Journal of Zhejiang University-SCIENCE A (Applied Physics & Engineering)
ISSN 1673-565X (Print); ISSN 1862-1775 (Online)
www.zju.edu.cn/jzus; www.springerlink.com
E-mail: jzus@zju.edu.cn

A comparative study of differential evolution and genetic algorithms for optimizing the design of water distribution systems*

Xiao-lei DONG†, Sui-qing LIU, Tao TAO, Shu-ping LI, Kun-lun XIN
(*College of Environmental Science and Technology, Tongji University, Shanghai 200092, China*)
†E-mail: francisco_dxl@hotmail.com

Received Mar. 18, 2012; Revision accepted July 13, 2012; Crosschecked Aug. 9, 2012

Abstract: The differential evolution (DE) algorithm has been received increasing attention in terms of optimizing the design for the water distribution systems (WDSs). This paper aims to carry out a comprehensive performance comparison between the new emerged DE algorithm and the most popular algorithm—the genetic algorithm (GA). A total of six benchmark WDS case studies were used with the number of decision variables ranging from 8 to 454. A preliminary sensitivity analysis was performed to select the most effective parameter values for both algorithms to enable the fair comparison. It is observed from the results that the DE algorithm consistently outperforms the GA in terms of both efficiency and the solution quality for each case study. Additionally, the DE algorithm was also compared with the previously published optimization algorithms based on the results for those six case studies, indicating that the DE exhibits comparable performance with other algorithms. It can be concluded that the DE is a newly promising optimization algorithm in the design of WDSs.

Key words: Differential evolution (DE), Genetic algorithms (GAs), Optimization, Water distribution systems (WDSs)
doi:10.1631/jzus.A1200072 **Document code:** A **CLC number:** TU991.3

1 Introduction

Providing customers with drinking water at adequate pressure and proper quality at a minimal cost is the goal of all water providers. Considering the high costs associated with the construction of water distribution systems (WDSs), related research in this field has been dedicated to the development of techniques for minimizing the capital costs associated with such infrastructure. This process has been called "optimal design" or "optimization" of WDSs. WDSs optimization problems can be divided in two different types: the design of new WDSs and the expansion of existing WDSs.

Considering the nonlinear relationship between pipe discharge, head loss, and the availability of discrete pipe sizes, optimal WDSs design poses challenges for optimization algorithms. Linear programming (LP) and non-linear programming (NLP) (Schaake and Lai, 1969; Bhave and Sonak, 1992; Varma *et al.*, 1997) techniques were first applied. Since 1990, a number of evolutionary algorithms (EAs) have been applied to the optimization of WDSs. The search strategy of EAs differs from that of the traditional optimization techniques (such as LP or NLP) in that it explores broadly the search space based on stochastic evolution rather than on gradient information. EA techniques include genetic algorithm (GA), ant colony optimization (ACO), shuffled frog leaping algorithm (SFLA), and particle swarm optimization (PSO). GA is an adaptive stochastic algorithm based on natural selection and genetics (Goldberg, 1989), and it has been successfully applied to optimal WDSs design (Simpson *et al.*, 1994; Savic and Walters, 1997; Gupta *et al.*, 1999; Prasad and Park, 2004; Kadu *et al.*, 2008). Besides the GA, other

* Project (No. 2008AA06A413) supported by the National High-Tech R&D (863) Program of China

EAs also have been employed to optimize the design of WDSs. For example, adaptability of ACO for optimal WDSs design was demonstrated by Maier *et al.* (2003). Tabu search was used by da Conceição Cunha and Rebeiro (2004). Geem (2006; 2009) developed harmony search (HS) and particle-swarm harmony search (PSHS) models. Eusuff and Lansey (2003) proposed an SFLA model. Suribabu and Neelakantan (2006) introduced particle swarm optimization (PSO). Tolson *et al.* (2009) developed a hybrid discrete-dynamically dimensioned search (HD-DDS) algorithm for WDSs optimization. Mohan *et al.* (2010) developed honey-bee mating optimization. Zheng *et al.* (2011; 2012) developed a combined NLP-differential evolution algorithm (NLP-DE) and self-adaptive differential evolution algorithm (SADE) for WDSs optimization. Among the EAs, GAs are probably the most popular evolutionary optimization technique. However, they are computationally expensive, especially when dealing with large-scale WDSs.

To overcome the GA limitation, a new stochastic search method, the differential evolution (DE) algorithm, is proposed in this paper. This algorithm, as a stochastic search method, was proposed by Storn and Price (1995). The objective of the present paper is to apply the DE to deal with two different types of WDSs optimization problems. Then, comparisons of performance are made between DE and EAs for the optimal design of WDSs. Six famous benchmark WDS case studies, including the expansion of two existing WDSs and four new designs, are investigated to demonstrate the effectiveness of the proposed optimization method.

2 Methods

2.1 Optimization model for WDSs design

The optimal design of a WDS is often viewed as a least cost optimization problem. The decision variables are the diameters of each pipe in the WDS. The optimal solution is obtained by minimizing the total cost. For a given layout, the source head, elevation and demand values of nodes, pipe lengths and pipe roughness are known in advance. The objective is to find a combination of different sizes of pipe that can satisfy the nodal head constraints at the lowest cost. The objective function used to minimize the cost for a WDS is given by

$$F=\sum_{i=1}^{n}C(D_i)L_i, \tag{1}$$

where D_i is the diameter of pipe i, L_i is the pipe length, $C(D_i)$ is the unit cost of pipe diameter D_i, and n is the total number of pipes in the network.

Typically, the WDSs optimization constraints include flow continuity at each node, energy conservation in each primary loop, and the minimum allowable head requirement at each node. The constraints can be mathematically expressed as

$$q_j^{\text{in}}-q_j^{\text{out}}-q_j=0,\quad j=1,2,\cdots,\text{nd}, \tag{2}$$

$$H_j\ge H_j^{\min},\ j=1,2,\cdots,\text{nd}, \tag{3}$$

$$\left(\sum_{i=1}^{\text{np}_L}\text{HL}_i\right)_L=0,\ L=1,2,\cdots,\text{nL}, \tag{4}$$

$$D_{\min}\le D\le D_{\max}, \tag{5}$$

where q_j^{in} is the flow entering node j, q_j^{out} is the flow leaving node j towards the downstream nodes, q_j is the demand at node j, H_j is the hydraulic head available at node j, $H_j^{\min}$ is the minimum hydraulic head required at node j, nd is the number of demand nodes, HL_i is the head loss in pipe i, np_L is the number of pipes in a loop, nL is the number of loops in the WDS, and in this context, $D_{\min}$ and $D_{\max}$ are the minimum and maximum allowable pipe sizes, respectively. The loop refers to the closed circuit formed by the pipes. Eq. (2) is referred to as the nodal mass balance equation. Eq. (3) is the minimum hydraulic head requirement constraint. Eq. (4) is referred to as the loop energy balance equation, and Eq. (5) is the constraint for the pipe diameters.

2.2 Differential evolution algorithm

The DE algorithm is a population-based stochastic method for global optimization. DE maintains a pair of vector populations where both contain N and D-dimensional vectors of real-valued parameters. N competitions are held in each generation to determine the composition of the next generation. The population is always expressed as

$$\begin{aligned}&\boldsymbol{P}_{x,g}=(\boldsymbol{x}_{i,g}),\ i=0,1,\cdots,N-1,\ \ g=0,1,\cdots,g_{\max},\\&\boldsymbol{x}_{i,g}=(\boldsymbol{x}_{j,i,g}),\ j=0,1,\cdots,D-1,\end{aligned}\tag{6}$$

where N is the number of population vectors, g is the generation counter, and D is the dimensionality, i.e., the number of parameters.

The DE algorithm includes mutation, crossover, and selection operators. The DE process will be described in the next section.

2.2.1 Initialization

Both the upper and lower bounds of each parameter must be specified in advance. These 2D values will be collected by two D-dimensional initialization vectors $b_{\min}$ and $b_{\max}$ in which the subscripts min and max indicate the lower and upper bounds, respectively. Once the initialization bounds have been specified, a random number generator assigns each parameter of every vector a value from within the prescribed range. Normally, the initial value (g=0) of the jth parameter of the ith vector is

$$\boldsymbol{x}_{j,i,0}=\text{rand}_j(0,1)\cdot\left(b_{j,\mathrm{U}}-b_{j,\mathrm{L}}\right)+b_{j,\mathrm{L}},\tag{7}$$

where rand$_j$(0,1) is a uniformly distributed random number from within the range [0,1), b_{L} and b_{U} indicate the lower and upper bounds of the parameter vectors $\boldsymbol{x}_{i,j}$, respectively, and j is a new random value generated for each parameter.

2.2.2 Mutation

The mutation strategy in DE is different from that in GAs. A scaled, randomly sampled vector difference is added to a third vector to produce a population of N mutant vectors. The following formula is used frequently:

$$\boldsymbol{V}_{i,g}=\boldsymbol{x}_{r_0,g}+F\cdot\left(\boldsymbol{x}_{r_1,g}-\boldsymbol{x}_{r_2,g}\right),\tag{8}$$

where $\boldsymbol{V}_{i,g}$ is the mutant vector with respect to the target vector $\boldsymbol{x}_{i,g}$ at generation g. The random indexes r_0, r_1, and r_2 should be mutually exclusive. The mutation weighting factor F is a positive real number that controls the rate at which the population evolves.

2.2.3 Crossover

Uniform crossover is employed after the mutation. The operator builds a trial vector $\boldsymbol{U}_{i,g}$ from the parameter values copied from the two different vectors. In particular, DE crosses each vector with a mutant vector:

$$\boldsymbol{u}_{i,g}=\boldsymbol{u}_{j,i,g}=\begin{cases}\boldsymbol{v}_{j,i,g}, & \text{if } \text{rand}_j[0,1)\le\text{CR},\\ \boldsymbol{x}_{j,i,g}, & \text{otherwise},\end{cases}\tag{9}$$

where $\boldsymbol{u}_{j,i,g}$, $\boldsymbol{v}_{j,i,g}$, and $\boldsymbol{x}_{j,i,g}$ are the jth parameters for the ith trial, mutant, and target vectors, respectively. The crossover probability (CR), CR∈[0, 1] is a user-defined value that controls the fraction of parameter values copied from the mutant. If the random number is less than or equal to CR, the trial parameter is inherited from the mutant $\boldsymbol{v}_{i,g}$; otherwise, the parameter is copied from vector $\boldsymbol{x}_{i,g}$.

2.2.4 Selection

After the crossover, DE uses simple one-to-one survivor selection where trial vector $\boldsymbol{u}_{i,g}$ competes against target vector $\boldsymbol{x}_{i,g}$. The vector with the lowest objective function value survives into the next generation g+1 by

$$\boldsymbol{x}_{i,g+1}=\begin{cases}\boldsymbol{u}_{i,g}, & \text{if } f(\boldsymbol{u}_{i,g})\le f(\boldsymbol{x}_{i,g}),\\ \boldsymbol{x}_{i,g}, & \text{otherwise},\end{cases}\tag{10}$$

where $\boldsymbol{x}_{i,g+1}$ is the ith individual at generation g+1.

Once the new population is installed, the process of mutation, crossover, and selection is repeated until the optimal individual is located or a pre-specified termination criterion is satisfied, e.g., all the individuals are the same. Fig. 1 shows the flowchart of the proposed DE algorithm.

The continuous variables of available discrete pipe sizes are introduced. The continuous pipe sizes are adjusted to the nearest commercially available pipe diameter after application of the mutation operator, obtained from Eq. (8). First, each mutant vector element is checked. If its value is smaller or larger than the minimum or maximum allowable pipe size, then the minimum or maximum allowable pipe size is respectively assigned. If its value is between two sequentially discrete pipe diameters, the closest discrete pipe diameter is assigned. In DE, constraint tournament selection (Deb, 2000) is used to handle head constraints.

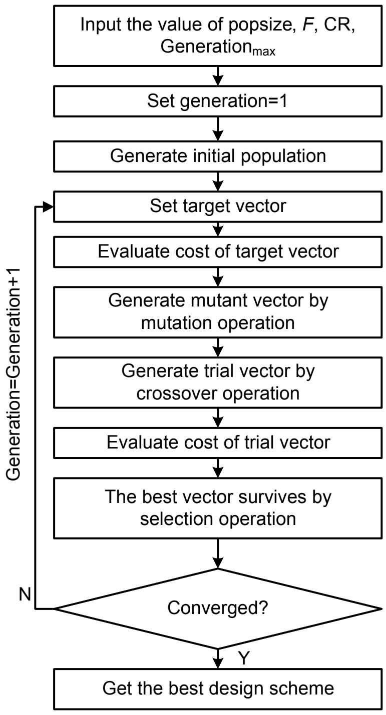

Fig. 1 Flowchart of the proposed DE algorithm

The N, F, and CR parameters are the most important DE parameters. Tuning the three main control variables N, F, and CR and finding the boundaries of their values has been a topic of intensive research. Storn and Price (1995) recommended the DE parameter ranges $D \le N \le 10D$, $0.3 \le F \le 0.9$, and $0.5 \le \mathrm{CR} \le 1.0$ because a DE with these parameter ranges shows generally favorable performance in terms of convergence properties.

3 Case studies

In this section, the DE, GA and other EAs performances in the optimization design of WDSs are compared. Six well-known benchmark WDSs were used to verify the effectiveness of the proposed optimization approach, including the two-loop network, the Goyang water distribution network, the BakRyun water distribution network, the New York tunnels problem (NYTP), the Hanoi problem (HP) and the Balerma network (BN). The Hazen-Williams formula was used to calculate the head loss. The DE and GA used here have been coded in C++. The application combined the EPANET 2.0 solver (Rossman, 2000).

For each case study in the present paper, a preliminary sensitivity analysis was performed to determine the effective N, F, and CR values based on the range given by Storn and Price (1995) for each parameter. The same process was also implemented in the GA application.

3.1 Case study 1: Two-loop network (two-loops, eight pipes)

The two-loop network (Fig. 2) was presented by Alperovits and Shamir (1977). The network consists of seven nodes and eight pipes, fed by a single reservoir with a head of 210 m. The minimum head requirement of the other nodes is 30 m above ground level. The Hazen-Williams coefficient for each new pipe is 130. The set of commercially available diameters is S=[1, 2, 3, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24] in inches (1 inch=2.54 cm). Thus, the total search space is 14^8 (1.48×10^9).

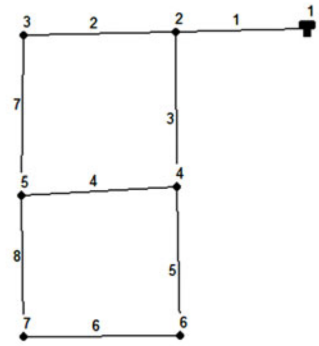

Fig. 2 Layout of the two-loop network

3.2 Case study 2: Goyang water distribution network (Goyang, 30 pipes)

The Goyang water distribution network (Fig. 3) was first presented by Kim *et al.* (1994). The network consists of 22 nodes and 30 pipes, and is fed by a pump (4.52 kW) from a reservoir with a head of 71 m. The minimum head requirement of the other nodes is 15 m above ground level. The Hazen-Williams coefficient for each new pipe is 100. The set of commercially available diameters is S=[80, 100, 125, 150, 200, 250, 300, 350] in mm. Thus, the total search space is 8^{30} (1.24×10^{27}).

3.3 Case study 3: BakRyun water distribution network (BakRyun, nine pipes)

The BakRyun water distribution network (Fig. 4) was presented by Lee and Lee (2001). The network consists of 35 nodes and 58 pipes, and is fed by a single reservoir with a head of 58 m. The minimum head requirement of the other nodes is 15 m above ground level. The Hazen-Williams coefficient for each new pipe is 100. The objective of the problem is to determine the diameters of new pipes (pipes 1–3) and parallel pipes (pipes 4–9) in addition to the existing network. The set of commercially available diameters is *S*=[300, 350, 400, 450, 500, 600, 700, 800, 900, 1000, 1100] in mm. Thus, the total search space is 11^9 (2.36×10^9).

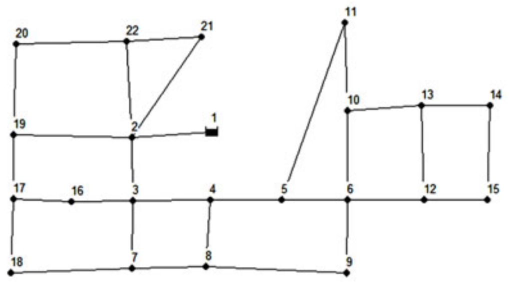

Fig. 3 Layout of the Goyang water distribution network

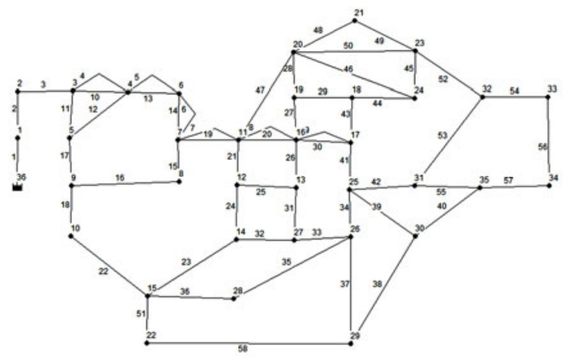

Fig. 4 Layout of the BakRyun water distribution network

3.4 Case study 4: New York tunnels problem (NYTP, 21 pipes)

The New York City water tunnels problem was presented by Schaake and Lai (1969). Fig. 5 shows the layout of the system. The network has 20 nodes and 21 pipes fed by a single reservoir with a head of 300 ft (1 ft=30.48 cm). The objective is to determine if a new pipe is to be laid parallel to an existing pipe and the diameter of a parallel pipe. A selection of 15 pipe diameters is available for the NYTP. A zero pipe size provides an additional option giving a total of 16 for each link. Therefore, the set of commercially available diameters is *S*=[0, 36, 48, 60, 72, 84, 96, 108, 120, 132, 144, 156, 168, 180, 192, 204] in inches. Thus, the total search space is 16^{21} (1.934×10^{25}).

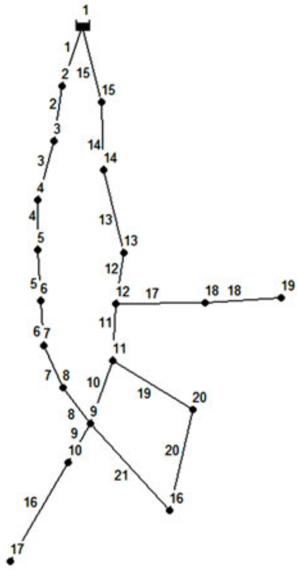

Fig. 5 Layout of the New York tunnels

3.5 Case study 5: Hanoi problem (HP, 34 pipes)

The water distribution network in Hanoi, Vietnam was presented by Fujiwara and Khang (1990). The network consists of 32 nodes and 34 pipes, fed by a single reservoir with a head of 100 m (Fig. 6). The minimum head requirement of the other nodes is 30 m above ground level. The set of commercially available diameters is *S*=[12, 16, 20, 24, 30, 40] in inches. The Hazen-Williams coefficient for each new pipe is 130. If only discrete pipe diameters are considered, the total search space presents 6^{34} (2.87×10^{26}) possible network designs. The Hanoi network is famous for having a largely infeasible search space with a small region of feasible solutions near the maximum pipe sizes, making the finding of an optimal solution to the problem difficult.

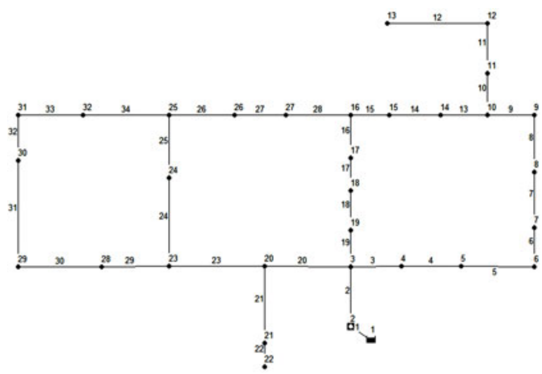

Fig. 6 Layout of the Hanoi network

3.6 Case study 6: Balerma network (BN, 454 pipes)

The Balerma network is an irrigation water distribution network located in the province of Almeria (Spain). It was first investigated by Reca and Martínez (2006). It is a multi-source network containing 443 demand nodes (hydrants), fed by four source nodes. It has 454 pipes and eight loops (Fig. 7). The pipeline database is composed of 10 commercial poly(vinyl chloride) (PVC) pipes with nominal diameters ranging from 125 mm to 630 mm. Thus, the search space is 10^{454}. The absolute roughness coefficient k is equal to 0.0025 mm. The minimum pressure limitation at each node is 20 m above ground level. The pipe costs are provided by Reca and Martínez (2006).

Fig. 7 Layout of the Balerma network

4 Results and discussion

In the present research, DE and GA were applied to the optimal design for all the cases. The DE strategy DE/rand/1/bin (Storn and Price, 1997) was used to generate trial vectors. The identical parameters, which include population size (N), maximum allowable number of evaluations (MAE), F, and CR, were applied in the DE applications. An integer coding method was used in the GA applications. The tournament selection and uniform crossover operators were used in the GA. The identical parameters, which included N, MAE, probability of crossover (P_c), and probability of mutation (P_m), were implemented in all GA applications.

All the algorithms result statistics were recorded. The statistical indicators include the best solution found, percentage of trials in which the current best solution was found, average cost solution, percentage of trials with the best solution found, and the average number of evaluations conducted to obtain the best solution based on different runs.

4.1 Two-loop network problem

For DE, a sensitivity analysis of parameters was executed first. The parameter values of N=20, MAE =10000 were set initially. F was varied from 0.6–0.9 in 0.1 increments. The CR was varied from 0.5–0.6 in increments of 0.1. Eight different combinations of constants were considered from the above range. Then a total of 30 different DE runs using different initial random number seeds were performed for all the combinations.

Table 1 shows the results of different trial runs. It was clear the trial run with the parameter values of F=0.7 and CR=0.5 found the best results. Then a total of 100 different DE runs using different initial random number seeds were performed.

For the GA, the parameters used were N=20, MAE=10000, P_c=0.6, and P_m=0.05. A total of 100 different GA runs using different initial random number seeds were also performed.

The best known solution for the two-loop case study was 419000 USD, first found using the GA technique (Savic and Walters, 1997). This best known solution was also established by the DE and GA optimization techniques proposed in the present study. For this case study, Table 2 shows that the DE obtained the current best solution with a frequency of 40%. It was better than other algorithms. The percentage of best solutions found by DE becomes 100% if the value of N is no less than 100. The DE average cost values were also smaller than those of GA. The proposed DE was able to determine the best known solution after 5987 evaluations, which is fewer than for all the other algorithms' evaluations except for HS and PSHS.

4.2 Goyang water distribution network

For DE, a sensitivity analysis of parameters was carried out. The parameter values of N=50, MAE= 25000 were set. The F was varied from 0.6–0.8 and similarly the CR was varied from 0.5–0.9. Seven different combinations of constants were considered from the above range. Then a total of seven different DE runs using different initial random number seeds were performed for all the combinations.

Table 3 shows the results of different trial runs. It was obvious that the trial run with the parameter values of F=0.6, and CR=0.5 found the best results. Then a total of 27 different DE runs using different initial random number seeds were performed.

For the GA, the parameters used were N=50, MAE=25 000, P_c=0.8, and P_m=0.02. Then 27 different trial runs were also performed with different initial random number seeds.

The current best known solution was found by the proposed DE optimization techniques in the present study. A comparison of the performances of the EAs was applied to the case study. Table 4 shows that the DE obtained the current best solution with a frequency of 52%. This was better than other algorithms.

The values of percentage of best solution found and the average cost values of DE were also better than those of the other optimization techniques. The proposed DE was able to determine the best known solution after 8750 evaluations, which is fewer than those of the other algorithms.

Table 1 Sensitivity analysis for the two-loop network case study

Trial run number	Crossover probability (CR)	Weighting factor (F)	Percentage of trials with best solution found (%)	Average cost solution (USD)
1	0.5	0.6	47	424 833
2	0.5	0.7	53	423 300
3	0.5	0.8	37	423 067
4	0.5	0.9	50	422 133
5	0.6	0.6	40	425 467
6	0.6	0.7	50	422 633
7	0.6	0.8	43	423 967
8	0.6	0.9	40	422 800

Table 2 Solutions for the two-loop network case study

Reference	Algorithm	Best solution found (USD)	Percentage of trials with best solution found (%)	Average cost solution (USD)	Average number of evaluations to find best solutions
Present work	DE	419 000	40	423 860	5987
Present work	GA	419 000	3	471 444	5739
Savic and Walters (1997)	GA	419 000	N/A	N/A	65 000
Geem (2009)	HS	419 000	13	N/A	2891
Geem (2009)	PSHS	419 000	13	N/A	233

Table 3 Sensitivity analysis for the Goyang case study

Trial run number	Crossover probability (CR)	Weighting factor (F)	Percentage of trials with best solution found (%)	Average cost solution (USD)
1	0.5	0.6	60	177 010
2	0.7	0.6	30	177 011
3	0.7	0.8	60	177 016
4	0.8	0.7	40	177 016
5	0.8	0.8	30	177 021
6	0.9	0.6	30	177 034
7	0.9	0.7	30	177 012

Table 4 Solutions for the Goyang case study

Reference	Algorithm	Best solution found (USD)	Percentage of trials with best solution found (%)	Average cost solution (USD)	Average number of evaluations to find best solutions
Present work	DE	177 010	52	177 013	8750
Present work	GA	177 061	4	177 706	12 683
Kim (1994)	NLP	177 143	N/A	N/A	N/A
Geem (2006)	HS	177 136	4	N/A	10 000

4.3 BakRyun water distribution network

For DE, a sensitivity analysis of parameters was accomplished. The parameter values of N=50, MAE =5000 were set initially. The F was varied from 0.5–0.7 in 0.1 increments and the CR was varied from 0.5–0.7 in 0.1 increments. Nine different combinations of constants were considered from the above range. Then a total of 20 different DE runs using different initial random number seeds were performed for all the combinations. Table 5 shows the results of different trial runs. It was clear that the trial run with the parameter values of F=0.7, and CR=0.6 found the best results in fewer evaluations. Then a total of 100 different DE runs using different initial random number seeds were performed.

For the GA, the parameters used were N=100, MAE=50000, P_c=0.8, and P_m=0.01. A total of 100 different GA runs using different initial random number seeds were also performed.

The best known solution for the case study was 903620 USD, first found by the GA technique (Lee and Lee, 2001). This best known solution was also established by the DE and GA optimization techniques proposed in the present study. A comparison of the performances of the EAs was applied to the case studies. For the BakRyun case study, Table 6 shows that all the indicators of DE were better than those of other algorithms.

4.4 New York tunnels problem

For DE, a sensitivity analysis of parameters was carried out. The parameter values of N=50, MAE =50000 were set initially. F was varied from 0.7–0.9 in 0.1 increments and similarly the CR was varied from 0.5–0.7 in increments of 0.1. Nine different combinations of constants were considered from the above range. Then a total of 100 different DE runs using different initial random number seeds were performed for all the combinations. Table 7 shows the results of different trial runs. It was obvious that the trial run with the parameter values of F=0.9, and CR=0.5 found the best results. Then a total of 100 different DE runs using different initial random number seeds were performed.

For the GA, the parameters used were N=100, MAE=100000, P_c=0.8, and P_m=0.03. A total of 100 different GA runs using different initial random number seeds were also performed.

The best known solution for the NYTP case study was 38.64×10^6 USD, first found using the ACO technique (Maier *et al.*, 2003). This best known solution was also established by the DE and GA optimization techniques proposed in the present study. A comparison of the performances of the EAs was applied to the NYTP. Although the result (Savic and Walters, 1997) is lower than 38.64×10^6 USD, the solution was infeasible, as determined by EPANET 2.0 (Maier *et al.*, 2003).

Table 5 Sensitivity analysis for the BakRyun case study

Trial run number	Crossover probability (CR)	Weighting factor (F)	Percentage of trials with best solution found (%)	Average number of evaluations to find best solutions
1	0.5	0.5	100	2735
2	0.5	0.6	100	2870
3	0.5	0.7	100	2888
4	0.6	0.5	95	2815
5	0.6	0.6	100	2733
6	0.6	0.7	100	2610
7	0.7	0.5	100	2705
8	0.7	0.6	100	3000
9	0.7	0.7	95	2745

Table 6 Solutions for the BakRyun case study

Reference	Algorithm	Best solution found (USD)	Percentage of trials with best solution found (%)	Average cost solution (USD)	Average number of evaluations to find best solutions
Present work	DE	903620	100	903620	2555
Present work	GA	903620	7	921733	5796
Lee and Lee (2001)	GA	903620	41	917493	N/A
Geem (2006)	HS	903620	89	904366	5000

For the NYTP case study, Table 8 shows that the DE obtained the current best solution with a frequency of 99%. It was better than the other algorithms except GHEST. The percentage of best solution found by GHEST rises to 100% if limited to the optimal set of parameters. The DE average cost values were lower than those of the other optimization techniques. The proposed DE was able to determine the best known solution after 18271 evaluations, which was fewer only than those of GA and HD-DDS.

4.5 Hanoi network

For DE, a sensitivity analysis of parameters was executed. The parameter values of N=100, MAE=100000 were set initially. The F was varied from 0.6 to 0.7 in 0.1 increments. The CR was varied from 0.5 to 0.8 in increments of 0.1 similarly. Eight different combinations of constants were considered from the above range. Then a total of 100 different DE runs using different initial random number seeds were performed for all the combinations.

Table 9 shows the results of different trial runs. It was obvious that the trial run with the parameter values of F=0.6, and CR=0.7 got the best results. Then a total of 100 different DE runs using different initial random number seeds were performed.

For the GA, N=100, MAE=500000, P_c=0.8, and P_m=0.03 were used. A total of 100 different GA runs using different initial random number seeds were also conducted.

Table 10 shows the results when the same statistical indicators were used to compare the DE, GA and other EAs performances in the Hanoi network case. The current best known solution for the HP case study with a value of 6.081×10^6 USD was found first by Reca and Martínez (2006) using a GA variant (GENOME). This solution was also arrived at by our proposed DE optimization approach (Table 10). In addition, the average cost solutions of DE were better than those of the other algorithms. Moreover, the DE located the best solution with a frequency of 98%. The value of the average number of evaluations to find the

Table 7 Sensitivity analysis for the NYTP case study

Trial run number	Crossover probability (CR)	Weighting factor (F)	Percentage of trials with best solution found (%)	Average cost solution (×10^6 USD)
1	0.5	0.7	94	38.66
2	0.5	0.8	95	38.65
3	0.5	0.9	99	38.65
4	0.6	0.7	93	38.66
5	0.6	0.8	94	38.65
6	0.6	0.9	97	38.65
7	0.7	0.7	87	38.67
8	0.7	0.8	87	38.66
9	0.7	0.9	93	38.66

Table 8 Solutions for the NYTP case study

Reference	Algorithm	Best solution found (×10^6 USD)	Percentage of trials with best solution found (%)	Average cost solution (×10^6 USD)	Average number of evaluations to find best solutions
Present work	DE	38.64	99	38.65	18271
Present work	GA	38.64	64	38.96	26944
Savic and Walters (1997)	GA	37.13	N/A	N/A	10000
Geem (2009)	HS	38.64	N/A	N/A	5991.5
Geem (2009)	PSHS	38.64	N/A	N/A	5923.5
Tolson *et al.* (2009)	HD-DDS	38.64	86	38.65	47000
Dandy *et al.* (2010)	GA	38.64	N/A	38.70	500000
Zheng *et al.* (2010)	GA	38.64	45	39.00	49950
Bolognesi *et al.* (2010)	GHEST	38.64	92	38.64	11464
Zheng *et al.* (2011)	NLP-DE	38.64	99	38.80	8277
Zheng *et al.* (2012)	SADE	38.64	92	38.64	6598

best solution was only 31618 for DE, which was bigger only than those of the HS and PSHS.

Fig. 8 shows the comparison of the DE and GA evolution processes for the HP case for a trial run. Obviously, only the DE is able to obtain the best solution. Moreover, the convergence speed of DE was much faster than that of GA. At the start of the evolution, DE and GA evolved with good convergence properties. Then, the DE identified the best solution quickly. By contrast, the convergence speed of GA declined to a great extent. The GA was not able to locate the best solution despite its more costly evaluations.

Fig. 9 shows the DE and GA evolution processes for a trial run corresponding to the generation number. The DE solutions tended to converge at the same final solution (the best solution), which is DE's most significant advantage over GA. Second, almost all DE solutions converged more quickly than those of the GA. For GA, the evolution process is similar to that of DE during the starting period; however, during the middle and later periods, its convergence speed becomes very slow. In addition, the GA cannot obtain the best solution. Fig. 9 also shows that DE is more robust in that it produced far fewer differences in solutions in the trial run than did GA.

4.6 Balerma network

For DE, a sensitivity analysis of parameters was executed. The parameter values of N=500, MAE=500000 were set initially. The F was set as 0.3 and similarly the CR was varied from 0.5–0.9 in increments of 0.1. Five different combinations of constants were considered from the above range. Then a total of 10 different DE runs using different initial random number seeds were performed for all the combinations.

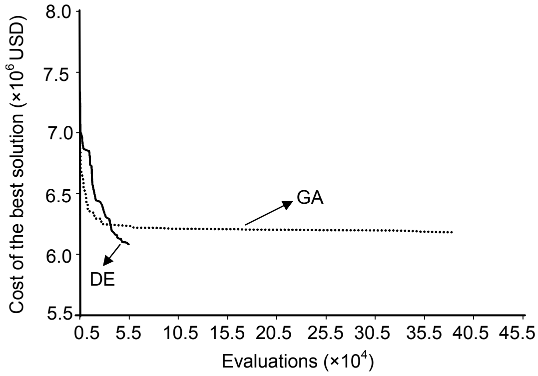

Fig. 8 Evolution process for the HP case study

Table 9 Sensitivity analysis for the HP case study

Trial run number	Crossover probability (CR)	Weighting factor (F)	Percentage of trials with best solution found (%)	Average cost solution (×10^6 USD)
1	0.5	0.6	91	6.089
2	0.5	0.7	81	6.111
3	0.6	0.6	91	6.091
4	0.6	0.7	87	6.097
5	0.7	0.6	98	6.081
6	0.7	0.7	98	6.082
7	0.8	0.6	90	6.084
8	0.8	0.7	90	6.086

Table 10 Solutions for the HP case study

Reference	Algorithm	Best solution found (×10^6 USD)	Percentage of trials with best solution found (%)	Average cost solution (×10^6 USD)	Average number of evaluations to find best solutions
Present work	DE	6.081	98	6.081	31618
Present work	GA	6.135	1	6.284	272830
Reca and Martínez (2006)	GENOME	6.081	10	6.248	N/A
Geem (2009)	HS	6.081	N/A	6.319	27721*
Geem (2009)	PSHS	6.081	N/A	6.340	17980*
Dandy *et al.* (2010)	GA	6.126	N/A	6.214	500000
Bolognesi *et al.* (2010)	GHEST	6.081	38	6.175	50134
Zheng *et al.* (2011)	NLP-DE	6.081	98	6.100	42782
Zheng *et al.* (2012)	SADE	6.081	84	6.090	60532

* The minimum number of evaluations in this study

Table 11 shows the results of different trial runs. It was evident that the trial run with the parameter values of F=0.3 and CR=0.5 found the best results. Then the parameter values of N=500, MAE=1 000 000 were set and a total of 15 different DE runs using different initial random number seeds were performed.

For the GA, N=500, MAE=10 000 000, P_c=0.8, and P_m=0.03 were used. A total of 15 different GA runs using different initial random number seeds were also conducted.

Table 12 shows that with the best solution, the average cost value obtained by DE was better than those of other algorithms except NLP-DE. The convergence speed of DE was less than those of NLP-DE and SADE. Although Zheng *et al.* (2011) found better solutions for the BN case study, the DE algorithm used in their method was seeded by approximate optimal solutions obtained by the NLP, while the DE used in this study is seeded by randomly generated solutions. This may results in that the final optimal solutions found by the NLP-DE method were better than those obtained by the DE used in this study.

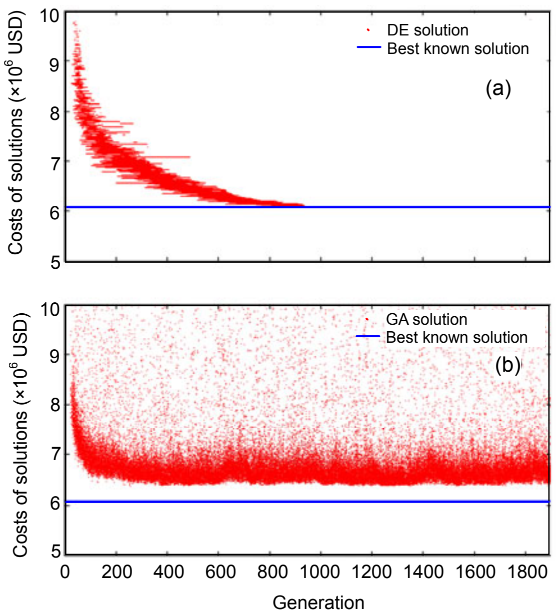

Fig. 9 Results of DE (a) and GA (b) applied to the HP case study

5 Conclusions

DE, a novel optimization technique, was applied in the optimization design of WDSs in this paper. DE was applied to six well-known benchmark WDSs. The performance of the proposed optimization technique was compared with those of GAs and other EAs. The results showed that the DE technique has better convergence properties than all the GAs. The DE technique could locate the current best solution with a higher frequency than the GAs in all cases. DE is also robust; it is able to reproduce the same results over many trials, whereas the GA performance is more dependent on the randomized initialization of the individual parameters.

The efficiency of the algorithm was also proven in larger networks, e.g., the BN case. In the case study, the DE attained the current best solution better than did GAs. Furthermore, the proposed method

Table 11 Sensitivity analysis for the BN case study

Trial run number	Crossover probability (CR)	Average cost solution (×10^6 EUR)
1	0.5	1.958
2	0.6	1.960
3	0.7	1.964
4	0.8	1.969
5	0.9	2.017

Table 12 Solutions for the BN case study

Reference	Algorithm	Best solution found (×10^6 EUR)	Percentage of trials with best solution found (%)	Average cost solution (×10^6 EUR)	Average number of evaluations to find best solutions
Present work	DE	1.955	7	1.958	313 000
Present work	GA	2.104	7	2.144	1133
Reca and Martínez (2006)	GENOME	2.302	10	2.334	N/A
Geem (2009)	HS	2.601	N/A	N/A	45 400
Geem (2009)	PSHS	2.633	N/A	N/A	45 400
Bolognesi *et al.* (2010)	GHEST	2.002	10	2.055	254 400
Zheng *et al.* (2011)	NLP-DE	1.923	10	1.927	1 427 850
Zheng *et al.* (2012)	SADE	1.983	10	1.995	1 200 000

obtained the optimal solutions with a faster convergence speed compared with GAs.

Thus, we conclude that the DE performance in the case studies is outstanding in comparison with that of the GAs. In addition, DE also exhibits comparable performance with other EAs. The DE is simple and robust, converges fast, and finds the optimum solution in most trial runs. Compared with GAs and other EAs, DE can rightfully be regarded as an excellent first choice for the least cost design of WDSs. Developing a multi-objective DE algorithm to optimize WDSs will be our future research goal because the DE has been demonstrated to be effective in finding the least cost solution for WDSs design.

References

Alperovits, E., Shamir, U., 1977. Design of water distribution systems. *Water Resource Research*, **13**(6):885-900. [doi:10.1029/WR013I006P00885]

Bhave, P.R., Sonak, V.V., 1992. A critical study of the linear programming gradient method of optimal design of water supply networks. *Water Resource Research*, **28**(6):1577-1584. [doi:10.1029/92wr00555]

Bolognesi, A., Bragalli, C., Marchi, A., Artina, S., 2010. Genetic heritage evolution by stochastic transmission in the optimal design of water distribution networks. *Advances in Engineering Software*, **41**(5):792-801. [doi:10.1016/j.advengsoft.2009.12.020]

da Conceição Cunha, M., Rebeiro, L., 2004. Tabu search algorithms for water network optimization: simulated annealing approach. *European Journal of Operational Research*, **157**(3):746-758. [doi:10.1016/S0377-2217(03)00242-X]

Dandy, G.C., Wilkins, A., Rohrlach, H., 2010. A methodology for Comparing Evolutionary Algorithms for Optimizing Water Distribution Systems. Proceedings of the 12th Water Distribution System Analysis Symposium, Tucson, USA. American Society of Civil Engineers, Reston, USA, p.786-798. [doi:10.1061/41203(425)73]

Deb, K., 2000. An efficient constraint handling method for genetic algorithms. *Computer Methods in Applied Mechanics and Engineering*, **186**(2-4):311-338. [doi:10.1016/s0045-7825(99)00389-8]

Eusuff, M.M., Lansey, K.E., 2003. Optimisation of water distribution network design using shuffled frog leaping algorithm. *Journal of Water Resource Planning and Management*, **129**(3):210-225. [doi:10.1061/(ASCE)0733-9496(2003)129:3(210)]

Fujiwara, O., Khang, D.B., 1990. A two-phase decomposition method for optimal design of looped water distribution networks. *Water Resource Research*, **26**(4):539-549. [doi:10.1029/WR026i004p00539]

Geem, Z.W., 2006. Optimal cost design of water distribution networks using harmony search. *Engineering Optimization*, **38**(3):259-280. [doi:10.1080/03052150500467430]

Geem, Z.W., 2009. Particle-swarm harmony search for water network design. *Engineering Optimization*, **41**(5):297-311. [doi:10.1080/03052150802449227]

Goldberg, D.E., 1989. Genetic Algorithms in Search, Optimization and Machine Learning. Addison-Wesley Longman Publishing Company, Boston, MA, USA.

Gupta, I., Gupta, A., Khanna, P., 1999. Genetic algorithm for optimization of water distribution systems. *Environmental Modeling & Software*, **14**(5):437-446. [doi:10.1016/S1364-8152(98)00089-9]

Kadu, M.S., Gupta, R., Bhave, P.R., 2008. Optimal design of water networks using a modified genetic algorithm. *Journal of Water Resource Planning and Management*, **134**(2):147-160. [doi:10.1061/1031(ASCE)0733-9496(2008)134:2(147)]

Kim, J.H., Kim, T.G., Kim, J.H., Yoon, Y.N., 1994. A study on the pipe network system design using non-linear programming. *Journal of Korean Water Resource Association*, **27**(4):59-67.

Lee, S.C., Lee, S.I., 2001. Genetic algorithms for optimal augmentation of water distribution networks. *Journal of Korean Water Resource Association*, **34**(5):567-575.

Maier, H.R., Simpson, A.R., Zecchin, A.C., Foong, W.F., Phang, K.Y., Seah, H.Y., Tan, C.L., 2003. Ant colony optimization for the design of water distribution systems. *Journal of Water Resource Planning and Management*, **129**(3):200-209. [doi:10.1061/(ASCE)0733-9496(2003)129:3(200)]

Mohan, S., Jinesh Babu, K.S., 2010. Optimal water distribution network design with Honey-Bee mating optimization. *Journal of Computing in Civil Engineering*, **24**(1):117-126. [doi:10.1061/(ASCE)CP.1943-5487.0000018]

Prasad, D.T., Park, N.S., 2004. Multiobjective genetic algorithms for design of water distribution networks. *Journal of Water Resource Planning and Management*, **130**(1):73-82. [doi:10.1061/(ASCE)0733-9496(2004)130:1(73)]

Reca, J., Martínez, J., 2006. Genetic algorithms for the design of looped irrigation water distribution networks. *Water Resource Research*, **42**:W05416. [doi:10.1029/2005WR004383]

Rossman, L.A., 2000. EPANET2-User Manual. National Risk Management Research Laboratory. Office of Research and Development, US Environmental Protection Agency, Cincinnati.

Savic, D.A., Walters, G.A., 1997. Genetic algorithms for least-cost design of water distribution networks. *Journal of Water Resource Planning and Management*, **123**(2):67-77. [doi:10.1061/(ASCE)0733-9496(1997)123:2(67)]

Schaake, J., Lai, D., 1969. Linear Programming and Dynamic Programming Applications to Water Distribution Network Design. Research Report No. 116, Hydrodynamics Laboratory, Department of Civil Engineering, Massachusetts Institute of Technology, Cambridge, Massachusetts.

Simpson, A.R., Dandy, G.C., Murphy, L.J., 1994. Genetic algorithms compared to other techniques for pipe optimization. *Journal of Water Resource Planning and Management*, **120**(4):423-443. [doi:10.1061/(ASCE)0733-9496(1994)120:4(423)]

Storn, R., Price, K., 1995. Differential Evolution—A Simple and Efficient Adaptive Scheme for Global Optimization over Continuous Space. Technical Report TR-95-012, International Computer Science Institute, Berkeley, CA.

Storn, R., Price, K.V., 1997. Differential evolution—a simple and efficient heuristic for global optimization over continuous spaces. *Journal of Global Optimization*, **11**(4): 341-359. [doi:10.1023/A:1008202821328]

Suribabu, C.R., Neelakantan, T.R., 2006. Design of water distribution networks using particle swarm optimization. *Journal of Urban Water*, **3**(2):111-120. [doi:10.1080/15730620600855928]

Tolson, B.A., Asadzadeh, M., Maier, H.R., Zecchin, A.C., 2009. Hybrid discrete dynamically dimensioned search (HD-DDS) algorithm for water distribution system design optimization. *Water Resources Research*, **45**:W12416. [doi:10.1029/2008WR007673]

Varma, K.V.K., Narasimhan, S., Bhallamudi, S.M., 1997. Optimal design of water distribution systems using an NLP method. *Journal of Environmental Engineering*, **123**(4):381-388. [doi:10.1061/(ASCE)0733-9372(1997)123:4(381)]

Zheng, F., Simpson, A.R., Zecchin, A.C., 2010. A Method of Assessing the Performance of Genetic Algorithm Optimization Water Distribution Design. Proceedings of the 12th Water Distribution System Analysis Symposium, Tucson, USA. American Society of Civil Engineers, Reston, USA, p.771-785. [doi:10.1061/41203(425)72]

Zheng, F., Simpson, A.R., Zecchin, A.C., 2011. A combined NLP-differential evolution algorithm approach for the optimization of looped water distribution systems. *Water Resources Research*, **47**:W08531. [doi:10.1029/2011WR010394]

Zheng, F., Zecchin, A.C., Simpson, A.R., 2012. A self-adaptive differential evolution algorithm applied to water distribution system optimization. *Journal of Computing in Civil Engineering ASCE*, in press. [doi:10.1061/(ASCE)CP.1943-5487.0000208)]