DEPARTMENT OF ENVIRONMENT,
TECHNOLOGY AND TECHNOLOGY MANAGEMENT

Optimisation of water distribution network design: a critical review

Annelies De Corte & Kenneth Sörensen

UNIVERSITY OF ANTWERP
Faculty of Applied Economics

EARNED EXCELLENCE
AACSB ACCREDITED
EST. 1916
THE BEST BUSINESS SCHOOLS IN THE WORLD

Stadscampus
Prinsstraat 13, B.226
BE-2000 Antwerpen
Tel. +32 (0)3 265 40 32
Fax +32 (0)3 265 47 99
http://www.ua.ac.be/tew

FACULTY OF APPLIED ECONOMICS

DEPARTMENT OF ENVIRONMENT,
TECHNOLOGY AND TECHNOLOGY MANAGEMENT

Optimisation of water distribution network design: a critical review

Annelies De Corte & Kenneth Sörensen

RESEARCH PAPER 2012-016
AUGUST 2012

University of Antwerp, City Campus, Prinsstraat 13, B-2000 Antwerp, Belgium
Research Administration – room B.226
phone: (32) 3 265 40 32
fax: (32) 3 265 47 99
e-mail: joeri.nys@ua.ac.be

D/2012/1169/016

Optimisation of water distribution network design: a critical review

University of Antwerp Operations Research Group – ANT/OR

Annelies De Corte * Kenneth Sörensen

August 21, 2012

This paper surveys the literature on the optimisation of water distribution network design. The water distribution network design (WDND) optimisation problem entails finding the type and diameter of each pipe in the network so that the total cost of the network is minimised without violating any hydraulic constraints. This is a difficult combinatorial optimisation problem, in which decision variables are discrete and both cost function and constraints are non-linear. Over the past 30 years, a large number of methods, especially in the field of (meta)heuristics, have been developed to solve this problem, most of which obtain good results on the available benchmark networks. In addition to outlining the basic features of each method, a detailed computational comparison is presented. The need for an adequate set of benchmark instances and the desirable properties of an instance set generator, are discussed. The paper concludes by pointing out some issues with the current state of the art in this domain and presents some suggestions for future research in this domain.

1 Introduction

A water distribution system is a network that consists of different components (pumps, reservoirs, pipes, valves, ...) that are used to transport drinking water from one or more resource nodes to multiple demand nodes (domestic, commercial, and industrial consumers). The water must be supplied in sufficient quantities and at an adequate pressure: water that is delivered at poor pressures limits the use of several applications, pressures that are too high can damage pipes in the network. Water distribution systems are generally owned and maintained by local governments as natural monopolies. In Flanders (Belgium), e.g., water management is organised by ten drinking water companies, that are each responsible for a certain region. International organisations such

*annelies.decorte@ua.ac.be

as the European Federation of National Associations of Water and Wastewater Services (Eureau) and the International Water Agency (IWA) aim to develop an international (global) drinking water policy.

Water distribution networks require decisions in four different phases with a different time horizon, as can be seen in Table 1. During the *layout* phase, the structure of the network is defined. In this phase it is determined where the different pipelines will be constructed, and where pumps, valves, water towers, reservoirs, and other components of the network should be built. In the *design* phase the type of each pipe is chosen, i.e., for each pipe in the network its optimal diameter and material are chosen. The *programming* phase is the only phase that involves socio-economic criteria and aims to set up a priority order in which water users should be connected to the water distribution net, see Roy et al. [1992]. The *planning* phase groups all decisions that are taken on a daily basis and that are concerned with the functioning of valves and pumps to ensure (among many other things) that sufficient water is available in all nodes of the network, see Słowiński [1986], Burgschweiger et al. [2009] and Verleye and Aghezzaf [2011].

Table 1: Optimisation of water distribution networks: different phases

Phase	**Decision level**	**Decision variables**
Layout	Strategic	System connectivity, topology
Design	Tactical	Pipe diameter, pipe roughness, ...
Programming	Tactical/Operational	Priority order users, ...
Planning	Operational	Pump and valve control, ...

In this paper, a thorough review is presented of the research that has been done on the second phase, the *design* of water distribution networks. In this phase, the physical structure of the network is assumed given. With all demand nodes, resource nodes, adjacency matrix and pipe lengths assumed known, the aim of this phase is to select both the diameter and the material of each pipe in such a way that the total cost of the network is minimised and all pressure-related constraints are fulfilled. This results in an optimisation problem that is both non-linear and highly constrained, and in which the decision variables are the type of each pipe (determined by diameter and material). These variables are discrete, since the type of each pipe has to be chosen from a set of commercially available types. The unit cost per meter pipe depends on the characteristics of the pipe type.

The remainder of this paper is organised as follows. In the next section, the water distribution network design (WDND) optimisation problem is presented. In Section 3, an extensive overview of existing methods for the WDND optimisation problem is given. A computational comparison of these methods can be found in Section 4. Section 5 lists the main shortcomings of previous work, including the absence of an adequate set of benchmark instances. In Section 6, an overview of previous attempts to automatically generate realistic networks of arbitrary size with varying characteristics is given. Section 7 concludes and gives some pointers for future research.

2 The WDND optimisation problem

A water distribution network is usually represented as a graph in which the pipes are represented by the set of arcs P and the set of nodes N represents points of demand, points of supply (reservoirs), and junctions, i.e., points connecting two or more pipes. The objective of the WDND optimisation problem is to minimise the total cost of the network design. The cost of an individual pipe depends on the *type* t that is chosen for this pipe from a list of commercially available types T. The type of a pipe determines both its diameter and the material of which it is made, which in term determine its hydraulic properties. If the cost per meter of a pipe of type t is represented by C_t and the length of pipe p is represented as L_p, the objective function of the WDND optimisation problem can be written as:

$$\text{Minimise total cost} = \sum_{p \in P} \sum_{t \in T} x_{tp}.C_t.L_p \tag{1}$$

where x_{tp} is a binary decision variable that determines whether pipe p is chosen from type t ($x_{tp} = 1$) or not ($x_{tp} = 0$).

The objective function is conditioned by physical mass and energy conservation laws, and by minimum head requirements in the demand nodes.

The *mass conservation law* must be satisfied for each node $n \in N$. This law states that the volume of water flowing into a node in the network per unit of time must be equal to the volume of water flowing out of this node. Let Q_{ij} represent the water flowing from node i to node j, and let S_n be the supply and D_n the demand of node n (all expressed in $m^3 s^{-1}$) then the following should hold:

$$\sum_{i \in N \setminus n} Q_{in} - \sum_{j \in N \setminus n} Q_{nj} = D_n - S_n \qquad \forall n \in N \tag{2}$$

Furthermore, for each closed loop $l \in O$, the *energy conservation law* must be satisfied. This law states that the sum of pressure drops in a closed loop is zero. Pressure drops (also called head losses) in piping systems are caused by wall shear in pipes and friction caused by piping components such as junctions, valves, and bends. In past research, only the first type of friction losses (in the pipes) are taken into account. Let ΔH_p represent the head loss in a pipe p (in m) that connects nodes n_1 and n_2 and H_{n_1} the pressure head in node n_1, then:

$$\Delta H_p = H_{n_1} - H_{n_2}$$

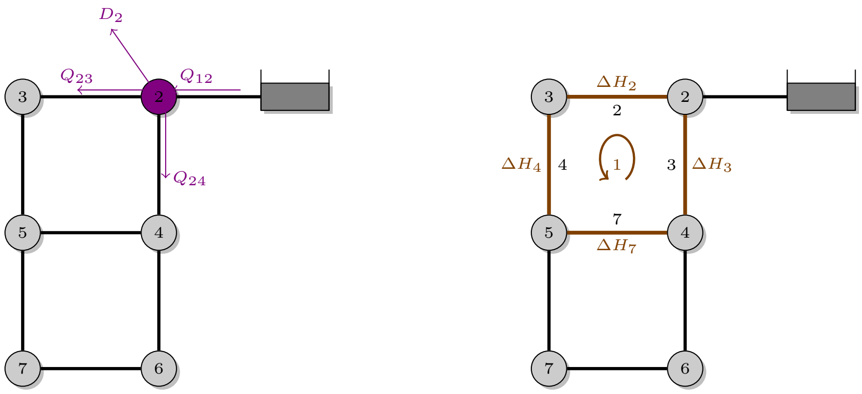
For the closed loop l, the energy conservation law can therefore be stated as:

$$\sum_{p \in \text{loop } l} \Delta H_p = 0 \qquad \forall l \in O \tag{3}$$

Approximating the head losses in the pipes of the network is usually done using the Hazen–Williams approximation:

$$\Delta H_p = w \frac{Q^a L_p}{C^a D^b} \tag{4}$$

In this equation, Q is the water flow rate in (m^3s^{-1}), L_p is the pipe length (in m), C is the Hazen–Williams roughness coefficient (unitless), D is the pipe diameter (in m), and a, b, and w are unitless parameters. Parameters D and C are determined by the type of a pipe and are assumed given for each available type. An example of both the mass and the energy conservation law can be seen in Figure 1.



Mass conservation law for node 2:
$Q_{12} - (Q_{24} + Q_{23}) = D_2$

Energy conservation law for loop 1:
$\Delta H_2 + \Delta H_4 - \Delta H_7 - \Delta H_3 = 0$

Figure 1: Example of the mass (left) and the energy conservation law (right)

Finally, *minimum pressure head requirements* exist for every (demand) node $n \in N$. Let H_n be the pressure head in node n (in m) and H_n^{min} the minimum pressure head in node n (in m). This constraint therefore can be represented as:

$$H_n \geq H_n^{min} \qquad \forall n \in N \tag{5}$$

The WDND optimisation problem assumes that pipe layout, nodal demands, and minimum head requirements in the nodes are known and constant. Although water distribution networks are dynamic systems, in which nodal demands and head requirements evolve over time, most research has focused on this static problem. Other components frequently found in real-life water distribution networks, such as pumps, valves, water towers, etc., are also disregarded in research on the WDND optimisation problem.

Most papers use the hydraulic solver EPANET2.0 to check hydraulic feasibility of the solutions they generate, although some papers use the previous version (EPANET1.0). EPANET software employs the gradient method proposed by Todini and Pilati [1987] to solve the mass and energy conservation laws. The head loss equation parameters in version 2.0 of EPANET are $a = 1.852$; $w = 10.6668$ (for SI units) and $b = 4.871$. The Hazen–Williams expression then reduces to:

$$\Delta H_p = 10.67 \frac{Q^{1.852} L_p}{C^{1.852} D^{4.871}} \tag{6}$$

EPANET2.0 is hydraulically stricter than the previous version (EPANET1.0), which uses $w = 10.5088$.

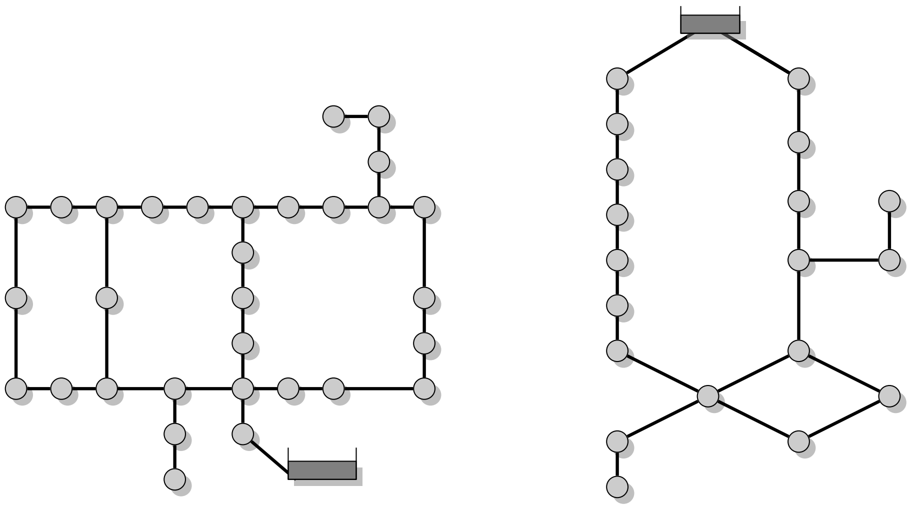


Figure 2: The "Hanoi network" (left) [Fujiwara and Khang, 1990] and the “New York City Tunnels” network (right) [Schaake and Lai, 1969]. The “two loop” network of Alperovits and Shamir [1977] can be found in Figure 1.

Figure 2 shows two examples of typical WDND instances found in the literature. The two loop network consists of one gravity-fed reservoir and eight pipes that can be optimised. The New York City Tunnels network consists of one gravity-fed reservoir and twenty-one pipes.

3 State of the art

The (re)construction of a water distribution network generally entails a significant capital investment, especially in an urban environment. Hence, a lot of research has been dedicated to the modelling and optimisation of the design of water distribution networks. The main focus of this research lies in the development of efficient optimisation methods, a large number of which have been developed and applied on a limited set of test problems. In the following paragraphs, each of these methods is briefly discussed. A comparison of results can be found in Tables 3 to 7.

3.1 Early methods

Alperovits and Shamir [1977] present one of the first contributions in the field of WDND optimisation. In contrast to later work, the authors aim to optimise sizing of components and take operational decisions simultaneously, and use a linear objective function. The *linear programming gradient* (LPG) method is applied. This method decomposes the optimisation problem into a hierarchy of two levels. On the first level an equation that gives the optimal cost of the network for any given feasible water flow is derived. On the second level, the water flow is systematically changed with the aim of improving cost.

Only local optima are reached by this search procedure. The authors propose to repeat the procedure using different starting points to increase the probability of finding the global optimum.

The approach of Alperovits and Shamir [1977] has been extended in several contributions. The optimisation model of Quindry et al. [1981], e.g., is an extension of this linear programming approach. The authors include extra terms in the gradient derivation to express the interaction between paths, which was neglected by Alperovits and Shamir [1977]. This correction improves the performance of the linear programming gradient method for the tested examples. Kessler and Shamir [1989] successfully analyse and improve the linear programming gradient method. Fujiwara and Khang [1990] use a two-phase decomposition method, which extends the linear programming gradient method of Alperovits and Shamir [1977] to nonlinear modelling.

Complete enumeration (CEn) evaluates every possible solution, and therefore automatically generates the global optimum. This technique is applied by Simpson et al. [1994], but has the major disadvantage of requiring very large computing times. Therefore, complete enumeration is not applicable on large networks.

In order to avoid the large computing times encountered when performing complete enumeration, Gessler [1985] and Loubser and Gessler [1990] use *selective enumeration* (SEn) which is faster, since not every possible solution is considered. Gessler [1985] bases the pruning of the search space on experience, whereas Loubser and Gessler [1990] set up three guidelines for pruning: (1) grouping sets of pipes and assuming that a single diameter will be used for each group, (2) progressively storing the lowest cost solution which satisfies the constraints and eliminating all the other solutions of higher cost and (3) checking on combinations that violate constraints and eliminating all combinations that include the same or smaller pipe sizes.

The non-linear nature of the problem has spawned a number of approaches based on *non-linear programming* (NLP) techniques. Shamir [1974] uses the *generalised reduced gradient method*. A Lagrangian function is constructed by penalizing the objective function for violation of the constraints (in this case the flow equations). This method distributes the problem to a sequence of subproblems that use a linear approximation of the constraints. Duan et al. [1990] also use the generalised reduced gradient method. Disadvantages of the NLP technique reported by the author are that the solution is dependent on the initial solution, and that there is a reasonable probability of converging to a local optimum (which is also true for most heuristic methods). Moreover, the available pipe sizes are considered as a continuous variable, whereas in reality they are taken from a set of discrete diameter sizes. The optimal values will therefore not necessarily conform with these discrete values. As a result, rounding the values is necessary, which can influence the results significantly. El-Bahrawy and Smith [1987], e.g., develop a non-linear package, MINOS, to solve water network layout problems in which continuous solutions are modified by a discretisation procedure to arrive at the equivalent discrete solution. Simpson et al. [1994] perform the round off manually.

3.2 Heuristics and metaheuristics

Starting in the early nineties, the different existing (meta)heuristics frameworks have been applied to solve the WDND optimisation problem. This paper follows the taxonomy of Sörensen and Glover [To appear], who divide metaheuristics frameworks into three classes. *Local search* metaheuristics operate on a single complete solution and iteratively improve it by making small adjustments called moves. *Population-based metaheuristics* operate on a set of solutions and find better solutions by combining solutions from that set into new ones. Finally, *constructive metaheuristics* build a solution by working with a single, unfinished, solution and adding one solution element at a time. Metaheuristic algorithms derived from a metaheuristic framework are always heuristic in nature, i.e., they do not guarantee to find the global optimal solution. For most real-life combinatorial optimisation problems, however, their flexibility and ability to find close-to-optimal solutions has made metaheuristics the preferred methods.

3.2.1 Local search metaheuristics

Simulated annealing (SA), introduced by Kirkpatrick et al. [1983], is a heuristic method based on the physical annealing of crystals to low energy states. The method uses a random move strategy, in which the current candidate solution is compared to random solutions in its neighbourhood. Potential solutions that perform better (i.e., that have a lower objective function value) are always accepted. Worse solutions are accepted with a certain probability, which is dependent on a time-varying and endogenous parameter called the *temperature*. The acceptance probability decreases as the decrease in solution quality is larger. The initial temperature is set to a high value and decreases during the optimisation process. This decrease implies a reduction in the probability of accepting solutions that have a higher cost. The process terminates when the temperature drops below the stopping value, which is usually close to zero. Simulated annealing as an optimisation method for the design of water distribution networks has been applied by Loganathan et al. [1995], Cunha and Sousa [1999] and Cunha and Sousa [2001].

Tabu search (TS) is a another local search technique, introduced by Glover in 1986 and formalised in Glover [1989]. The basic principle of tabu search is the use of memory structures. A *tabu list* keeps record of previously examined solutions. These recent candidate solutions become tabu for a certain number of iterations and cannot be visited again as long as they are on the list. The length of this list is defined by the *tabu tenure* parameter. Cunha and Ribeiro [2004] propose a tabu search algorithm to find the least-cost design for water distribution networks. They apply both a tabu search with a fixed value (TS 1) and a tabu search with a variable random value (TS 2) for the tabu tenure parameter.

3.2.2 Population-based metaheuristics

Genetic algorithms (GA) were the first population-based heuristic technique to be applied to the optimisation of water distribution networks. A genetic algorithm mimics the principles of natural evolution, such as variation, selection, recombination and mutation to exchange attributes between two or more good solutions. Basic concepts were developed by Holland [1975], Holland [1992] and Goldberg [1989]. Murphy and Simpson [1992] were the first to apply a genetic algorithm on water distribution networks, followed by Simpson et al. [1994]. All of these authors use a binary encoding in which each pipe diameter is assigned a binary code. A chromosome or candidate solution is then represented as a binary string. For example, if eight possible diameters are available for twelve pipes, the binary string will consist of twelve sets of three bits, each set representing the diameter of a pipe. Both Dandy et al. [1996] and Savic and Walters [1997] use an improved genetic algorithm (iGA), which uses a Gray coding instead of the traditionally used binary coding. The use of Gray coding overcomes the Hamming cliff effect (i.e., more than one bit needs to be flipped to move to the nearest neighbour), since adjacent integers differ in only one bit position. Gupta et al. [1999] avoid coding pipe diameter sizes in a binary alphabet by storing the set of candidate solutions as a vector of discrete pipe sizes. Moreover, in order to prune the solution space, they stratify the network in different diameter sets, based on the judgement of a design engineer. For example, pipes that are located close to the source and hence have to transport bigger flows are grouped into higher dimensional sizes.

A similar method is applied by Vairavamoorthy and Ali [2000] and Reca et al. [2008]. They also use real coding (RCGA) to avoid the Hamming cliff and to overcome redundant values generated by using a binary alphabet. Moreover, the authors use a variable penalty coefficient to obtain the objective function (which is a combination of the cost of the network and a penalty for the pressure violation). This variable penalty coefficient is dependent on the degree of violation of each of the pressure constraints, contrary to the fixed penalty coefficient used in previous research. A third change is the limited use of a hydraulic solver: Vairavamoorthy and Ali [2000] apply a linear transfer function that approximates the hydraulic behavior of the network. If the pressure violations exceed a certain value, the current setting is used as starting solution for the hydraulic solver. The authors conclude that this technique took on average only 70 percent of the computational time that it would take if the hydraulic solver had been used exclusively.

Differential evolution (DE) is a more recently developed population-based technique [Storn and Price, 1997]. The process starts with a population of vectors, which represent the candidate solutions. New solutions are generated by adding the weighted difference between two vectors to a third vector, an operation that is called the mutation process. In a next phase, called crossover, the parameters of these mutated vectors are mixed with the parameters of another vector, the target vector (another candidate solution), to yield a trial vector (the new candidate solution). In the selection process, the trial vectors are compared to the target vectors. If the trial vector yields a lower cost function than the target vector, the trial vector replaces the target vector in the next generation (iteration). Vasan and Simonovic [2010] apply this technique to the WDND optimisation problem.

A *memetic algorithm* (MA) [Moscato, 1989] is a population-based algorithm that combines evolutionary operators and local search techniques. Each individual (candidate solution) is subjected to a local search operator in order to improve (i.e., obtain a lower value for the objective function). After the local search, the individual interacts with other members of the population in a way that is similar to the crossover applied in evolutionary algorithms. These interactions result in the creation of new individuals (candidate solutions). This process of local search and evolutionary operations repeats until a stopping criterion is satisfied. Memetic algorithms are applied on the water distribution network optimisation problem in Baños et al. [2007] and Baños et al. [2010]. The population is initialised by generating random individuals, which are optimised by a hill climbing (local search) optimiser. A reproduction phase generates children by applying a variation operation. The children are optimised by the applied local search technique and the population is updated by combining previous and new individuals.

The *Cross-entropy* method (CE) is an adaptive algorithm based on variance minimisation [Rubinstein, 1999]. Perelman and Ostfeld [2007] apply this method to the optimisation of water distribution networks. In a first phase, a probability vector is initialised. This vector represents the probability that a certain diameter size is chosen for a certain pipe. A population of sample vectors (or candidate solutions) is generated by Monte Carlo sampling, using this probability vector. A performance value (which is equal to the total investment cost plus penalty for hydraulic infeasibility) is associated with each sample vector and all sample vectors are sorted in order of their performance value. The best vectors (best candidate solutions) are used to update the probability vector: the probability of choosing a certain diameter for a certain pipe increases as the number of times it was chosen from the top previous performance vectors enlarges. This process continues until a stopping criterion is satisfied.

Scatter search (SS) Glover [1977] employs only a small reference set (set of candidate solutions) and makes limited use of randomisation as a proxy for diversification. The diversification phase generates a population of trial solutions, using an initial or seed solution as its input. Each trial solution is a string of integers, the place of the integer in the string specifies the pipe and the integer represents the corresponding pipe diameter. In the improvement phase, the trial solutions are enhanced by local search techniques. The following step is an update of the reference set with the best solutions in terms of quality and diversity. A subset is generated, based on the reference set. In the final solution combination phase, the solutions of the subset are combined to one or more solution vectors (new candidate solutions). This scatter search procedure is applied by Lin et al. [2007].

The *immune algorithm* (IA) is inspired by theoretical immunology and is based on the affinity between antigens (substances that provoke an immune response) and antibodies (proteins used by the immune system to identify and neutralize foreign substances). Chu et al. [2008] develop an immune algorithm and a modified immune algorithm (mIA) for the optimisation of water distribution networks. In their algorithm, the antigen represents the configuration of variables in the optimal solution of the optimisation problem. The repertoire of antibodies represents a population of candidate solutions. Each antibody consists of real coded pipe diameters for each pipe. These antibodies (candidate solutions) are tested for their affinity with the antigen (value of the objective function). A lower investment cost generates a higher affinity, whereas a pressure violation has a negative effect on the affinity. The best antibodies (in terms of higher antibody-antigen affinity) are cloned by genetic operations such as crossover and mutation. These newly generated antibodies are tested again for their affinity with the antigen and for their antibody-antibody affinity. This antibody-antibody affinity is based on the Hamming distance: the higher the Hamming distance, the smaller the affinity. If they possess a higher antigen-affinity than the antibodies that are member of the memory set, these current members are replaced by the newly generated ones with the lowest antibody-antibody affinity, to retain diversity. This procedure is repeated until a stopping criterion (in this case a maximum number of evaluations) is met.

The *shuffled frog leaping algorithm* (SFLA), introduced in Eusuff and Lansey [2003], is a memetic metaheuristic that is designed to seek a globally optimal solution by performing an informed heuristic search using a heuristic function. The basic idea resembles the idea underlying genetic algorithms and starts from a virtual population, consisting of individuals that are characterised by their memetic vector. The difference between the two methods is that under the memetic representation, ideas are passed between all individuals (possible solutions), whereas only parent-sibling interactions are allowed in a genetic algorithm. After a defined number of memetic evolution steps (number of iterations), a shuffling process gathers and re-sorts the evolved sets (candidate solutions). This process repeats until the stopping criterion is met.

3.2.3 Constructive metaheuristics

Ant colony optimisation algorithms are inspired by the behaviour of ants and were first described by Dorigo et al. [1996] and further modified by Stützle and Hoos [2000], a.o.. Colonies of ants use pheromone trails to optimise their search for food. ACO algorithms mimic the behaviour of a colony of cooperating ants that look for the optimal path (optimal solution) to reach the food source. They adopt a stochastic solution construction process, in which each ant constructs a solution from not yet selected elements based on information collected by the colony. This information is stored by assigning to each potential solution element an amount of pheromone, which is proportional to the quality: better paths (i.e., solutions with a better value for the objective function) maintain higher pheromone levels on their elements. The higher the level, the higher the probability that this solution element will be selected again in the future. Periodically, the pheromone level is reduced to reflect evaporation. This process is repeated multiple times. Max-Min Ant System, developed by Stützle and Hoos [2000], provides dynamically evolving bounds on the pheromone trail intensities, such that the pheromone intensity on all paths is always within a specified limit of the path with the greatest pheromone intensity. Therefore, no pheromone level will reduce to zero, which means that every path has a non-trivial probability of being selected, which leads to a wider exploration of the search space and avoids premature convergence. Maier et al. [2003] use ant colony optimisation for the optimisation of water distribution networks. Ant Systems (AS) are used by Zecchin et al. [2005]. In Zecchin et al. [2006], both Ant Systems and the Max-Min Ant System (MMAS) are studied. The water distribution network problem is converted to an unconstrained problem by the use of a penalty function that expresses pressure violations. In this unconstrained case, the optimisation takes the form of minimising the sum of the network design cost plus the penalty cost. A network design is the WDND equivalent for a path and therefore, higher pheromone levels will be assigned to paths with lower design costs.

Particle swarm optimisation (PSO) was introduced by Kennedy and Eberhart [1995] and is based on the behaviour of a bird flock (see Poli et al. [2007] for a recent overview). A swarm of particles (set of candidate solutions) is placed in the search space of the optimisation problem. Each particle (possible solution) is determined by its current position, previous best position and velocity. These parameters are updated according to the relationship between the particles parameters and the solutions found by individuals of their topological neighbourhood. The search is biased towards better solutions: particles will flock (solutions will evolve) to better regions of space, representing the best solutions.

Geem [2009] introduced particle swarm harmony search optimisation (PSHS), a method that combines harmony search optimisation with particle swarm optimisation. The method starts with a randomly generated set of solution vectors, stored in the harmony memory. New vectors (possible solutions) are generated from this set, by applying one of three possible operations on every element of the solution vector (solution). These operations are: random selection (where a new value is chosen out of the candidate set), memory consideration (where a new value is chosen out of the harmony memory and eventually replaced by the best value with a certain probability (the so called "particle swarm rate") and pitch adjustment (where a new value is chosen out of the harmony memory and adjusted to a neighbouring value, using a certain probability, the so called "pitch adjusting rate"). If the newly generated vector is better than the worst vector in the harmony memory (with respect to the objective function), the latter is replaced by the former. This procedure continues until a stopping criterion (a maximum number of newly generated vectors) is met.

3.3 Multi-objective optimisation

More recently, research in the field of design of water distribution networks has extended from single-objective optimisation — finding the optimal design that does not violate (hydraulic) constraints — to a multi-objective perspective. The multi-objective problem is most frequently stated as a bi-objective one, where the minimisation of the total network design cost and the minimisation of pressure deficit represent the two conflicting objectives [See, e.g., Keedwell and Khu, 2006, di Pierro et al., 2009]. After applying multi-objective optimisation on a water distribution network, the decision maker can use the resulting Pareto-front to investigate the trade-off between solutions that fully satisfy (hydraulic) constraints, but come with a higher cost, and solutions that are characterised by both slight violation of the pressure requirement and significant cost reductions. This second solution could be a useful consideration if the pressure violations are small and concern non-strategic nodes. Some authors add a third objective, e.g., Montalvo et al. [2010] add a reliability assessment from an economic point of view. This objective quantifies the costs of the water not delivered due to disruptions in the systems and the associated repair costs and has to be minimised. A description of the multi-objective algorithms developed in this context, is beyond the scope of this paper.

4 Computational comparison

In this section, the developed metaheuristics for the WDND optimisation problem are compared. Table 2 describes the column headings and symbols used in the results tables. Detailed results can be found in Tables 3 to 7.

Table 2: Column headings and symbols used in the results tables

M	Applied method
A	Author(s)
w	Hydraulic w-coefficient
F	Feasibility under EPANET2.0 (F: feasible, IF: infeasible)
TC	Total network cost (in $mUSD$)
#	Mean number of evaluations (* : only one value reported)
T	Calculation time (in s) and computing infrastructure

Table 3: Overview of results: “two loop” network Alperovits and Shamir [1977]

M	A	w	F	TC	#	T
LP	Alperovits and Shamir [1977]	na	F	0.480	19*	4.05 (IBM 370/168)
LP	Quindry et al. [1981]	10.9031	F	0.442	na	na
LP	Kessler and Shamir [1989]	na	IF	0.418	na	na
GA1	Savic and Walters [1997]	10.5088	F	0.419	65,000*	< 600 (PC 486/DX2 50 MHz)
GA2	Savic and Walters [1997]	10.9031	F	0.420	¡ 250,000	< 600 (PC 486/DX2 50 MHz)
SA	Cunha and Sousa [1999]	10.5088	F	0.419	25,000*	40 (Pentium PC 166 MHz)
GA	Gupta et al. [1999]	na	IF	0.408	na	na
GA	Wu and Simpson [2001]	10.5088	F	0.419	74,67*	?
SFLA	Eusuff and Lansey [2003]	10.6668	F	0.419	11,323	?
TS1	Cunha and Ribeiro [2004]	na	F	0.420	na	16 (Pentium PC 433 MHz)
TS2	Cunha and Ribeiro [2004]	na	F	0.420	na	16 (Pentium PC 433 MHz)
HS	Geem [2006]	10.6668	F	0.419	2,891	2 (Intel Celeron 1.8 GHz)
CE	Perelman and Ostfeld [2007]	10.6668	F	0.419	35,000*	na
SS	Lin et al. [2007]	10.5088	F	0.419	3,482	1.58 (Pentium PC 2,4 GHz)
SS	Lin et al. [2007]	10.6668	F	0.419	3,215	1.61 (Pentium PC 2,4 GHz)
SS	Lin et al. [2007]	10.675	F	0.419	3,183	1.77 (Pentium PC 2,4 GHz)
SS	Lin et al. [2007]	10.9031	F	0.419	5,274	1.84 (Pentium PC 2,4 GHz)
PSHS	Geem [2009]	10.6668	F	0.419	233	1 (Intel Celeron 1.8 GHz)
MA	Baños et al. [2010]	10.6668	F	0.419	na	39 (workstation 2 GHz)

Table 4: Overview of results: 14 pipe problem by Gessler [1985]

M	A	w	F	TC	#	T
SEn	Gessler [1985]	na	F	1.834	900	?
CEn	Simpson et al. [1994]	10.6750	F	1.750	11,940,000	295,200 (SUN 4/280 16.67 MHz)
NLP	Simpson et al. [1994]	10.6750	F	1.750	na	480 (SUN 4/280 16.67 MHz)
GA	Simpson et al. [1994]	10.6750	F	1.750	20,790	2,700 (SUN 4/280 16.67 MHz)
ACO	Maier et al. [2003]	10.6668	F	1.750	8,509	na
TS1	Cunha and Ribeiro [2004]	na	?	1.728	na	46 (Pentium PC 433 MHz)
TS2	Cunha and Ribeiro [2004]	na	?	1.728	na	55 (Pentium PC 433 MHz)

Table 5: Overview of results: New York City Tunnels problem by Schaake and Lai [1969]

M	A	w	F	TC	#	T
na	Fujiwara and Khang [1990]	10.5088	IF	36.10	na	na
iGA	Dandy et al. [1996]	-		38.80	96750*	<1,500 (Sun Sparc 1+ 25 MHz)
GA	Savic and Walters [1997]	10.5088	IF	37.13	na	na
GA	Savic and Walters [1997]	10.9031	F	40.42	na	na
GA	Lippai et al. [1999]	10.5088	IF	38.13	46,016*	
RCGA	Vairavamoorthy and Ali [2000]	na	IF	37.09	800	660 (UNIX mainframe computer)
SA1	Cunha and Sousa [2001]	10.5088	IF	37.10	na	162 (Pentium PC 400 MHz)
SA1	Cunha and Sousa [2001]	10.9031	IF	40.40	na	162 (Pentium PC 400 MHz)
SA2	Cunha and Sousa [2001]	10.6792	IF	38.80	na	162 (Pentium PC 400 MHz)
ACO	Maier et al. [2003]	10.6668	F	38.64	13,928	na
SFLA	Eusuff and Lansey [2003]	10.6688	IF	38.13	31,267*	na
SFLA	Eusuff and Lansey [2003]	10.6688	F	38.80	21,569*	na
TS1	Cunha and Ribeiro [2004]	na	IF	37.13	na	54 (Pentium PC 433 MHz)
TS2	Cunha and Ribeiro [2004]	na	IF	37.13	na	60 (Pentium PC 433 MHz)
AS	Zecchin et al. [2005]	10.6688	F	38.64	22,052	na
MMAS	Zecchin et al. [2006]	10.6688	F	38.64	30,711	na
HS	Geem [2006]	10.6688	F	38.64	5,991	7 (Intel Celeron 1.8 GHz)
SS	Lin et al. [2007]	10.5088	IF	36.68	43,913	78 (Pentium PC 2,4 GHz)
SS	Lin et al. [2007]	10.6668	IF	38.13	46,603	71 (Pentium PC 2,4 GHz)
SS	Lin et al. [2007]	10.675	F	38.64	57,58	73 (Pentium PC 2,4 GHz)
SS	Lin et al. [2007]	10.9031	F	40.42	56,552	92 (Pentium PC 2,4 GHz)
IA	Chu et al. [2008]	10.9031	F	40.42	75,829	395 (AMD-1.39GHz PC platform)
IA	Chu et al. [2008]	10.5088	IF	37.13	73,928	393 (AMD-1.39GHz PC platform)
mIA	Chu et al. [2008]	10.9031	F	40.42	54,827	259 (AMD-1.39GHz PC platform)
mIA	Chu et al. [2008]	10.5088	IF	37.13	56,062	339 (AMD-1.39GHz PC platform)
PSHS	Geem [2009]	10.6688	F	38.64	5,923	8 (Intel Celeron 1.8 GHz)
DE	Vasan and Simonovic [2010]	10.6668	F	38.64	30,701*	30.7 (Pentium PC 3.0 GHz)

Table 6: Overview of results: Hanoi problem by Fujiwara and Khang [1990]

M	A	w	F	TC	#	T
na	Fujiwara and Khang [1990]	10.5088	IF	5.562	na	na
GA	Savic and Walters [1997]	10.5088	IF	6.073	na	<216,000 (PC 486/DX2 50 MHz)
GA	Savic and Walters [1997]	10.9031	F	6.195	na	<216,000 (PC 486/DX2 50)
SA	Cunha and Sousa [1999]	10.5088	IF	6.056	53,000*	7,200 (Pentium PC 166 MHz)
RCGA	Vairavamoorthy and Ali [2000]	na	IF	6.056	1,600	1500 (UNIX mainframe PC)
fmGA	Wu and Simpson [2001]	?	F	6.182	113,600*	
SA1	Cunha and Sousa [2001]	10.5088	IF	6.073		360 (Pentium PC 400 MHz)
SA1	Cunha and Sousa [2001]	10.9031	F	6.183		360 (Pentium PC 400 MHz)
SA2	Cunha and Sousa [2001]	10.6792	F	6.093		360 (Pentium PC 400 MHz)
SFLA	Eusuff and Lansey [2003]	10.6668	IF	6.073	27,546	na
TS1	Cunha and Ribeiro [2004]	na	IF	6.056	na	391 (pentium pc 433 MHz)
TS2	Cunha and Ribeiro [2004]	na	IF	6.056	na	445 (pentium pc 433 MHz)
AS	Zecchin et al. [2005]	10.6668	F	6.367	67,136	na
MMAS	Zecchin et al. [2006]	10.6668	F	6.134	85,571	na
HS	Geem [2006]	10.6668	F	6.081	27,721*	35 (Intel Celeron 1.8 GHz)
CE	Perelman and Ostfeld [2007]	10.6668	F	6.081	97,000*	na
SS	Lin et al. [2007]	10.5088	IF	6.056	40,864	114 (Pentium pc 2,4 GHz)
SS	Lin et al. [2007]	10.6668	IF	6.081	43,149	123 (Pentium pc 2,4 GHz)
SS	Lin et al. [2007]	10.675	IF	6.097	59,047	172 (Pentium pc 2,4 GHz)
SS	Lin et al. [2007]	10.9031	F	6.183	59,561	175 (Pentium pc 2,4 GHz)
SA	Reca et al. [2008]	10.6668	F	6.333	25,000*	na
GA	Reca et al. [2008]	10.6668	F	6.173	26,457*	na
PSHS	Geem [2009]	10.6668	F	6.081	17,980*	34 (Intel Celeron 1.8 GHz)
MA	Baños et al. [2010]	10.6668	F	6.231	na	143 (Workstation of 2 GHz)
DE	Vasan and Simonovic [2010]	10.6668	F	6.195	56,201*	108.3 (pentium IV pc 3.0 GHz)

Table 7: Overview of results: Balerma network by Reca and Martínez [2006]

M	A	w	F	TC	#	T
MA	Baños et al. [2010]	10.6668	-	3.120	na	3,202 (workstation of 2 GHz)
PSHS	Geem [2009]	10.6668	-	2.633	45,400	na
SA	Reca et al. [2008]	10.6668	F	3.476	25,000*	na
GA	Reca et al. [2008]	10.6668	F	3.738	24,000*	na

As can be derived from Tables 3 to 7, nearly all methods obtain good solutions on the applied benchmark test problems. Comparison of the developed methods, however, remains difficult for multiple reasons:

- Only few case study benchmark networks are publicly available and moreover, only few authors apply their methods on every available benchmark network. Therefore, optimisation results remain very case-specific, and the robustness of the developed methods when applied to different benchmark instances, has not been demonstrated.

- Even when the same networks are used as basis of comparison, a lack of uniformity remains, due to the use of different hydraulic solvers. EPANET by Rossman

[2000] is the most frequently used hydraulic solver, but two versions of it exist. These versions use different values for parameter w in the Hazen–Williams equation (EPANET1: $w = 10.5088$ and EPANET2: $w = 10.6688$), which influences optimal solutions significantly, since the latter is hydraulically stricter and leads to higher cost alternatives.

As a result, it is questionable that scientifically valid conclusions on the performance of the developed methods can be made, based on these case-study and solver-specific results. For this reason, no attempt is made in this paper to derive a global performance measure, as this might give the (erronous) impression that such a measure can be meaningfully analysed.

5 Potential improvements in the state of the art

In this section, some potential issues related to much of the published research on WDND optimisation are identified, and some ways in which these issues can be avoided in future research are suggested.

Broadly speaking, two major shortcomings can be identified that characterise most papers published on the topic. First, most of the developed methods are not based on modern, established principles of metaheuristic design. Most of them are black-box procedures, that hardly use any problem-specific information, and are often much more complicated than strictly necessary. Second, and most importantly, the developed methods are not adequately tested, and thus no strong conclusions can be drawn on their performance.

Early research in the field of metaheuristics used to focus on the development and application of different methods, based on ever more outlandish metaphors. The motivation for these early metaheuristics did not extend beyond the natural process that they supposedly imitated. More recent research, however, has convincingly demonstrated that the best-performing metaheuristics are those based on a profound analysis of the problem at hand. In more established application domains of metaheuristics research, such as vehicle routing [a complete ranking of the best-performing methods can be found in Szeto et al., 2011], vehicle routing with time windows [Braysy and Gendreau, 2005], graph coloring [Galinier and Hertz, 2006], or project scheduling [Kolisch and Hartmann, 2006], the list of best-performing metaheuristics is dominated by methods that invariably use problem-specific operators and black-box application of a nature-inspired metaheuristic frameworks are rare exceptions. The large "competition" between different researchers attempting to develop the best methods, and the high standard of testing (see further) in these application domains have ensured that only the best and most robust methods are published. In these application domains, there is now also considerably attention to the theoretical study of the relationship between problem characteristics and metaheuristics performance [see, e.g., Watson et al., 2003, 2006].

In contrast, research on WDND optimisation has focused on developing metaheuristics based on complicated nature-inspired metaheuristic frameworks without much attempt to understand the relationship between the nature of the problem and the performance of the developed metaheuristics. This leads to black-box implementations that are generally much more complicated than strictly necessary to achieve similar performance. The result of these added levels of complexity is that the methods are much more difficult to implement and maintain, and suffer from a lack of robustness with respect to different input data sets.

A second criticism that can be levied at the current state of the art in WDND optimisation research is that the metaheuristics developed in this field are not adequately tested. The benchmark networks used for testing purposes in this field are both very simple and very small, and it is no surprise that most of the developed methods can find the optimal solution or a solution that is very close to optimal. Additionally, the number of available instances on which algorithms can be tested is limited to only a handful.

The problem size of a WDND optimisation problem instance is determined by the size of the solution space and the size of the set of equations that needs to be solved by the network simulator (EPANET). The size of the set of equations is determined by the number of junctions and loops in the network[1], whereas the size of the solution space, i.e., the number of potential solutions, is determined by the number of pipes and the number of possible pipe types[2]. Table 8 compares the problem size of some of the best-known problem instances in the WDND optimisation literature to the size of some real-life instances.

Table 8: Dimensions of benchmark and realistic networks

	Two loop	14 pipes	New York	Hanoi	Balerma	Apulia	Realistic
Junctions	7	12	20	32	447	24	j
Loops	2	3	2	3	8	11	(6,166 + 1 - j)
Pipes	8	14	21	34	454	34	6,166
Cost function	1	1	1	1	1	1	1
Available pipes	14	8	16	6	10	9	90
equations	18	30	44	70	910	70	12,334
solution space	14^8	8^8	16^{21}	6^{34}	10^{464}	9^{34}	$90^{6,166}$
	1.5×10^9	1.7×10^7	1.9×10^{25}	2.9×10^{26}	1×10^{464}	1.0×10^{56}	$5.3 \times 10^{10,826}$

In conclusion, it can be stated that the field of metaheuristics for WDND optimisation is still somewhat in its infancy. Research has focused on the development of complicated nature-inspired black-box methods, that are not adequately tested on a representative set of benchmark instances, at least partly because such a set does not exist. It is very likely that much simpler and much more robust methods can be developed according to modern insights in the design of metaheuristics. It follows that the applicability of most of these earlier developed methods as a decision support tool for the optimisation

[1]The *number of equations* N_e is determined by the number of pipes P (head loss equations), junctions J (mass conservation equations), loops L (energy conservation equations) and cost (objective) functions C: $N_e = P + J + L + C$. Since Euler's law for planar graphs requires that $J - P + L = 1$, and the number of cost functions is usually 1, it follows that $N_e = 2P + 2$. In other words, the number of equations that need to be solved to evaluate a given solution is roughly equal to twice the number of pipes in the network.

[2]The *number of possible solutions* is determined by the number of pipes (P) in the network, and the number of available pipe types (T), determined by diameter and material. If every pipe can take every available configuration, the number of possible solutions is C^T

of real-life water distribution networks, is very limited.

6 Algorithmic generation of benchmark instances

For various reasons, many water distribution companies like to keep their network layout confidential. Therefore, only a few real-life networks are available for research ends. Nevertheless, it is of crucial importance that researchers can test their developed methods on multiple realistic networks. This is necessary in order to be able to verify the performance on real-life settings, their performance in comparison to other methods and their sensitivity and robustness. The need for adequate test instances has been recognised in the literature. Cunha and Ribeiro [2004], e.g., state that "it would be very useful to build up a library containing a significant number of case studies to permit the comparison of the results provided by different solution methods".

The following desirable properties of a virtual water distribution network generation tool can be extracted from the literature:

- The development of virtual water distribution networks should be done *algorithmically*, in order to be able to generate an extensive library of diverse problem instances.
- The tool should be able to generate realistic networks of arbitrary size and characteristics, corresponding to the networks of different sizes and characteristics found in real life. This can be done by defining different water distribution network patterns (industrial, rural, commercial, residential), which could be combined in networks of any size. Each of these patterns has its own characteristics, which can be reflected in adjustment of parameter settings. Possible parameters could be the number of demand nodes, supply nodes, loops, demand patterns, etc.
- The database of generated networks should be *freely available online*, so that every researcher can use the instances. This will foster scientific developments in this research area.
- The generated networks should be available in *EPANET format*, since this is the most frequently used hydraulic solver.

Several attempts to generate artificial water distribution networks have already been made. Brumbelow et al. [2007] develop two virtual cities, called Micropolis and Mesopolis. Micropolis is based on the development pattern of a small city of 5,000 inhabitants, whereas Mesopolis represents a larger city of about 100,000 residents. Both cities resemble a very realistic network setting. The main shortcoming of this approach is, as stated by Sitzenfrei [2010], the very time consuming manual construction process. This limits the generation of an extensive library of test networks. Möderl et al. [2011] develop the Modular Design System, which algorithmically generates a variety of different water distribution networks. The main advantage of this technique is that this systematic generation process allows many different types of networks to be generated. The main drawback of this technique is that it makes significant assumptions and simplifications: no pumps, tanks nor valves are present and the network layout includes a large amount of unrealistic symmetry. These simplifications do not strike with reality.

7 Conclusions and future research

The optimisation of water distribution networks has received a large amount of attention of the research community over the past thirty years. This paper provides an elaborate overview of the main research efforts in this area and a critical review, together with some recommendations for further development in this area of research.

Numerous heuristic methods have been developed, most of which obtain good results on the available benchmark networks in terms of reaching the reported minimal cost. This paper has argued that this fact is probably the result of the simplicity of the benchmark instances used, rather than of the intrinsic quality of the developed methods. Two significant shortcomings have been identified that characterise the state of the art in the field of WDND optimisation: (1) the developed methods are not based on established principles of metaheuristic design, and (2) the developed methods are not adequately tested. As a result, it is not very likely that the developed optimisation methods can be successfully applied to the optimisation of large and complex real-life WDND optimisation problems.

In order to be able to test the proposed methods accurately, an extensive library of high-quality networks should be developed. This virtual water distribution network generation tool should have at least the following characteristics: algorithmic network development, generation of realistic networks, on-line availability, and availability in the EPANET-format. The development of such a library is crucial, since it would not only enhance research quality in this domain significantly, but also foster the development of decision support tools that could be applied in realistic settings.

Further research should also focus on the discrepancy that exists between the theoretical WDND optimisation problem described in this paper and its counterpart in real life. In the WDND optimisation problem, e.g., all demand is assumed to be situated in the *nodes* of the network. The EPANET solver, too, only allows demand to be situated in a node. In real life, demand is often more or less uniformly spread along the length of the pipe, for example in a street where each house is a point of demand. Although it is in principle possible to create a node for each access point in the systems, the number of nodes, and therefore the size of the problem, quickly grows to an intractable size. It is worthwhile to study how the number of access points in a realistic water distribution network can be dramatically reduced to a more tractable number of demand nodes in the corresponding WDND optimisation problem. Secondly, water distribution networks can be truly enormous. Solving a WDND optimisation problem with millions of demand points and many thousands of pipes is, of course, intractable. Therefore, the network on a regional or national scale needs to be decomposed into more manageable sub-networks, perhaps on different scales (e.g., neighborhood – commune – region), but an adequate methodology to do this is still lacking. Thirdly, realistic water distribution networks consist of more than reservoirs, pipes and demand nodes. Pumps, e.g., are important components of every modern network, and need to be incorporated in the WDND optimisation problem. Finally, the WDND optimisation problem described in this paper assumes that all demand is constant, which is not the case in reality. Demand patterns can be created that reflect the water demand of a family, an industrial company, a school, etc., and these demand patterns can be used to find a robust water distribution network design that has a low cost and meets all pressure constraints during each point of the day. However, working with water demand patterns increases the computational complexity of the WDND optimisation problem considerable, and further research is necessary to develop efficient methods to solve this more realistic problem efficiently.

References

A. Alperovits and U. Shamir. Design of optimal water distribution systems. *Water Resources Research*, 13(6):885–900, 1977.

R. Baños, C. Gil, J. Agulleiro, and J. Reca. A memetic algorithm for water distribution network design. *Soft Computing in Industrial Applications*, pages 279–289, 2007.

R. Baños, C. Gil, J. Reca, and F.G. Montoya. A memetic algorithm applied to the design of water distribution networks. *Applied Soft Computing*, 10:261–266, 2010.

O. Braysy and M. Gendreau. Vehicle routing problem with time windows, part ii: Metaheuristics. *Transportation Science*, 39(1):119–139, 2005.

J. Brumbelow, J. Torres, S. Guikema, E. Bristow, and L. Kanta. Virtual cities for water distribution and infrastructure system research. In *World Environmental and Water Resources Congress*, 2007.

J. Burgschweiger, B. Gnädig, and M.C. Steinbach. Optimization models for operative planning in drinking water networks. *Optimization and Engineering*, 10:43–73, 2009.

C.-W. Chu, M.-D. Lin, G.-F. Liu, and Y.H. Sung. Application of immune algorithms on solving minimum-cost problem of water distribution network. *Mathematical and Computer Modelling*, pages 1888–1900, 2008.

M. Cunha and L. Ribeiro. Tabu search algorithms for water network optimization. *European Journal of Operational Research*, 157(3):746–758, 2004.

M. Cunha and J. Sousa. Water distribution network design optimization: simulated annealing approach. *Journal of Water Resources Planning and Management*, 125(4), 1999.

M. Cunha and J. Sousa. Hydraulic infrastructures design using simulated annealing. *Journal of Infrastructure Systems*, 7(1):32–39, 2001.

G.C. Dandy, A.R. Simpson, and L.J. Murphy. An improved genetic algorithm for pipe network optimisation. *Water Resources Research*, 32(2):449–458, 1996.

F. di Pierro, S.T. Khu, D. Savic, and L. Berardi. Efficient multi-objective optimal design of water distribution networks on a budget of simulations using hybrid algorithms. *Environmental Modelling & Software*, 24(2):202–213, 2009.

M. Dorigo, V. Maniezzo, and A. Colorni. Ant system: optimization by a colony of cooperating agents. *IEEE Transactions on Systems, Man and Cybernetics, Part B: Cybernetics*, 26:29–41, 1996.

N. Duan, L.W. Mays, and K.E. Lansey. Optimal reliability-based design of pumping and distribution systems. *Journal of Hydraulic Engineering*, 116(2):249–268, 1990.

A.N. El-Bahrawy and A.A. Smith. A methodology for optimal design of pipe distribution networks. *Canadian Journal of Civil Engineering*, 14(2):207–215, 1987.

M.M. Eusuff and K.E. Lansey. Optimization of water distribution network design using the shuffled frog leaping algorithm. *Journal of Water Resources Planning and Management*, 129(3), 2003.

O. Fujiwara and D.B. Khang. A two-phase decompostion method for optimal design of looped water distribution networks. *Water Resources Research*, 26(4):539–549, 1990.

P. Galinier and A. Hertz. A survey of local search methods for graph coloring. *Computers & Operations Research*, 33(9):2547–2562, 2006.

Z.W. Geem. Optimal cost design of water distribution networks using harmony search. *Engineering Optimization*, 38(3):259–280, 2006.

Z.W. Geem. Particle-swarm harmony search for water network design. *Engineering Optimization*, 41(4):297–311, 2009.

J. Gessler. Pipe network optimization by enumeration. *Computer Applications in Water Resources*, pages 572–581, 1985.

F. Glover. Heuristics for integer programming using surrogate constraints. *Decision Sciences*, 8:156–166, 1977.

F. Glover. Tabu search - part i. *ORSA Journal on Computing*, 1:190–206, 1989.

D.E. Goldberg. *Genetic algorithms in search, optimization and machine learning.* Addison-Wesley, 1989.

I. Gupta, A. Gupta, and P. Khanna. Genetic algorithm for optimization of water distribution systems. *Environmental Modelling & Software*, 14(5):437–446, 1999.

J.H. Holland. Adaptation in natural and artificial systems. *Michigan Press*, 1975.

J.H. Holland. Genetic algorithms. *Scientific American*, 267:66–72, 1992.

E. Keedwell and S.-T. Khu. A novel evolutionary metaheuristic for the multi-objective optimization of real-world water distribution networks. *Engineering Optimization*, 38 (3):1–18, 2006.

J. Kennedy and R. Eberhart, editors. *Particle swarm optimization*, volume 4, 1995. IEEE International Conference on Neural Networks.

A. Kessler and U. Shamir. Analysis of the linear programming gradient method for optimal design of water supply networks. *Water Resources Research*, 25(7):1469–1480, 1989.

S. Kirkpatrick, C.D. Gelatt, and M.P. Vecchi. Optimization by simulated annealing. *Science*, 220:671–680, 1983.

R. Kolisch and S. Hartmann. Experimental investigation of heuristics for resource-constrained project scheduling: An update. *European Journal of Operational Research*, 174(1):23–37, 2006.

M.D. Lin, Y.H. Liu, G.F. Liu, and C.W. Chu. Scatter search heuristic for least-cost design of water distribution networks. *Engineering Optimization*, 39(7):857–876, 2007.

I. Lippai, J.P. Heaney, and M. Laguna. Robust water system design with commercial intelligent search optimizers. *Journal of Computing in Civil Engineering*, 13(3), 1999.

G.V. Loganathan, J.J. Greene, and T.J. Ahn. Design heuristic for globally minimum cost water distribution systems. *Journal of Water Resources Planning and Management*, 121(2):182–192, 1995.

B.F. Loubser and J. Gessler. Computer-aided optimization of water distribution networks. *The Civil Engineer of South Africa*, 32:413–422, 1990.

H.R. Maier, A.R. Simpson, A.C. Zecchin, W.K. Foong, K.Y.Phang, H.Y.Seah, and C.L. Tan. Ant colony optimization for design of water distribution systems. *Journal of Water Resources Planning and Management*, 129(3):200–209, 2003.

I. Montalvo, J. Izquierdo, S. Schwarze, and R. Pérez-García. Multi-objective particle swarm optimization applied to water distribution systems design: An approach with human interaction. *Mathematical and Computer Modelling*, 52:1219–1227, 2010.

P. Moscato. On evolution, search, optimization, genetic algorithms and martial arts, towards memetic algorithms. Caltech concurrent computation program 158-79, California Institute of Technology, 1989.

L.J. Murphy and A.R. Simpson. Genetic algorithms in pipe network optimisation. Research Report R93, University of Adelaide, 1992.

M. Möderl, R. Sitzenfrei, T. Fetz, E. Fleischhacker, and W. Rauch. Systematic generation of virtual networks for water supply. *Water Resources Research*, 47, 2011.

L. Perelman and A. Ostfeld. An adaptive heuristic cross-entropy algorithm for optimal design of water distribution systems. *Engineering Optimization*, 39(4):413–428, 2007.

R. Poli, J. Kennedy, and T. Blackwell. Particle swarm optimization. *Swarm intelligence*, 1(1):33–57, 2007.

G.E. Quindry, J.C. Liebman, and E.D. Brill. Optimization of looped water distribution systems. *Journal of the Environmental Engineering Division*, 107(4):665–679, 1981.

J. Reca and J. Martínez. Genetic algorithms for the design of looped irrigation water distribution networks. *Water Resources Research*, 42, 2006.

J. Reca, J. Martínez, C. Gil, and R. Baños. Application of several meta-heuristic techniques to the optimization of real looped water distribution networks. *Water Resources Management*, 22:1367–1379, 2008.

L.A. Rossman. *EPANET 2: users manual*. Environmental Protection Agency, 2000.

B. Roy, R. Słowiński, and W. Treichel. Multicriteria programming of water systems for rural areas. *Water Resources Bulletin*, 28:13–31, 1992.

R. Rubinstein. The cross-entropy method for combinatorial and continuous optimization. *Methodology and Computing in Applied Probability*, 1:127–190, 1999.

D. Savic and G. Walters. Genetic algorithms for least-cost design of water distribution networks. *Journal of Water Resources Planning and Management*, 123(2):67–77, 1997.

J. Schaake and D. Lai. Linear programming and dynamic programming applications to water distribution network design. Report 116, Department of Civil Engineering, MIT, Cambridge USA, 1969.

U. Shamir. Optimal design and operation of water distribution systems. *Water Resources Research*, 10(1):27–36, 1974.

A.R. Simpson, G.C. Dandy, and L.J. Murphy. Genetic algorithms compared to other techniques for pipe optimization. *Journal of Water Resources Planning and Management*, 120(4):423–443, 1994.

R. Sitzenfrei. *Stochastic generation of urban water systems for case study analysis*. Dissertation, Leopold Franzens Universität Innsbruck, Innsbruck, 2010.

R. Słowiński. A multicriteria fuzzy linear programming method for water supply system development planning. *Fuzzy Sets and Systems*, 19:217–237, 1986.

R. Storn and K. Price. Differential evolution - a simple and efficient heuristic for global optimization over continuous spaces. *Journal of Global Optimization*, 11:341–359, 1997.

T. Stützle and H.H. Hoos. Max-min ant system. *Future Generation Computer Systems*, 16:889–914, 2000.

W.Y. Szeto, Y. Wu, and S.C. Ho. An artificial bee colony algorithm for the capacitated vehicle routing problem. *European Journal of Operational Research*, 215(1):126–135, 2011.

K. Sörensen and F. Glover. Metaheuristics. In S. Gass and M. Fu, editors, *Encyclopedia of OR/MS, 3rd edition*, New York, To appear. Springer.

E. Todini and S. Pilati. A gradient method for the analysis of pipe networks. In *International Conference on Computer Applications for Water Supply and Distribution*. Leicester Polytechnic, 1987.

K. Vairavamoorthy and M. Ali. Optimal design of water distribution systems using genetic algorithms. *Computer-Aided Civil and Infrastructure Engineering*, pages 374–382, 2000.

A. Vasan and S.P. Simonovic. Optimization of water distribution network design using differential evolution. *Journal of Water Resources Planning and Management*, 136(2): 279–287, 2010.

D. Verleye and E.H. Aghezzaf. Modeling and optimization of production and distribution of drinking water at vmw. *Network Optimization*, pages 315–326, 2011.

J.P. Watson, J.C. Beck, A.E. Howe, and L.D. Whitley. Problem difficulty for tabu search in job-shop scheduling. *Artificial Intelligence*, 143(2):189–217, 2003.

J.P. Watson, A.E. Howe, and L. Darrell Whitley. Deconstructing nowicki and smutnicki's i-tsab tabu search algorithm for the job-shop scheduling problem. *Computers & Operations Research*, 33(9):2623–2644, 2006.

Z.Y. Wu and A.R. Simpson. Competent genetic-evolutionary optimization of water distribution systems. *Journal of Computing in Civil Engineering*, 15, 2001.

A.C. Zecchin, A.R. Simpson, H.R. Maier, and J.B. Nixon. Parametric study for an ant algorithm applied to water distribution system optimization. *IEEE Transactions on Evolutionary Computation*, 9(2):175–191, 2005.

A.C. Zecchin, A.R. Simpson, H.R. Maier, M. Leonard, A.J. Roberts, and M.J. Berrisford. Application of two ant colony optimisation algorithms to water distribution system optimisation. *Mathematical and Computer Modelling*, 44:451–468, 2006.

Acknowledgements

This research was partially supported by the Fund for Scientific Research Flanders (FWO).