Optimising water distribution systems using a weighted penalty in a genetic algorithm

M van Dijk[1]*, SJ van Vuuren[1] and JE van Zyl[2]

[1] *Department of Civil & Biosystems Engineering, University of Pretoria, Pretoria 0001, South Africa*
[2] *Department of Civil Engineering Science, University of Johannesburg, PO Box 524, Auckland Park 2006, South Africa*

Abstract

Genetic algorithms (GAs) have become the preferred water system design optimisation technique for many researchers and practitioners. The main reason for using GAs is their ability to deal with nonlinear complex optimisation problems. The optimal decision in terms of designing, expansion/extending, addition or rehabilitation of water supply systems has to review possible options and select a cost-effective and efficient solution. This paper presents a new approach in determining a penalty value depending on the degree of failure, of the set pressure criteria, and the importance of the link supplying a specific node. Further modifications are also made in the cross-over and mutation procedures to ensure an increase in algorithm convergence. EPANET, a widely used water distribution network simulation model, is used in conjunction with the proposed newly developed GA for the optimisation of water distribution systems. The developed GA procedure has been incorporated in a software package called GANEO, which can be used to design new networks, analyse existing networks and prioritise improvements on existing networks. The developed GA has been tested on several international benchmark problems and has proved to be very efficient and robust. The EPANET hydraulic modelling software as well as the developed GANEO software, which performs the optimisation of the water distribution network, is freeware. The software provides a tool for consulting engineers to optimise the design or rehabilitation of a water distribution network.

Keywords: optimising, water distribution system, genetic algorithm

Introduction

As a vital part of water supply systems, water distribution networks represent one of the largest infrastructure assets of industrial society. Simulation of hydraulic behaviour within a pressurised, looped pipe network is a complex task, which effectively means solving a system of non-linear equations. The South African objective to provide 'water for all' makes it essential that the limited capital has to be employed in such a way to provide the maximum benefit. Optimising a relatively small system will require numerous repetitive calculations. The discrete nature of the network optimisation problem (pipe diameters) and the size of the solution space make it virtually impossible to apply any of the conventional optimisation techniques to find the global optimum.

The development of hydraulic models in the last two decades improved the ability to simulate hydraulic behaviour of large water distribution networks (Rossman, 2000). Most optimisation techniques are applicable when continuous variables are evaluated. GAs are applicable when a large solution space has to be searched, consisting of discreet variables, and it is now accepted by most experts that the GA is the best technique for network optimisation (Van Vuuren et al., 2005).

According to Michalewicz (1994), GAs can basically be described as artificial evolution search methods based on the theories of natural selection and mechanisms of population genetics. GAs apply the principle of survival of the fittest in a mathematical sense. In the evolutionary process, all species develop in such a way to improve their chances of survival and quality of living. It therefore means that all species are striving to a certain optimum, being physical, behavioural or otherwise. In this paper a traditional GA is developed which incorporates a unique penalty structure and tailored cross-over and mutation procedures.

Problem formulation

The difficulty of optimising water distribution systems is mainly due to the discrete nature of the variables and the size of the solution space. Many optimisation techniques can only be applied to problems, which have continuous variables, unlike the pipe diameter variables in the network optimisation problem. The size of the solution space (the total number of possible solutions to the problem) for the network optimisation problem can be calculated as the number of possible discrete pipe diameters to the power of the number of pipes in the network.

A major problem for water supply authorities is to select those components in a network that should be changed, increased or replaced to ensure a sustainable service to the consumers at the lowest cost. Water supply authorities are also faced with problems other than that of network design such as network calibration, operation and reliability. For each type of optimisation, the main objective function, possible variables and the main constraints are summarised in Table 1.

A water supply distribution system consists of a complex network of interconnected pipes, service reservoirs and pumps that deliver water from the treatment plant to a consumer. The distribution of water through the network is governed by complex, non-linear, non-convex and discontinuous hydraulic equations (Keedwell and Khu, 2005).

*   To whom all correspondence should be addressed.
☎ +2712 420 3176; fax: +2712 362 5218;
    e-mail: marco.vandijk@up.ac.za

TABLE 1
Types of optimisation for water distribution systems (Van Vuuren et al., 2005)

Optimisation type	Objective	Possible variables	Main constraints
Design	Minimise cost	Pipe layout; pipe diameters; pipe rehabilitation	Min. level of service; available diameters; rehabilitation options; available budget; LCC
Operation	Minimise operational cost	Pump controls; reservoir levels; sources and capacity	Min. level of service; number of pump switches; source capacity; pump capacity
Calibration	Minimise difference between model and observed values	Pipe roughness; pipe diameter; valve settings; leakage; demands	System layout; available data
Level-of-service	Maximise level of service, e.g. pressure, water quality or reliability	All above	System configuration; budget
Monitoring system design	Minimise cost of monitoring system	Number and position of monitoring points	System configuration; budget
Network testing	Find critical sets of events that may cause a system to fail.	Fires; pipe failures; power failures; contamination events	System configuration; number of simultaneous events

Two equations, which are used to determine if a network is hydraulically balanced, are the continuity and energy equations (Eqs. (1) and (2) respectively).

$$\sum_{i=1}^{n} q_i = 0 \qquad (1)$$

The continuity equation is applied to each node with q_i the flow rate (in and out of the node) and n the number of pipes joined at the node.

$$\sum_{i=1}^{m} h_i = 0 \qquad (2)$$

The energy equation is applied to each loop in the network with h_i the head loss in each pipe and m the number of pipes in the loop. The head loss is the sum of the local head losses and the friction head losses. The friction head loss can be calculated using the Darcy-Weisbach (Eq. (3)) or Hazen-Williams (Eq. (4)) empirical equations or something similar.

$$h_f = \frac{\lambda L V^2}{2gD} \qquad (3)$$

$$h_f = \frac{6.8193L}{D^{1.167}} \left(\frac{V}{C}\right)^{1.852} \quad or \quad h_f = \frac{\omega L Q^{1.852}}{C^{1.852} D^{4.871}} \qquad (4)$$

where:
L = length (m)
D = internal diameter (m)
V = velocity (m/s)
g = gravitational acceleration (m/s²)
λ = Darcy Weisbach friction factor
C = Hazen-Williams friction factor
ω = numerical conversion constant (in this paper $\omega = 10.667$)

The network can be hydraulically balanced utilising Eqs. (1) to (4) and methods such as the Hardy-Cross and nodal methods. The aim of designing a new network is to obtain a system that will meet the demand at each node at a required minimum pressure. Similar for the rehabilitation or improvement of a network, the optimum system is achieved when the system components are identified, to replace or rehabilitate, which will provide the level of service required.

Genetic algorithm

Genetic algorithms (GAs) have been developed (Holland, 1975) to assist in searching through complex solution spaces for the optimum solution. GAs have been applied as search techniques for various engineering problems such as, structural design optimisation, water distribution network evaluation, pump scheduling, hydrological runoff predictions and resource utilisation. According to Michalewicz (1994), GAs can basically be described as artificial evolution search methods based on the theories of natural selection and mechanisms of population genetics. GAs emulate nature's optimisation technique of evolution, based on:
• Survival and reproduction of the fittest members of the population
• The maintenance of a population with diverse members
• The inheritance of genetic information from parents
• The occasional mutation of genes.

A GA evolves optimal solutions by sampling from all the possible solutions. The best of these solutions are then combined, using the genetic operators of cross-over and mutation, to form new solutions. The identification of these best solutions is done based on a set objective function. This process continues until some termination condition is fulfilled. A flow diagram of the basic GA process is given in Fig. 1.

The objective function in the optimisation of a water distribution system is usually the minimisation of total cost. The total actual cost is a combination of the capital costs and operating and maintenance costs. In this paper only the capital cost of the pipes (supply, lay and jointing) are considered and hence it can be generally expressed as:

$$f_{cost}(D) = \sum_{j=1} K L_j D_j^n \qquad (5)$$

The pipe costs per unit length usually vary nonlinearly with its diameter and in Eq. (5) it is assumed that it can be expressed as a single term for all diameters, where $f_{cost}(D)$ is the cost of the pipes, L_j and D_j are the lengths and diameters of the jth pipe and K and n are constants that will depend on local conditions (Vairamoorthy and Ali, 2000).

Figure 1
Basic genetic algorithm process
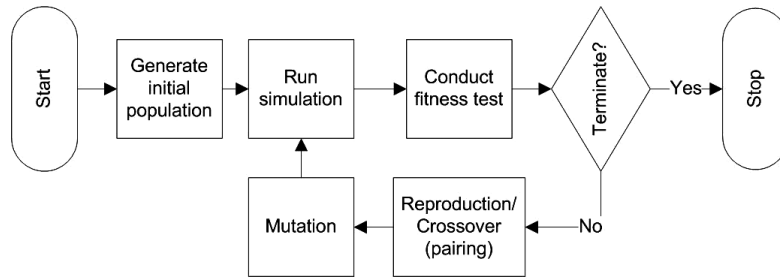
Genetic algorithm optimisation model

GAs can be used to optimise various different parameters in water distribution systems. Various types of optimisation can be identified for water distribution systems as shown in Table 1. In many cases the types and subsequent objectives of optimisations are in conflict with each other. For example, attempts to minimise operational cost will generally place the system in a more vulnerable state and less able to handle abnormalities such as pipe bursts, thus reducing the level of service (Jowitt et al., 1988). In such cases it is necessary to strike a balance between the objectives. This can be done by either defining the balance before hand (for example using an objective function equally weighed for cost and level of service) or by using multi-objective optimisation. In this paper only a single objective, in this case minimising cost, will be set for a water distribution system (new or upgrading/rehabilitation). The genetic algorithm process with its newly developed penalty function is summarised in Table 2.

Benchmarking the model

The developed genetic algorithm optimisation model was tested against benchmark problems in order to establish its functionality and efficiency. Software was developed (called GANEO) and

TABLE 2
Genetic algorithm process for optimising water distribution systems with weighted penalty function

Process	Description
Create the initial population	Make a random selection of pipe diameters, from a selected list of available pipes, for the pipe network to create a string (possible network solution). Repeat this process to generate the entire population of network solutions.
Hydraulic analysis	Perform a hydraulic analysis on each of the population's strings (using hydraulic modelling package such as EPANET) to determine the pressure and supply at each node in the network as well as the flow rate in each pipe.
Fitness of each string (solution)	If a node does not meet the minimum pressure requirement the pipes supplying that node are penalised. If nodes have negative pressures the pipes supplying these nodes are penalised extensively to emphasise the poor results thereof. The cost of a pipe that results in a node not meeting the minimum pressure requirement will be calculated by modifying Eq. (5) as follows (Eq. (6)): $$PC_{Hmin-j} = C_P^{1+b} \ x \ L_P \quad \text{and} \quad b = PF\left[\frac{P_{required} + \left(P_{required} - P_{calculated}\right)}{P_{required}}\right]\left[\frac{Q_j}{Q_{node}}\right]$$ where: PC_{Hmin-j} = Cost of pipe 'j' with added penalty cost due to minimum pressure constraint C_P = Cost of pipe per unit length, which is KD_j^n L_P = Length of the installed pipe b = Penalty factor (b = 5 if $P_{calculated} < 0$) PF = User specified penalty factor (0.5 to 10) $P_{calculated}$ = Calculated pressure at node $P_{required}$ = Minimum residual pressure required at node Q_j = Flow in pipe 'j' Q_{node} = Total flow into node The aim of the weighted penalty cost structure as defined above is to increase the penalty on a system, the greater the pressure deficiency is and to add some proportional distribution of the importance of the supply pipe based on the flows in the pipes, to the cost. The more water a specific pipe supplies to the node the greater the importance of that pipe. The higher the user-specified penalty factor (PF) is the higher the cost component will be. The pressure penalties are subject to an if-then-else statement, which means that if the pressures fall within the specified boundaries, no penalties would be applied. The total cost of the network is the total cost of all the individual pipes (including penalties). A similar approach is followed in case a velocity criterion is not met.
Reproduction and cross-over (pairing)	In this proposed model, 75% of the top ranked solutions of the generation is retained and the worst solutions (25%) are discarded. A new set of strings (offspring) is generated from the remaining strings/solutions based on probabilities of their fitness values. Through a random process, or the spin of the roulette wheel, the new strings for the new generation are created. Thereafter a single point cross-over where the genes of the strings are transferred between parents is performed. The selection of parents for cross-over and determining the position of cross-over in each of the pairs is again a random process (although the developed software allows for other cross-over procedures).
Mutation	To force the solution to include gene strings from the total solution space and to steer away from the local optimum a mutation is performed with a probability equal to the mutation rate. Each gene (pipe) of each string (network solution) has in other words the probability of mutating and being replaced with a randomly selected gene from the available gene pool.
Termination	Following the selection, cross-over and mutation operators and introduction of the new child organisms into the population, the process is repeated until an appropriate termination condition is met. The simplest technique is to use a fixed number of generations or alternatively when complete convergence has occurred or no improvement in the fitness value of the best chromosome has occurred in some fixed number of the previous generations.

used to test common networks for which many optimisations have been performed; these include traditional and heuristic methods (Savic and Walters, 1997). Three systems were tested: New York Tunnels (Example 1), Hanoi (Example 2) and Two-loop network (Example 3) detailed below.

In all three examples the Hazen-Williams equation (Eq. (4) will be utilised to determine the friction loss in a pipe link between two nodes. Previous researchers have investigated these systems extensively and obtained numerous solutions that met the defined fitness function of minimum cost based on the constraint of required pressure and demand at every node. Due to different interpretations of the Hazen-Williams equation different researchers obtained different solutions (Savic and Walters, 1997) which made direct comparison not always possible. According to Savic and Walters (1997) the numerical conversion constant (ω), see Eq.(4), varied from 10.5088 to 10.9031. The consequence of this variation of the ω values used is that systems designed with $\omega = 10.5088$ calculate a lesser friction head loss when compared to $\omega = 10.9031$. This result in solutions meeting the set pressure criteria when analysed with the lower boundary of ω but when reanalysed with the upper boundary failing and thus providing an unfeasible solution. The value of ω as used in the EPANET software, which was used in the hydraulic analyses, is 10.667 (similar to Dandy et al. (1996), Montesinos et al. (1999), Wu and Simpson (2002) and Keedwell and Khu (2005)).

All three benchmark problems were analysed utilising a Pentium computer wit a 3.2 GHz Intel processor, 1024 MB RAM, using Microsoft Windows XP Professional as operating system. The GANEO software utilised, an example of an improvement to a distribution network (Example 1) and two examples of designing of new networks (Examples 2 and 3) are described in the following paragraphs.

GANEO software

The developed GA optimisation model as described above required the development of a software program to perform the computations. The developed program called GANEO requires six easy steps to optimise a water distribution system:

Step 1: Using EPANET, create a working network model and export the water distribution system into the correct format for importing in GANEO (*.inp file).

Step 2: Create a new project in GANEO and import the EPANET model (see Fig. 2).

Step 3: Create a pipe selection file in GANEO from which pipes (genes) will be selected for population of the network (strings).

Step 4: Set the GA parameters (cross-over method, mutation percentage, etc.) and boundaries (penalties) in GANEO.

Step 5: Run the GA optimisation analysis procedure (see Fig. 3).

Step 6: Evaluate the alternative options and export the result back to EPANET.

Example 1: New York tunnels

The New York water supply system has been studied by a number of researchers in the pipe network optimisation field. The aim of the various studies was to determine the most economical design to improve the existing system of tunnels that constituted the main water distribution system. Figure 4 is a general layout of the system indicating the pipes (Table 3), nodes (Table 4) and single supply reservoir. The existing system was found to be inadequate due to ageing and an increase in demands, in terms of the pressure requirements.

Figure 2
GANEO input screen
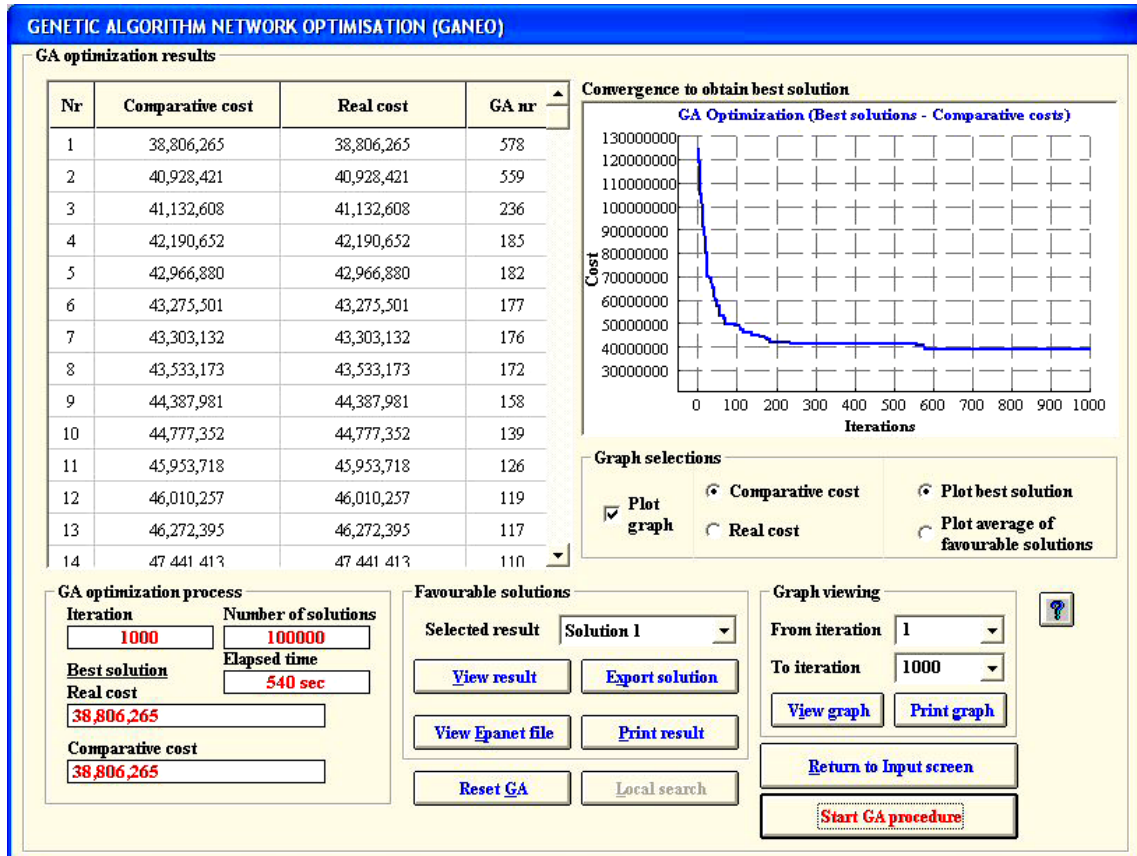
Figure 3
GANEO results screen
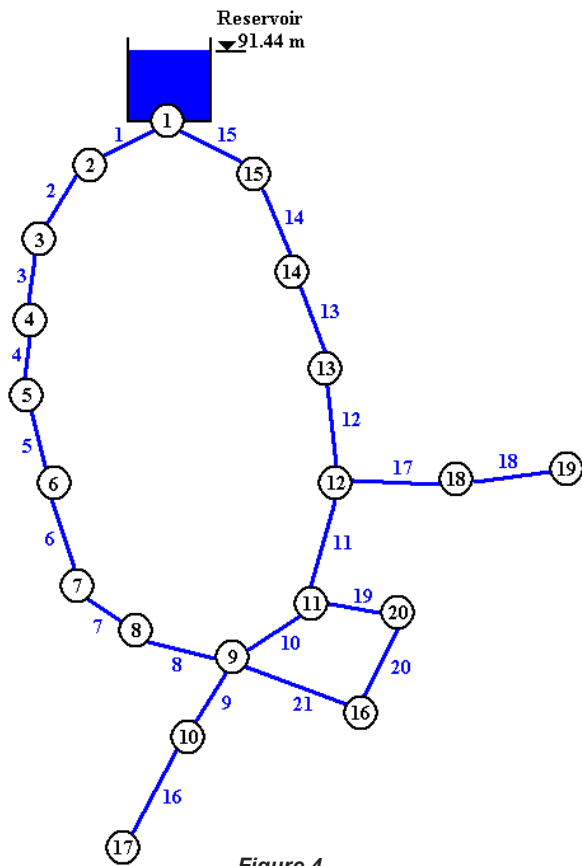
Figure 4
General layout of New York tunnels water supply system

Pipe	Length (m)	Diameter (m)
1	3 536.59	4.57
2	6 036.59	4.57
3	2 225.61	4.57
4	2 530.49	4.57
5	2 621.95	4.57
6	5 823.17	4.57
7	2 926.83	3.35
8	3 810.98	3.35
9	2 926.83	4.57
10	3 414.63	5.18
11	4 420.73	5.18
12	3 719.51	5.18
13	7 347.56	5.18
14	6 432.93	5.18
15	4 725.61	5.18
16	8 048.78	1.83
17	9 512.20	1.83
18	7 317.07	1.52
19	4 390.24	1.52
20	11 707.32	1.52
21	8 048.78	1.83

TABLE 3
Pipe data of the existing system/network

TABLE 4		
Node data of the existing system/network		
Node	Minimum required pressure level (m)*	Demand (m³/s)
1	91.44	Reservoir
2	77.72	2.62
3	77.72	2.62
4	77.72	2.50
5	77.72	2.50
6	77.72	2.50
7	77.72	2.50
8	77.72	2.50
9	77.72	4.81
10	77.72	0.03
11	77.72	4.81
12	77.72	3.32
13	77.72	3.32
14	77.72	2.62
15	77.72	2.62
16	79.25	4.81
17	83.15	1.63
18	77.72	3.32
19	77.72	3.32
20	77.72	4.81

*   *The ground elevation is 0.0 m.*

TABLE 5			
Available pipes from which the pipe selection will be made			
Pipe no	Diameter (mm)	Unit pipe cost ($/m)	Reference diameter (inches)
1	0.0	0.00	0
2	914.4	306.75	36
3	1 219.2	439.63	48
4	1 524.0	577.43	60
5	1 828.8	725.07	72
6	2 133.6	875.98	84
7	2 438.4	1 036.75	96
8	2 743.2	1 197.51	108
9	3 048.0	1 368.11	120
10	3 352.8	1 538.71	132
11	3 657.6	1 712.60	144
12	3 962.4	1 893.04	156
13	4 267.2	2 073.49	168
14	4 572.0	2 260.50	180
15	4 876.8	2 447.51	192
16	5 181.6	2 637.80	204

The method utilised to improve the system was to lay parallel pipes between certain nodes. There are 15 available commercial diameters (Table 5), which could be used as well as a so-called 'do nothing' option for each of the 21 pipes in the system. The cost per unit length as shown in Table 5 is similar to other researchers' defined cost functions (see Eq. (5) with L = 1 m, K = 0.06537, n = 1.24 and D measured in mm).

A Hazen-Williams friction factor of $C_H = 100$ is assumed for both the old tunnels and the new pipes. It has been indicated that
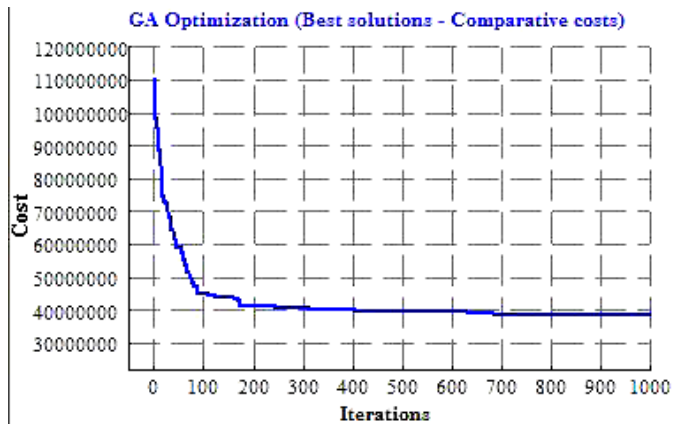
Figure 5
Convergence to optimum solution (New York Tunnels)

the conversion from imperial units to metric units could result in small differences and subsequent changes in optimum solutions (Savic and Walters, 1997).

The only system constraint that this network has is that the minimum head at each node should be as indicated in Table 4. Although the system is fairly unsophisticated and small in comparison to the internal distribution system there are 1.93 x E25 or 16[21] possible solutions for this system. It is thus impossible to analyse every single network improvement alternative.

The optimum obtained with the proposed GA procedure as used in the GANEO program was $M38.65. The number of iterations (generations) it took to obtain this optimum was 684. The convergence to the optimum is shown in Fig. 5. As can be seen the initial reduction in total cost of the system occurs fast after which this improvement process slows down.

The nodes with its pressure closest to the minimum required are Nodes 16, 17 and 19 with pressure elevations of 79.26 m, 83.16 m and 77.73 m respectively. The five best solutions, based on randomly selected initial seed values, and their costs are shown in Table 6. The following GA parameters were used: number of generations = 1 000, population = 100, penalty factor = 5 - 6 and mutation rate was set at 3%.

TABLE 6			
Optimisation results of the New York tunnel system			
Initial seed	Total cost ($M)	Optimum obtained during generation	Time to reach optimum solution (sec)
1	38.648	684	379
2	39.676	592	328
3	39.790	980	543
4	39.953	620	344
5	40.891	787	436
Average	39.792	733	406

These analyses in Table 6 compare favourably with the average cost of $M39.57 after 44 280 evaluations obtained by Montesinos et al. (1999) using different parameter sets and constraints. A direct comparison of the optimum solution obtained with the developed GA procedure and that obtained by other researchers is shown in Table 7.

As indicated by Savic and Walters (1997), GAs are stochastic-search techniques and the solution found will not always be

|	TABLE 7											
	Comparison with previous researchers' results											
Researcher	Schaake and Lai (1969)	Quindry et al. (1981)	Gessler (1985)	Bhave (1985)	Morgan and Goulter (1985)	Dandy et al. (1996)	Savic and Walters (1997)	Savic and Walters (1997)	Monte-sinos et al. (1999)	Vairava-moorthy & Ali (2000)	Wu & Simpson (2002)	Authors (2008)
ω		10.9031				10.6792	10.5088	10.9031	10.667	10.5088	10.6792	10.667
Pipe	Diameter (inches)											
1	52.02	0.00	0.00	0.0	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
2	49.90	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
3	63.41	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
4	55.59	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
5	57.25	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
6	59.19	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
7	59.06	0.00	100.00	0.00	144.00	0.00	108.00	0.00	0.00	96.00	0.00	144.00
8	54.95	0.00	100.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
9	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
10	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
11	116.21	119.02	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
12	125.25	134.39	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
13	126.87	132.49	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
14	133.07	132.87	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
15	126.52	131.37	0.00	136.43	0.00	120.00	0.00	144.00	120.00	0.00	120.00	0.00
16	19.52	19.26	100.00	87.37	96.00	84.00	96.00	84.00	84.00	100.00	84.00	96.00
17	91.83	91.71	100.00	99.23	96.00	96.00	96.00	96.00	96.00	96.00	96.00	96.00
18	72.76	72.76	80.00	78.17	84.00	84.00	84.00	84.00	84.00	84.00	84.00	84.00
19	72.61	72.64	60.00	54.40	60.00	72.00	72.00	72.00	72.00	72.00	72.00	72.00
20	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
21	57.82	54.97	80.00	81.50	84.00	72.00	72.00	72.00	72.00	72.00	72.00	72.00
Cost ($M)	**78.09**	**63.58**	**41.80**	**40.18**	**39.20**	**38.80**	**37.13**	**40.42**	**38.80**	**37.10**	**38.80**	**38.65**
Achieved at evaluation number	-	-	-	-	-	125 000	1 000 000*	1 000 000*	18 300	80 000	22 500	68 400

Number of generations allowed (exact evaluation number when optimum was achieved is unknown)

the same and a number of runs are required to ensure that the solutions identified are of good quality.

Example 2: Hanoi network

The Hanoi system, Vietnam, is a new network that should be optimised. It has a single fixed head source at elevation of 100 m. There are 34 pipes and 32 nodes as shown in Fig. 6 and listed in Tables 8 and 9.
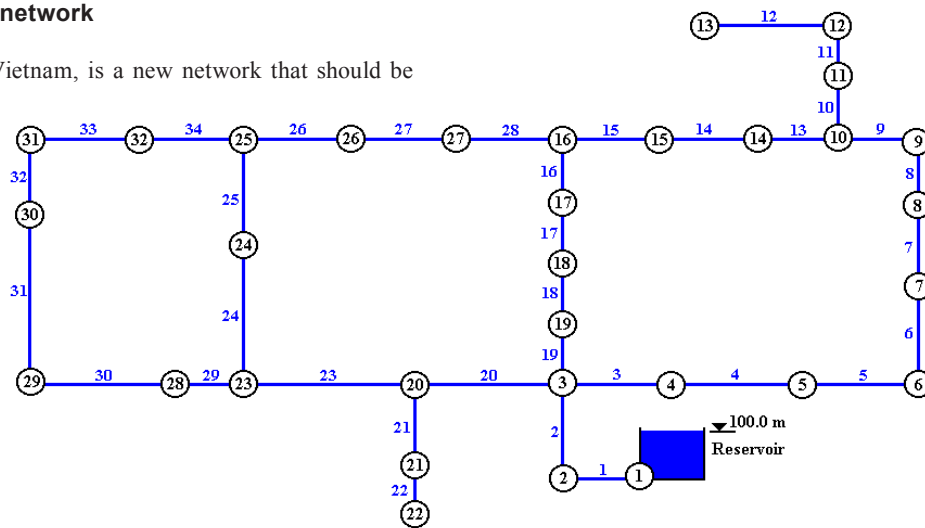
Figure 6
General layout of Hanoi water distribution network

TABLE 8 Pipe lengths of the existing network			
Pipe	Length (m)	Pipe	Length (m)
1	100	18	800
2	1 350	19	400
3	900	20	2 200
4	1 150	21	1 500
5	1 450	22	500
6	450	23	2 650
7	850	24	1 230
8	850	25	1 300
9	800	26	850
10	950	27	300
11	1200	28	750
12	3 500	29	1 500
13	800	30	2 000
14	500	31	1 600
15	550	32	150
16	2 730	33	860
17	1 750	34	950

TABLE 9 Node data of the existing network			
Node	Demand (m³/h)	Node	Demand (m³/h)
2	890	18	1 345
3	850	19	60
4	130	20	1 275
5	725	21	930
6	1 005	22	485
7	1 350	23	1 045
8	550	24	820
9	525	25	170
10	525	26	900
11	500	27	370
12	560	28	290
13	940	29	360
14	615	30	360
15	280	31	105
16	310	32	805
17	865		

The ground elevation of all the nodes is 0.0 m. The pipes that could be utilised in the design of the system are shown in Table 10.

TABLE 10 Available pipes from which the pipe selection will be made			
Nr	Diameter (inches)	Diameter (mm)	Unit cost (per meter length)
1	12	304.8	45.73
2	16	406.4	70.40
3	20	508.0	98.39
4	24	609.6	129.33
5	30	762.0	180.75
6	40	1 016.0	278.28

A Hazen-Williams friction factor of $C_H = 130$ is assumed for all new pipes

The head constraint for this system is 30 m, i.e. the pressure everywhere in the system should be greater than 30 m. The costs as indicated in Table 10 were determined with the cost function (Eq. 5) with K = 0.008593, n = 1.5 and D measured in millimetre.

The Hanoi network has been optimised by other researchers as shown in Table 11. Fujiwara and Khang (1990) did not use discrete values for the pipe diameters and for this reason direct comparison with their solution is not possible.

The following GA parameters were used: number of generations = 2 000, population = 100, penalty factor = 2 and mutation rate was set at 3%. The optimum obtained with the GANEO program was $M6.110. The number of iterations (generations) it took to obtain this optimum was 495 and it was achieved after 324 s.

According to Savic and Walters (1997) the solution of Fujiwara and Khang (1990) is based on a continuous cost function solution since their method could not handle discontinuous objective functions directly. Furthermore when the Fujiwara and Khang (1990) solution is reanalysed with the range values of ω = 10.5088 and ω =10.9031 it does not meet the minimum 30 m pressure requirement with pressures as low as 10.31 m and 7.69 m respectively.

As can be seen in Fig. 7 the initial reduction in total cost of the system occurs rapidly after which the improvement process slows down.
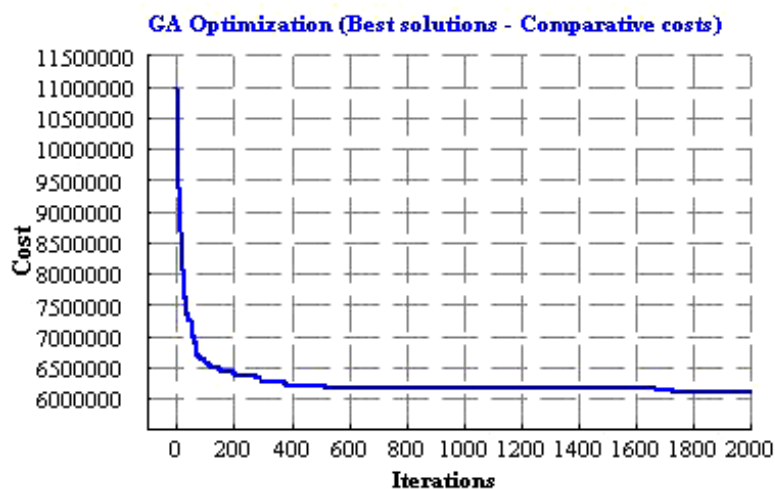
Figure 7
Convergence to optimum solution (Hanoi water distribution network)

TABLE 11 **Comparison with previous researcher's results (Hanoi water distribution system)**						
Researcher	Fujiwara and Khang (1990)	Savic and Walters (1997)		Vairava-moorthy and Ali (2000)	Liong et al. (2004) *	Wu and Walski (2005)	Authors (2008)
ω	10.5088	10.5088	10.9031	10.5088	10.667	10.667#	10.667
Pipe	Diameter (inches)						
1	40	40	40	40	40	40	**40**
2	40	40	40	40	40	40	**40**
3	38.8	40	40	40	40	40	**40**
4	38.7	40	40	40	40	40	**40**
5	37.8	40	40	40	40	40	**40**
6	36.3	40	40	40	40	40	**40**
7	33.8	40	40	40	40	40	**40**
8	32.8	40	40	40	30	40	**40**
9	31.5	40	30	40	30	40	**40**
10	25	30	30	30	3	30	**30**
11	23	24	30	24	30	24	**24**
12	20.2	24	24	24	24	24	**24**
13	19	20	16	20	16	20	**24**
14	14.5	16	16	16	12	16	**12**
15	12	12	12	12	12	12	**12**
16	19.9	12	16	12	24	12	**12**
17	23.1	16	20	16	30	16	**16**
18	26.6	20	24	20	30	20	**24**
19	26.8	20	24	20	30	20	**24**
20	35.2	40	40	40	40	40	**40**
21	16.4	20	20	20	20	20	**20**
22	12	12	12	12	12	12	**12**
23	29.5	40	40	40	30	40	**40**
24	19.3	30	30	30	30	30	**30**
25	16.4	30	30	30	24	30	**30**
26	12	20	20	20	12	20	**20**
27	20	12	12	12	20	12	**12**
28	22	12	12	12	24	12	**12**
29	18.9	16	16	16	16	16	**16**
30	17.1	16	16	12	16	12	**12**
31	14.6	12	12	12	12	12	**12**
32	12	12	12	16	16	16	**20**
33	12	16	16	16	20	16	**16**
34	19.5	20	20	24	24	24	**24**
Cost ($M)	**5.354**	6.073	6.187	6.056	6.220	6.056#	**6.110**

Shuffled complex evolution method was used

Assumed value of ω since EPANET was used

Vairavamoorthy and Ali obtained the optimum solution after approximately 25 min. Liong and Atiquzzaman analysis took approximately 11 min whilst the authors obtained the optimum after generation 495 and 6.4 min. The minimum pressure in the system was 30.045 m at node 32.

As detailed in Table 2 a new weighted penalty function is proposed by the authors. Wu and Walski (2005) provide a comparison of various constraint-handling techniques and this is reproduced in Table 12 (including the author's results).

As can be seen in Table 12 (next page) the proposed GA produced satisfactory results compared to others. Furthermore when the optimum Wu and Walski solution, Table 11, is reanalysed with ω = 10.667 as used in EPANET it does not meet the minimum pressure requirement of 30 m. Direct comparison is thus difficult
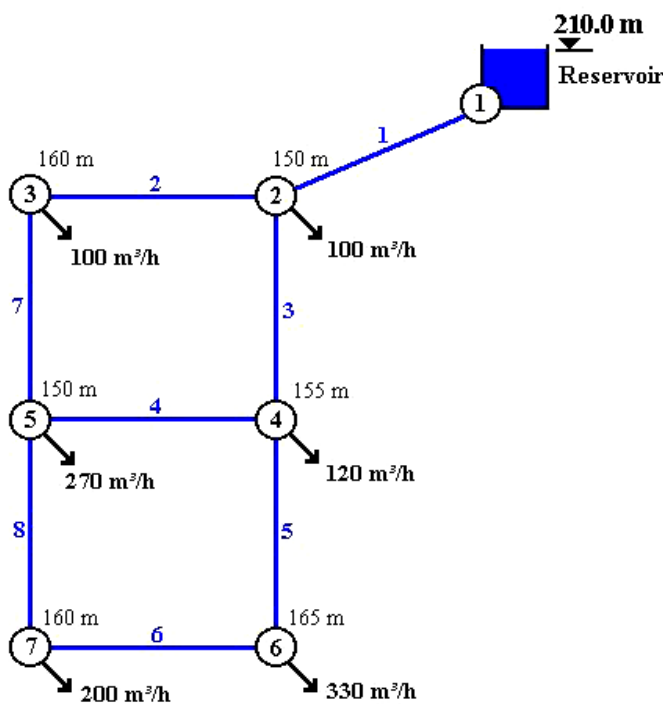
but the solutions obtained by the authors are competitive.

Example 3: Two-loop network

The two-loop network was first studied by Alperovits and Shamir (1977) and many others thereafter (Keedwell and Khu, 2005). The two-loop network consists of eight pipes, which are fed from a single fixed head reservoir to supply the demands as shown in Fig. 8 (next page).

The pipes in the network are all 1 000 m long and the only system constraint is the minimum pressure requirement for nodes 2 to 7 defined as 30 m. The available pipe diameters and costs that could be used in the design of the system are shown in Table 13.

TABLE 12
Comparison of weighted penalty with other constraint-handling techniques (Cost in 10^6)

Random seeds	Dynamic penalty (Joins and Houck, 1994)	Niched penalty (Deb and Agraval, 1999)	Self-organising penalty (Lin and Wu, 2004)	Self-adaptive penalty (Wu and Walski, 2005)	Weighted penalty (Authors, 2008)
0.1	6.56	6.3	6.40	6.16	6.15
0.2	6.29	6.58	6.41	6.19	6.15
0.3	6.34	6.11	6.15	6.06	6.15
0.4	6.61	6.35	6.60	6.56	6.15
0.5	6.38	6.3	6.39	6.30	6.31
0.6	6.34	6.27	6.40	6.23	6.11
0.7	6.64	6.20	6.58	6.24	6.35
0.8	6.42	6.27	6.27	6.16	6.38
0.9	6.41	6.27	6.30	6.30	6.33
Best	**6.29**	**6.11**	**6.15**	**6.06**	**6.11**
Average	**6.44**	**6.29**	**6.39**	**6.24**	**6.23**
Worst	**6.64**	**6.58**	**6.60**	**6.56**	**6.38**

Figure 8
Two-loop network

TABLE 13
Available pipes (Two-loop network)

Nr	Diameter (inches)	Diameter (mm)	Unit cost (per meter length)
1	1	25.4	2
2	2	50.8	5
3	3	76.2	8
4	4	101.6	11
5	6	152.4	16
6	8	203.2	23
7	10	254.0	32
8	12	304.8	50
9	14	355.6	60
10	16	406.4	90
11	18	457.2	130
12	20	508.0	170
13	22	558.8	300
14	24	609.6	550

The objective of the system was to determine the required pipe diameters that would yield the least total cost whilst still supplying the demand and adhering to the system constraint of minimum pressure at each node. Although the system seems extremely simple there are still 8^{14} possible combinations of pipes.

Savic and Walters (1997) found two solutions consistently at 419 000 and 420 000 cost units (depending on ω) which satisfied the demand and pressure requirements. According to Keedwell and Khu (2005) the algorithm was ran for 500 generations with a population size of 50, i.e. 25 000 network simulations. Using a cellular automata (CA), genetic algorithm (GA) combination approach Keedwell and Khu (2005) also analysed the two-loop network. The Cellular Automation for Network Design Algorithm combined with a GA (called by the authors CANDA-GA)

can be described as seeding the GA with CA solutions, i.e. providing a better initial population to start with. The optimisation results of Savic and Walters (1997), Keedwell and Khu (2005) and the authors are presented in Table 14.

As discussed by Savic and Walters (1997) other researchers also provided solutions in this range such as Kessler and Shamir (1989) with 402 352 and Eiger et al. (1994) with 417 500 but these solutions did not obtain the minimum pressure requirement of 30 m as every node. Some other researchers such as Goulter et al. (1986) with a cost of 435 015 and Alperovits and Shamir (1977) with a cost of 497 525 obtained feasible solutions but utilised split pipe solutions.

Keedwell and Khu (2005) reported that both algorithms (GA and CANDA-GA) fitness converged after 3 000 generations. If this is compared with what was obtained with GANEO it can be clearly seen that an optimum or near optimum is very quickly obtained. The results furthermore show an improvement of the average total cost of the five runs (different initial seed values) that were performed. The worst result obtained with GANEO was 9 000 cost units better than that obtained by Keedwell and Khu (2005) with the standard GA and 2 000 cost units better than that obtained with the CANDA-GA.

TABLE 14 Optimisation results of the two-loop network					
Researcher	Savic & Walters (1997)		Keedwell and Khu (2005)		Authors (2008)
	GA1	GA2	GA	CANDA-GA	GANEO
ω	10.5088	10.9031	10.667	10.667	10.667
Cost	419 000*	420 000	436 200#	432 400#	423 000#
Number of generations	500	500	10 000	10 000	1 000
Run time	10 min**	10 min**	-	-	2.7 min

* *The solution obtained will provide an unfeasible solution if ω = 10.9031, since the pressure at node 3 drops to 29.97 m (Coley, 2003)*
** *According to Savic and Walters (1997) each run took 10 min CPU time on a PC 486/DX 2 50 and one of these solutions were always identified when 10 runs were performed.*
\# *The average cost of a number of different runs. The optimum cost of 419 000 was also obtained, see Table 15.*

TABLE 15 Direct comparison of optimisation results of the Two-loop network					
Initial seed	Keedwell and Khu (2005)		Authors (2008)		
	GA	CANDA-GA			
	Cost	Cost	Cost	Generation nr	Time (s)
1	420 000	441 000	419 000	984	159
12	424 000	424 000	419 000	386	63
123	448 000	439 000	437 000	999	163
1234	442 000	420 000	420 000	662	108
12345	447 000	438 000	420 000	805	131
Average	436 200	432 400	423 000	529	124.8

The optimum solution of 419 000 cost units is obtained if the pipes as listed in Table 16 are used resulting in a minimum pressure in the system at node 6 being 30.44 m. These results were obtained with the following GA parameter: number of generations = 1 000, population = 100, penalty factor = 1.5 and mutation rate was set at 10%.

TABLE 16 Optimisation result of the two-loop network	
Pipe nr	Diameter (mm)
1	457.2
2	254.0
3	406.4
4	101.6
5	406.4
6	254.0
7	254.0
8	25.4

Features of GANEO

The developed GA was implemented in the software package GANEO to test its ability, and some of the features that makes the optimisation process more powerful are:
- Initial seed number – setting of the initial seed number for the pseudo random generator
- Selecting of cross-over method – single point, double point or uniform
- Selecting the mutation procedure and value – random or Min-Max
- Setting the termination criteria – number of generations,

time limit or when no change in fitness occurs for a set number of generations
- Setting of population size – four to one hundred
- Penalty factors – the penalty factor can be set for nodes not meeting the minimum pressure requirement as well as a penalty factor for velocities which are greater than a specified value
- Pipe/link fixing – when an existing network is analysed for rehabilitation purposes GANEO allows the fixing, restricting the changing or adding, of certain pipes. These pipes will thus be kept as is and won't be upgraded or improved although these will be included in the hydraulic analysis of the system. The reason for fixing a pipe is for instance when it is too costly or difficult to change/improve or if it is part of the existing system that must simply be analysed as part of the new extension of the network.

Conclusions

The developed genetic algorithm optimisation model was tested on three benchmark networks and it has been shown to produce good results in a limited number of generations (in relation to other GA-based methods). The weighted penalty function produced satisfactory final results and showed faster initial convergence. The developed GANEO program can be used in the design and analysis of a new network as well as providing suggestions on how to improve an existing network (adding additional pipes or replacing existing pipes). The EPANET software used for the hydraulic analysis of the systems is well-accepted, well-tested analysis software used in various other hydraulic analysis packages. The EPANET hydraulic modelling software as well as the developed GANEO software, which performs the optimisation of the water distribution network, is freeware. The software provides a tool for consulting engineers to optimise the

design or rehabilitation of a water distribution network.

Further research is envisaged to include local search procedures once near optimum solutions are obtained.

Acknowledgements

References

ALPEROVITS E and SHAMIR U (1977) Design of optimal water distribution systems. *Water Resour. Res.* **13** (6) 885-900.

BHAVE PR (1985) Optimal expansion of water distribution systems. *J. Environ. Eng.* ASCE **111** (2) 177-197.

COLEY DA (2003) *An Introduction to Genetic Algorithms for Scientists and Engineers.* World Scientific Publishing Co, Singapore. 227 pp.

DANDY GC, SIMPSON AR and MURPHY LJ (1996) An improved genetic algorithm for pipe network optimization. *Water Resour. Res.* **32** (2) 449-458.

DEB K and AGRAVAL S (1999) A niched-penalty approach for constrain handling in genetic algorithms. *Proc. of Int. Conf. on Artificial Neural Networks and Genetic Algorithms.* Control Theory and Applications Centre, Country Univ., UK.

EIGER G, SHAMIR U and BEN-TAL A (1994) Optimal design of water distribution networks. *Water Resour. Res.* **30** (9) 2637-2646.

FUJIWARA O and KHANG DB (1990) A two-phase decomposition method for optimal design of looped water distribution networks. *Water Resour. Res.* **26** (4) 539-549

GESSLER J (1985) Pipe network optimization by enumeration. *Proc. Spec. Conf. on Comp. Applications/Water Resourc.* ASCE, New York 572-581.

GOULTER IC, LUSSIER BM and MORGAN DR (1986) Implications of head loss path choice in the optimization of water distribution networks. *Water Resour. Res.* **22** (5) 819-822.

HOLLAND JH (1975) *Adaptation in Natural and Artificial Systems* (2nd edn.). University of Michigan Press, Michigan, USA. 1992.

JOINS J and HOUCK C (1994) On the use of non-stationary penalty functions to solve nonlinear constrained optimization problems with GAs. Z Michalewicz, JD Schaffer, HP Schwefel, DB Fogel and H Kitano (eds.) *Proc. 1st IEEE Int. Conf. on Evolutionary Computing.* IEEE Press, Piscataway, NJ. 579-584.

JOWITT P, GARRETT R, COOK S and GERMANOPOULOS G (1988) Real-time forecasting and control for water distribution. In: B Coulbeck and C Orr (eds.) *Computer Applications in Water Supply.* John Wiley & Sons, Letchworth, England. 329-355.

KEEDWELL E and KHU S-T (2005) A hybrid genetic algorithm for the design of water distribution networks. *Eng. Applic. Artificial Intell.* **18** 461-472.

KESSLER A and SHAMIR U (1989) Analysis of the linear programming gradient method for optimal design of water supply networks. *Water Resour. Res.* **25** (7) 1469-1480.

LIN LY and WU WH (2004) Self-organizing adaptive penalty strategy in constrained genetic search. *J. Struct. Multidisciplinary Optim.* **26** 417-428.

LIONG SY and ATIQUZZAMAN M (2004) Optimal design of water distribution network using shuffled complex evolution. *J. Inst. Eng.* **44** (1) 93-107.

MICHALEWICZ Z (1994) *Genetic Algorithms + Data Structures = Evolution Programs* (2nd edn.). Springer-Verslag, Berlin.

MORGAN DR and GOULTER IC (1985) Optimal urban water distribution design. *Water Resour. Res.* **21** (5) 642-652.

MOTESINOS P, GARCIA-GUZMAN A and AYUSO JL (1999) Water distribution network optimization using a modified genetic algorithm. *Water Resour. Res.* **35** (11) 3467-3473.

QUINDRY GE, BRILL ED and LIEBMAN JC (1981) Optimization of looped water distribution systems. *J. Environ. Eng. Div.* ASCE **107** (4) 665-679.

ROSSMAN L (2000) *EPANET Users Manual.* Environmental Protection Agency, Risk Reduction Engineering Laboratory, Cincinnati.

SAVIC D and WALTERS G (1997) Genetic Algorithms for the Least-cost Design of Water Distribution Networks. *J. Water Resour. Plann. Manage.* **123** (2) 67-77.

SCAAKE JC and LAI D (1969) Linear Programming and Dynamic Programming Applications to Water Distribution Network Design. Report 116 Hydrodyn, laboratory, Department of Civil Engineering, Massachusetts Institute of Technology, Cambridge, Massachusetts.

VAIRAVAMOORTHY K and ALI M (2000) Optimal design of water distribution systems using genetic algorithms. *Computer-Aided Civil and Infrastructure Eng.* **15** 374-382.

VAN VUUREN SJ, VAN ROOYEN PG, VAN ZYL JE and VAN DIJK M (2005) Application and Conceptual Development of Genetic Algorithms for Optimisation in the Water Industry. WRC Report No. 1388/1/05. Water Research Commission, Pretoria, South Africa.

WU ZY and SIMPSON AR (2002) A self-adaptive boundary search algorithm and its application to water distribution systems. *J. Hydraul. Res.* **40** 191-203.

WU ZY and WALSKI T (2005) Self-adaptive penalty approach compared with other constrain-handling techniques for pipeline optimization. *J. Water Resour. Plann. Manage.* **131** (3) 181-192.