe sizes are reduced one by one until any further reduction will result in a violation of any of the constraints. At every step, the pipe that results in the smallest pressure head drop at the most depressed node in the network ($\Delta hx$) per reduction in pipe cost ($\Delta$CP) when reduced to the next lower discrete size is selected as an index pipe. Hence, the selection parameter (SP) of each pipe $i$ is calculated as

$$SP_i = \frac{\Delta hx_i^\alpha}{\Delta CP_i^\beta}, \tag{1}$$

where $\alpha$ and $\beta$ are model parameters with a typical value of 1 (one) for all pipes but may sometimes take different values for different pipes. Another important factor in index pipe selection is a pressure drop at end nodes other than the most depressed node. In this optimizer, if $\Delta hx$ is lower than a certain small value (sv), the pressure drop in the other end nodes is considered. The sv is a model parameter with a typical value of 0 (zero) for looped networks.

## 4   Performance test

The performances of the optimizers have been tested on a number of benchmark problems. Here, only two of the problems are presented for the purpose of performance illustration. The first is a new branched network design, while the second is a simple yet typical looped network expansion problem.

### 4.1   Problem 1: branched network design

Figure 6 shows the layout of a hypothetical branched network, which was modified from Lansey and Mays (1989) and has 16 pipes, 16 junction nodes, 7 end nodes and a single source. The source pump at node 17 is assumed to operate at a steady condition with a flow of 6400 gpm and a total head of 220.2 ft. Network nodes and pipe data are given in Table 1. Thirty discrete pipe sizes of 1 in to 30 in are considered and the unit cost of the pipes (USD ft$^{-1}$) is assumed to be 3.45 $D$, if $D \le 10.93$ in, and 2.41$D^{1.15}$ otherwise. All

**Table 2.** A comparison of the results for the branched network design problem.

| Pipe | LP results | | GA results | OBORM results |
|------|------------|----|------------|----------------|
| No. | Diameter (in) | Length (ft) | Diameter (in) | Diameter (in) |
| 1 | 30 | 100 | 30 | 30 |
| 2 | 17, 16 | 9416, 2584 | 16 | 16 |
| 3 | 6, 5 | 998, 5002 | 6 | 6 |
| 4 | 6, 5 | 6119, 2881 | 6 | 6 |
| 5 | 14 | 6000 | 14 | 15 |
| 6 | 25, 24 | 5046, 6954 | 24 | 24 |
| 7 | 12 | 6000 | 13 | 12 |
| 8 | 9, 8 | 5840, 160 | 9 | 9 |
| 9 | 17 | 6000 | 17 | 17 |
| 10 | 25, 24 | 970, 11030 | 25 | 23 |
| 11 | 17, 16 | 2882, 3118 | 16 | 17 |
| 12 | 15 | 6000 | 14 | 15 |
| 13 | 8 | 6000 | 9 | 9 |
| 14 | 11 | 6000 | 12 | 12 |
| 15 | 15 | 6000 | 15 | 16 |
| 16 | 11 | 6000 | 11 | 13 |
| Pipe cost (USD) | 6 049 228 | | 6 096 012 | 6 112 920 |
| Execution time (s) | 0.03 | | 2.45 | 0.34 |

**Table 3.** A comparison of the results for problem 2.

| Solution method | Cost (USD) | Evaluations |
|-----------------|------------|-------------|
| Simpson et al. (1994) (GA – proportionate selection) | 1.750 | 20 790 |
| Simpson and Goldberg (1994) (GA – tournament selection) | 1.750 | 8700 |
| Maier et al. (2003) (ACOA – iteration best ant) | 1.750 | 8509 |
| GA optimizer | 1.750 | 5700 |
| OBORM optimizer | 1.750 | 1101 |

pipes are assumed to have a Hazen–Williams roughness coefficient of 120. A minimum pressure of 40 psi is required at all junction nodes, and the velocity of flow through the pipes is required to be between 0.2 and 2.0 m s$^{-1}$.

The LP model for this problem has 23 equations and 487 ($16 \times 30 + 7$) unknowns, including the slack variables. This is a "fairly large" problem, but the LP optimizer run on a 2.80 GHz desktop computer took only a fraction of a second to arrive at the optimal solution. The problem was also solved using the GA and OBORM optimizers, and the best results were found with the GA model parameters set to the values shown in Fig. 2 and the OBORM parameters set to $\alpha = 1$, $\beta = 1.5$ and $sv = 0.3$. Table 2 compares the results of the three optimizers. The results show that the LP is the most efficient, both in terms of cost and computation time, and is therefore the best choice for a branched network design. However, both the GA and OBORM optimizers were also able to find near-optimal solutions. This indicates that

both can be used for a branched network design if the split-pipe lengths of the LP are considered undesirable. On this and other test problems, the GA performed slightly better than the OBORM in terms of cost but required much more computation effort and more time and effort for parameter adjustment.

## 4.2 Problem 2: looped network expansion

The 14-pipe network shown in Fig. 7, which was first studied by Simpson et al. (1994), is used to illustrate the performance of the GA and OBORM in designing looped networks. As shown, 5 of the 14 pipes are new and have to be sized (with 1 of the 8 available discrete sizes), and 3 of the existing pipes may be duplicated with a new pipe in parallel, but not necessarily so; the remaining 6 existing pipes are to be left as they are. The network has to be designed to satisfy three demand patterns: a peak hour demand and two fire demands. With a solution space of $8^8$, this problem is relatively simple. But it is a typical problem that represents the case in many cities and towns in developing countries where the existing distribution networks should be expanded both in parallel and laterally to cope with rapid population growth and urbanization.

The problem was solved using the GA and OBORM optimizers, and both arrived at the best-known solution that was also found in previous studies. To see how fast the GA optimizer could converge on the optimal solution, the population size and maximum generation number were set to 50 and 100, respectively, and the optimizer was run 10 times with different seed numbers. The optimal solution was found in all 10 runs. The generation number at which the optimal solution
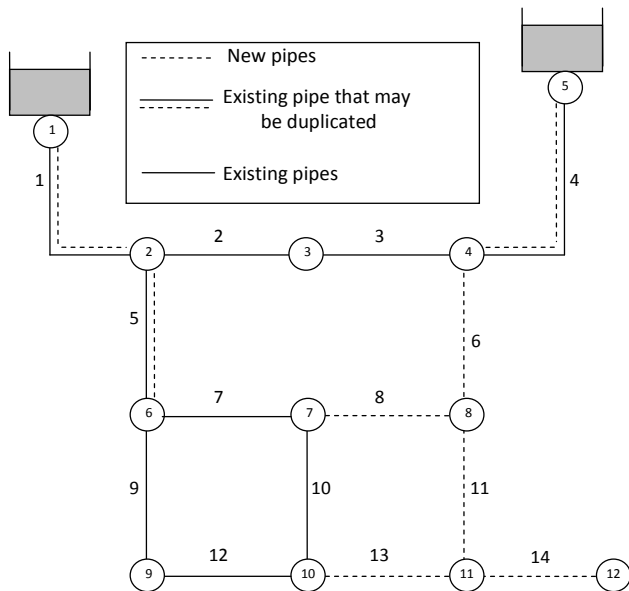
**Figure 7.** The network layout for problem 2 (after Simpson et al., 1994).

was attained varied from 19 to 59 and had an average value of 38. This makes the average number of evaluations needed to arrive at the optimal solution 5700 ($= 38 \times 50 \times 3$). Table 3 compares the optimal costs obtained and the number of evaluations needed by different researchers (models). From this table, we can see that the GA optimizer is computationally more efficient compared to the previous methods, and this improvement is attributed to the new heuristic method introduced to address the convergence issue of the standard GA. The OBORM optimizer, which was run with all the parameters set to their typical values, required by far the fewest evaluations. This clearly indicates that the OBORM outperforms the randomized search techniques in terms of computational efficiency. The execution times for our GA and the OBORM were 2.01 and 0.14 s, respectively.

## 5   Conclusion and remarks

With the aim of addressing the mathematical complexity and limited-performance trust issues associated with previous pipe network design models, we developed an easy-to-use, all-in-one model by combining three optimization techniques into a single package in which a user can easily choose what optimizer to use and compare the results for different optimizers to gain confidence in their performances. Based on its simplicity and the outstanding performance of the optimizers on several benchmark test problems, it is hoped that the model can be used for practical purposes, particularly in developing countries where a lot has to be done to reach the millions of people living without access to safe drinking water and to cope with the increasing demand in rapidly grow-

ing urban areas. However, the authors fully agree with the reviewers that the model should be tested in real-world case studies before it is made available for practical applications. An increase in the scale of the network to be designed will not cause much complexity in branched systems since this problem can be effectively handled using the LP optimizer. However, some challenges are expected in looped systems, such as the computational burden of using GA and the quality of the model outcome. The best way to tackle the challenges is first to solve the problem using the OBORM, which is much less sensitive to model parameters, and then refine the results using the GA optimizer. Research is currently underway to further improve the convergence property of the GA and to include other distribution system components, such as tanks and pumps, into the all-in-one model.

## References

Bragalli, C., D'Ambrosio, C., Lee, J., Lodi, A., and Toth, P.: On the optimal design of water distribution networks: A practical MINLP approach, Opt. Eng., 13, 219–246, 2012.

Calhoun, C.: Optimization of pipe systems by linear programming, in: Control of flow in Closed Conduits, edited by: Tullis, J. P., Colorado State University, Ft. Collins, 175–192, 1971.

Cunha, M. and Sousa, J.: Water distribution network design optimization: Simulated annealing approach, J. Water Resour. Plng. Mgmt., 125, 215–221, 1999.

Lansey, K. E. and Mays, L. W.: Optimization Models for Design of Water Distribution Systems, In: Reliability Analysis of Water Distribution Systems, ASCE, New York, USA, 37–84, 1989.

Maier, H. R., Simpson, A. R., Zecchin, A. C., Foong, W. K., Phang, K. Y., Seah, H. Y., and Tan, C. L.: Ant colony optimization for design of water distribution systems, J. Water Resour. Plng. Mgmt., 129, 200–209, 2003.

Savic, D. A. and Walters, G. A.: Genetic Algorithms for Least-cost Design of Water Distribution Networks, J. Water Resour. Plng. Mgmt., 123, 67–77, 1997.

Simpson, S. R., Dandy, G. C., and Murphy, L. J.: Genetic algorithms compared to other techniques for pipe optimization, J. Water Resour. Plng. Mgmt., 120, 423–443, 1994.