Tanyimboh, Tiku T. and Czajkowska, Anna M. (2018) Joint entropy and multi-objective evolutionary optimization of water distribution networks. Water Resources Management. ISSN 0920-4741 , http://dx.doi.org/10.1007/s11269-017-1888-y

This version is available at https://strathprints.strath.ac.uk/63023/

CrossMark

Joint Entropy Based Multi-Objective Evolutionary Optimization of Water Distribution Networks

Tiku T. Tanyimboh [1,2] • **Anna M. Czajkowska** [2,3]

Abstract It is essential to consider resilience when designing any water distribution network and surrogate measures of resilience are used frequently as accurate measures often impose prohibitive computational demands in optimization algorithms. Previous design optimization algorithms based on flow entropy have essentially employed a single loading condition because the flow entropy concept formally has not been extended to multiple loading conditions in water distribution networks. However, in practice, water distribution networks must satisfy multiple loading conditions. The aim of the research was to close the gap between the prevailing entropy-based design optimization approaches based on one loading condition essentially and water distribution practice that must address multiple loading conditions. A methodology was developed and applied to a real-world water distribution network in the literature, based on the concept of the joint entropy of independent probability schemes. The results demonstrated that the critical loading conditions were design specific. In other words, the critical loading and operating conditions cannot readily be determined beforehand. Consequently, maximizing the joint entropy provided the most consistently competitive solutions in terms of the balance between cost and resilience. The results were derived using a penalty-free genetic algorithm with three objectives. Compared to previous research using flow entropy based on a single loading condition and two objectives, there was a substantial increase of 274% in the number of nondominated solutions achieved.

✉ Tiku T. Tanyimboh
tiku.tanyimboh@wits.ac.za

1   School of Civil and Environmental Engineering, University of the Witwatersrand, Johannesburg WITS 2050,, South Africa

2   Department of Civil and Environmental Engineering, University of Strathclyde Glasgow, James Weir Building, 75 Montrose Street, Glasgow G1 1XJ, UK

3   RPS Group, Merchantile Building, 53 Bothwell Street, Glasgow G2 6TS, UK

🍀 Springer

1 Introduction

Water distribution networks consist of various components such as pipes, pumps, valves and tanks. The networks need to satisfy current and future water demands (Spiliotis and Tsakiris 2012; Korteling et al. 2013; Tsakiris and Spiliotis 2017). The main objective considered when designing water distribution networks is cost minimization subject to adequate water supply and pressure at the demand nodes. However, inevitably some network elements will be unavailable occasionally, for example, due to pipe breakage and pump failure. Therefore, some spare capacity needs to be included to enable the network to perform reasonably well under both normal and abnormal operating conditions. Thus, network resilience is important also (Harrison and Williams 2016; Herrera et al. 2016; Agathokleous et al. 2017). Resilience is characterised by redundancy, failure tolerance and reliability.

Optimizing the design of water distribution networks is an extremely complex problem that is classed as NP-hard (Yates et al. 1984). de Neufville et al. (1971) were among the first researchers who recognised that water system design involved at least two conflicting objectives related to cost minimization and network performance. For many years, conflicting objectives were often aggregated into a scalar function and solved as a single-objective optimization problem.

More recently, evolutionary algorithms have been preferred. Evolutionary algorithms are stochastic search techniques well suited to identifying the Pareto optimal sets in complex multi-objective optimization problems. Many population-based optimization techniques have been employed in the design of water distribution networks. Some examples are genetic algorithms (Holland 1975), differential evolution (Storn and Price 1997), particle swarm (Kennedy and Eberhart 1995), ant colony (Dorigo et al. 1999) and harmony search (Geem et al. 2001).

Elitism is an extremely important feature in evolutionary algorithms that ensures the best solutions in a population are retained for inclusion in the next generation. This ensures that the fittest candidates will be preserved to improve the convergence characteristics (Zitzler et al. 2000). Vector Evaluated Genetic Algorithm (Schaffer 1985), Vector-Optimized Evolution Strategy (Kursawe 1990), Multi-Objective Genetic Algorithm (Fonseca and Fleming 1993), Weight-Based Genetic Algorithm (Hajela and Lin 1992) and Non-Dominated Sorting Genetic Algorithm (Srinivas and Deb 1994) are some examples of non-elitist evolutionary algorithms. Examples of elitist evolutionary algorithms include the Non-Dominated Sorting Genetic Algorithm (NSGA) II (Deb et al. 2002), Distance-Based Pareto Genetic Algorithm (Osyczka and Kundu 1995), Strength Pareto Evolutionary Algorithm (Zitzler and Thiele 1998) and Pareto-Archive Evolution Strategy (Knowles and Corne 2000).

Genetic algorithms are widely employed in the design of water distribution networks (Wu and Simpson 2001; Vairavamoorthy and Ali 2005; Rao and Salomons 2007; Haghighi et al. 2011; Creaco and Pezzinga 2015). Examples include design and rehabilitation (Jayaram and Srinivasan 2008), pump operation scheduling (Goldberg and Kuo 1987), tank siting and sizing (Prasad 2010) and water quality optimization (Munavalli and Kumar 2003).

Many researchers in various disciplines including water distribution employed the nondominated sorting genetic algorithm NSGA II (Yan et al. 2016). The majority of the applications of NSGA II in

the area of water distribution considered design optimization through pipe sizing. Different measures of reliability were included in some cases, for example, Prasad and Park (2004), Farmani et al. (2006) and Saleh and Tanyimboh (2016). Some other applications of NSGA II included Farmani et al. (2006) and Prasad (2010) who designed the benchmark network called "Anytown". Jayaram and Srinivasan (2008) and Siew et al. (2014) investigated design and rehabilitation based on the life cycle cost. In the area of drinking water safety, Preis and Ostfeld (2008) and Weickgenannt et al. (2010) considered contamination detection whilst Jeong and Abraham (2006) focused on the consequences of external attacks.

Due to the high computational burden associated with the reliability calculations, surrogate measures (Forrester et al. 2008; Díaz et al. 2016) have been proposed, for example, flow entropy and resilience index (Awumah et al. 1990, 1991; Tanyimboh 1993; Tanyimboh and Templeman 1993a, b; Todini 2000). The advantage of using such measures is that they are relatively simple functions that do not require repetitive time-consuming hydraulic simulations and can be incorporated in optimization algorithms relatively easily.

The idea of incorporating Shannon's informational entropy measure of uncertainty in the design of water distribution networks was introduced by Awumah et al. (1990). Tanyimboh and Templeman (1993a, b) formalised the definition of the entropy function for flow networks using the conditional entropy concept (Khinchin 1953, 1957) and multiple probability spaces. Czajkowska and Tanyimboh (2013) demonstrated that the entropy-based solutions derived from multiple operating conditions outperformed those from a single operating condition.

It has been shown that networks designed to carry the maximum entropy flows are more reliable than the traditional minimum cost solution and the relationship between reliability and entropy is strong (Tanyimboh and Templeman 1993b, 2000; Tanyimboh and Kalungi 2008; Tanyimboh and Setiadi 2008). The evidence in the literature shows that higher entropy values increase the uniformity of the pipe diameters (Awumah et al. 1991; Tanyimboh and Templeman 1993b) which, therefore, increases the reliability by increasing the opportunities for flow re-routing (Prasad and Park 2004). Investigations by Gheisi and Naser (2015) and Tanyimboh et al. (2016) demonstrated that flow entropy correlated well with hydraulic reliability and failure tolerance while other surrogate measures were mostly inconsistent.

This article is concerned with the formulation of the maximum entropy model for water distribution networks with multiple operating conditions. The primary aim is to develop and demonstrate a rigorous approach using empirical results. Research on flow entropy hitherto has focused on the peak demands. However, different loading conditions may be critical at various points during the typical 24-h operating cycle and there is a clear need to develop more realistic maximum entropy models applicable to real rather than small hypothetical networks.

2 Entropy Maximizing Design Optimization Approach

Steady state simulation is widely used in the design of water distribution networks. However, there are some important situations in which steady state modelling has weaknesses, e.g. valve operation and analysing energy consumption. Extended period simulation is more realistic. It leads to a better understanding of the characteristics e.g. due to time-varying water demands, and tank water level variations. Moreover, the optimization of pump scheduling or tank sizing and siting require extended period simulations.

Previously it was thought that if the peak daily demands were satisfied then the other operating conditions would be satisfied as well (Vamvakeridou-Lyroudia et al. 2005).

However, Prasad (2010) demonstrated that the node pressure constraints have to be considered for all loading conditions. Alperovits and Shamir (1977) observed that the minimum demand periods have to be considered in addition to the maximum daily demand and fire flows. In addition to the fire-fighting and other emergency flows, the loading patterns that are frequently considered include:

(a). The maximum daily demand, i.e. the peak demand over a 24-h period.
(b). The peak hour demand that usually occurs in the evening of the maximum day.
(c). The average daily demand, i.e. the average demand for the average day's consumption.
(d). The minimum daily demand, i.e. the loading pattern when water consumption is at its lowest level and the tanks refill.

2.1 Informational Entropy under Multiple Operating Conditions

Alternative hypotheses for maximizing the entropy when there are multiple operating conditions include: (a) maximizing the largest entropy value across all the operating conditions; (b) maximizing the smallest entropy value across all the operating conditions, to alleviate the worst-case performance; and (c) maximizing the sum of the separate entropies from all the operating conditions.

Collectively, the entropy values for the various operating conditions constitute an entropy vector. Maximizing the largest entropy value implies maximizing the largest element in the entropy vector. On its own, the largest entropy element provides only a partial characterization and thus may not represent the overall performance adequately. It has the potential to overestimate the resilience and, if used in design optimization, could lead to suboptimal results (Czajkowska 2016). This option yields the following objective function.

$$\text{Maximize} f = Max(S_k, \quad \forall k) \tag{1}$$

where S_k is the entropy of the kth operating condition. Additional details on S_k are in the supplementary data.

On the other hand, maximizing the minimum element of the entropy vector would undoubtedly alleviate the worst-case scenario. However, it could underestimate the overall performance and, consequently, yield suboptimal or inconsistent results when used in optimization algorithms (Czajkowska 2016). The objective function based on the minimum entropy element is as follows.

$$\text{Maximize} f = Min(S_k, \quad \forall k) \tag{2}$$

where S_k is the entropy of the kth operating condition.

Maximizing the sum of the separate entropies aims to achieve a good performance for all the operating conditions collectively.

$$\text{Maximize} f = \sum_k S_k \tag{3}$$

where S_k is the entropy of the kth operating condition.

A basic property of entropy as a measure of uncertainty is that the joint entropy of two or more independent probabilistic schemes is the sum of their separate entropies (Shannon 1948, Tanyimboh 1993: 73-77). Accordingly, for independent operating conditions, the joint entropy

is the sum of the separate entropies. The appeal of this interpretation lies in the fact that it accounts for all the operating conditions. Furthermore, it does not require any additional assumptions or criteria. Therefore, viewed as a basic property of the network, the logical conclusion is that the sum of the entropies should be maximized. *This was the fundamental hypothesis of the research.*

2.2 Network Design Optimization Model

Flow entropy may be included in the design optimization of a water distribution network to minimize the cost without sacrificing resilience completely. In this way, redundancy is safeguarded and deployed to the best possible advantage, to help address any unanticipated flow re-routing and short-term increases in demand. The objectives were thus cost minimization and entropy maximization. Only the initial construction cost was considered in this research. Other costs and additional factors may be incorporated relatively easily (Siew et al. 2014). Though relevant, these wider considerations were not the main focus of the investigation.

The constraints were the constitutive equations (i.e. conservation of mass and energy) and the minimum node pressure constraints. The equations for the conservation of mass and energy were satisfied by embedding the EPANET 2 hydraulic simulation model (Rossman 2000) in the evolutionary algorithm. The minimum node pressure constraints were addressed by introducing an additional objective that considers the feasibility of the solutions.

Michalewicz (1995) classified the approaches for addressing constraints in evolutionary algorithms as: (a) repairing, (b) modifying, (c) rejecting and (d) penalizing strategies. The most common practice is to degrade infeasible solutions by applying penalties, with greater constraint violations incurring higher penalties. Excessively high penalties may confine the search to the feasible region of the solution space. However, searching through the feasible and infeasible regions improves the efficiency and yields better solutions than searching in the feasible regions only (Glover and Greenberg 1989). Designing penalty functions and calibrating the associated parameters is a complex task and requires extensive fine-tuning (Dridi et al. 2008). A penalty-free formulation was developed to obviate the difficulties.

According to the maximum entropy formalism, the entropy should be maximized subject to the relevant constraints without introducing any arbitrary assumptions (Jaynes 1957). Hence, the optimization problem may be summarized briefly as follows.

$$\text{Minimize the cost} \, f_1 = \sum_i^{np} C_i(d_i, L_i) \tag{4}$$

$$\text{Minimize the node pressure deficits} \, f_2 = Max\left\langle \max\left[0, \left(H_n^{req} - H_n\right)\right]; \quad \forall n \right\rangle \tag{5}$$

$$\text{Maximize the flow entropy} \, f_3 = S \tag{6}$$

$$\text{Subject to}: d_i \in D; \quad \forall i \tag{7}$$

$C_i(d_i, L_i)$ is the cost of pipe i with diameter d_i and length L_i while np is the number of pipes. The set D comprises the available discrete pipe diameter options. S is the flow entropy. H_n and H_n^{req}

are, respectively, the available and required residual heads at node n. The required head corresponds to the pressure above which the demand is satisfied in full. The decision variables are the pipe diameters. The hydraulic simulation model EAPANET 2, that ensures energy and flow conservation, also provides the nodal heads.

3 Solution Methodology

3.1 Solution of the Optimization Problem

NSGA II was employed as it is efficient and used widely by many researchers in various disciplines. The evolutionary optimization procedure in NSGA II (Deb et al. 2002) is based on Pareto-dominance and global elitism. It maintains diversity by seeking an even distribution of the nondominated solutions in the objective space using the crowding distance. The crowding distance is a measure of the spatial distribution of the solutions in the objective space and is based on the average distance between a solution and its neighbours.

The source code of NSGA II (in C++) was modified and linked to the hydraulic simulation model EPANET 2. EPANET 2 (Rossman 2000) is public domain software for modelling water distribution networks that performs both steady state and dynamic simulations, and water quality modelling. A procedure that calculates the flow entropy for any given network configuration was developed, tested and incorporated in the optimization algorithm. The multiobjective optimization algorithm thus produced can handle both single and multiple operating conditions.

While investigating the methodology it was observed that the number of infeasible solutions in the Pareto-optimal sets exceeded the number of feasible solutions. This may be because the nodal head deficit function f_2 in Eq. 5 does not distinguish between solutions with different levels of surplus head, as all feasible solutions have a deficit of zero. Hence, for *feasible solutions*, the Pareto-dominance is based on the cost and entropy only. By contrast, for *infeasible solutions*, the Pareto-dominance is based on three objectives. It is known that more solutions become nondominated as the number of objectives increases (Ishibuchi et al. 2015), and this may have contributed to the imbalance between the number of feasible and infeasible solutions in the Pareto-optimal sets achieved.

Therefore, on discarding the infeasible solutions at the end of each optimization run, all the *nondominated feasible* solutions achieved from the start to the end of the optimization were incorporated in the Pareto-optimal set. Membership of the final Pareto set of feasible solutions was based on Pareto-dominance and the selection was done using additional software developed (in Perl) in the research. Furthermore, the nondominated solutions from all the optimization runs were combined and sorted based on Pareto dominance to obtain a single unified nondominated set.

3.2 Resilience Evaluation Procedure

Two of the criteria used to assess the proposed maximum entropy design approach were the hydraulic capacity reliability and failure tolerance that emphasize different aspects of resilience (Tanyimboh et al. 2001). Pressure-driven analysis (Tsakiris and Spiliotis 2014) based on the logistic pressure-discharge relationship (Tanyimboh and Templeman 2010) was used to simulate the effects of pipe failures.

The definition used for the hydraulic capacity reliability is the network's ability to fulfil on average the required nodal demands at adequate pressure, under normal and abnormal operating conditions. The pipe failure model used was taken from Cullinane et al. (1992). Also, the pipe failure tolerance provides an estimate of the fraction of the total demand that the network can satisfy on average when one or more components are out of service and its importance was emphasized previously (Gheisi and Naser 2013, 2015).

4 Results and Discussion

4.1 Problem Specifications

This example is based on the network shown in Fig. 1 that serves part of the city of Ferrara (Creaco et al. 2010, 2012). It consists of 49 nodes, 76 pipes and 29 loops. The total demand is 367 l per second. Each reservoir has a head of 30 m. Manning's roughness coefficient is 0.015 for all the pipes the total length of which is about 25.2 km. The elevation and minimum required head at the demand nodes are 0 m and 28 m respectively.
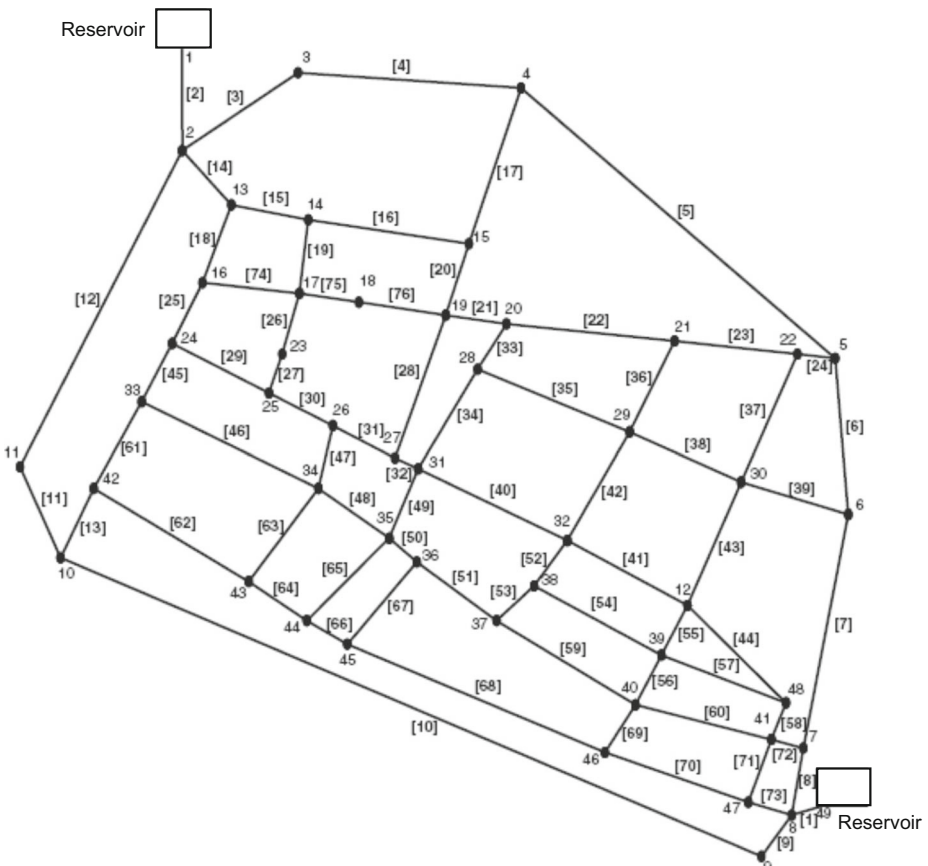

Fig. 1 Topology of the network investigated. The pipe identifiers are shown in square brackets

As the original data for the network had only one loading condition, two fire-flow scenarios were created with reference to the fire flows in Simpson et al. (1994). Under the fire flow conditions, the required residual head at the demand nodes was taken as 14 m, except for the node with a fire-fighting flow for which the required residual head was 8.4 m. The two fire flow conditions correspond to a fire at node 31 for Fire Flow 1 and node 11 for Fire Flow 2. Details of the three loading conditions and the pipe lengths are in Tables 1 and 2, respectively.

The available pipe diameter options and their costs per metre (in mm and €/m) are: {(150, 271.94), (200, 299.43), (250, 328.01), (300, 359.54), (350, 399.03), (400, 438.63), (450, 461.34), (500, 502.78)}. There are 8 pipe diameter options, so the solution space comprises 8^{76} or 4.313×10^{68} infeasible and feasible solutions. A 3-bit binary substring with 8 (i.e. 2^3) substrings was used. The GA parameters were: $N_E = 500,000$, $N_G = 1000$, $N_S = 500$, $N_R = 30$, $p_c = 1.0$ and $p_m = 1/228 = 0.004$ (this is the probability that any single bit would mutate by being reversed from 0 to 1 or vice versa) where 228 is the chromosome length. N_E is the maximum number of function evaluations or hydraulic simulations allowed; N_G is the number

Table 1 Nodal demands and required residual heads

Demand Nodes	Peak Demand		Fire Flow 1		Fire Flow 2	
	Demands (l/s)	Required heads (m)	Demands (l/s)	Required heads (m)	Demands (l/s)	Required heads (m)
2	5.000	28.0	5.000	14.0	5.000	14.0
3	3.000	28.0	3.000	14.0	3.000	14.0
4	6.500	28.0	6.500	14.0	6.500	14.0
5	10.50	28.0	10.50	14.0	10.50	14.0
6	9.500	28.0	9.500	14.0	9.500	14.0
7	6.500	28.0	6.500	14.0	6.500	14.0
8	8.500	28.0	8.500	14.0	8.500	14.0
9	9.500	28.0	9.500	14.0	9.500	14.0
10	10.00	28.0	10.00	14.0	10.00	14.0
11	4.500	28.0	4.500	14.0	18.00	8.40
12	5.500	28.0	5.500	14.0	5.500	14.0
13	3.500	28.0	3.500	14.0	3.500	14.0
14	4.000	28.0	4.000	14.0	4.000	14.0
15	5.000	28.0	5.000	14.0	5.000	14.0
16	4.000	28.0	4.000	14.0	4.000	14.0
17	8.000	28.0	8.000	14.0	8.000	14.0
18	6.500	28.0	6.500	14.0	6.500	14.0
19	7.500	28.0	7.500	14.0	7.500	14.0
20	6.000	28.0	6.000	14.0	6.000	14.0
21	8.000	28.0	8.000	14.0	8.000	14.0
22	8.500	28.0	8.500	14.0	8.500	14.0
23	2.500	28.0	2.500	14.0	2.500	14.0
24	4.500	28.0	4.500	14.0	4.500	14.0
25	4.500	28.0	4.500	14.0	4.500	14.0
26	7.500	28.0	7.500	14.0	7.500	14.0
27	6.000	28.0	6.000	14.0	6.000	14.0
28	6.000	28.0	6.000	14.0	6.000	14.0
29	8.000	28.0	8.000	14.0	8.000	14.0
30	7.125	28.0	7.125	14.0	7.125	14.0
31	9.000	28.0	39.00	8.40	9.000	14.0
32	8.625	28.0	8.625	14.0	8.625	14.0
33	5.000	28.0	5.000	14.0	5.000	14.0
34	11.50	28.0	11.50	14.0	11.50	14.0

Table 2 Pipe lengths

Pipes	Lengths (m)	Pipes	Lengths (m)	Pipes	Lengths (m)	Pipes	Lengths (m)
1	1.0000	20	230.65	39	239.08	58	98.510
2	1.0000	21	159.46	40	404.28	59	458.41
3	260.00	22	456.91	41	367.51	60	386.40
4	600.00	23	313.99	42	391.18	61	135.01
5	1200.0	24	73.800	43	229.78	62	559.91
6	600.00	25	192.73	44	404.76	63	319.10
7	800.00	26	176.53	45	215.20	64	193.48
8	200.00	27	124.94	46	575.37	65	393.04
9	500.00	28	478.71	47	210.89	66	60.000
10	2100.0	29	282.03	48	257.55	67	243.45
11	500.00	30	262.90	49	277.95	68	646.80
12	1000.0	31	202.43	50	48.100	69	236.22
13	304.16	32	52.910	51	309.11	70	422.08
14	341.16	33	154.54	52	150.25	71	244.93
15	210.50	34	352.16	53	174.42	72	5.0000
16	422.28	35	165.08	54	396.55	73	31.840
17	529.66	36	325.05	55	169.47	74	258.71
18	254.00	37	377.94	56	178.62	75	154.70
19	252.90	38	275.39	57	339.77	76	229.87

of generations; N_S is the population size; N_R is the number of independent runs with random initial populations; and p_c and p_m are the crossover and mutation probabilities. A single-point crossover operator was used to produce two offspring from two parents. The average CPU time for a single optimization run was about 90 min on a PC (Intel Core 2 Duo, 3.5 GHz, 3GB RAM); in other words, 45 h or 1.9 days in total for the 30 optimization runs. The results are summarised in the supplementary data that is available online, in Table C1 in Appendix C.

4.2 Effectiveness of the Optimization Approach

It was thought that the surplus heads at the critical nodes would increase as the entropy increased. However, there was no correlation between the surplus heads and entropy or cost as shown in Fig. 2. The surplus heads were relatively small with an average of 0.027 m. As the pipe diameters were discrete and thus small surpluses were virtually unavoidable given that only eight pipe diameter options were available, the methodology developed can be considered satisfactory with respect to the optimality of the solutions.

The results of individual optimization runs revealed apparent gaps in some of the Pareto fronts and an imbalance between the feasible and infeasible solutions. It is possible that the number of feasible solutions for this network is relatively small in comparison to the size of the solution space. It may be noted that the maximum headloss allowed between the supply and demand nodes was only 2 m based on the available head of 30 m at the supply nodes and the stipulated minimum residual head of 28 m at the demand nodes. In fact the present results are consistent with Saleh and Tanyimboh (2016). While the cost and entropy values (for the peak demands) were similar, the present formulation achieved 74 nondominated feasible solutions with 15 million function evaluations in total compared to only 27 nondominated feasible solutions in Saleh and Tanyimboh (2016) with 20 million evaluations in total. Also, the average surplus head was 2.7 cm compared to 3.5 cm in Saleh and Tanyimboh (2016).
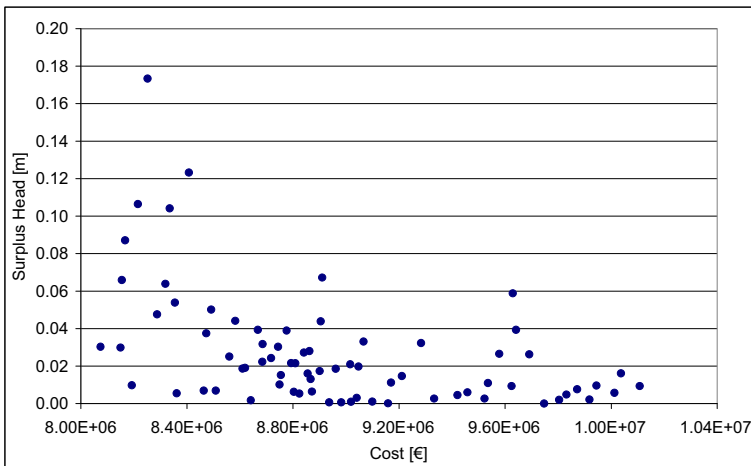
(a) Surplus head versus cost
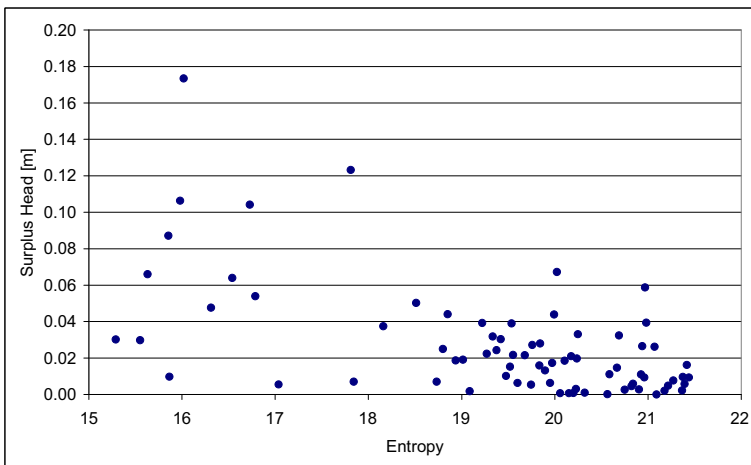


(b) Surplus head versus entropy

Fig. 2 Quality of the solutions achieved. The plots are based on 30 optimization runs

In Saleh and Tanyimboh (2016) the least expensive feasible solution cost was €8.011 million and the entropy was 4.7968. The cost of the most expensive solution was €10.285 million and the entropy was 7.2989. Herein, the corresponding values were: €8.074 million and 5.125, for the peak demands, for the least expensive solution; and €10.108 million and 7.127, for the peak demands, for the most expensive solution. Thus the solutions from the proposed methodology are very competitive. In any case, Saleh and Tanyimboh (2016) did not consider multiple operation conditions; so the present solutions have more advantages.

The value of the entropy increased rapidly from the start of the optimization, followed by a second phase with relatively steady progress as shown in Fig. 3a. Only the feasible solutions were included in the results shown in Fig. 3a. The reason is that the entropy values for infeasible solutions are unrealistic, as the flow rates derived from demand-
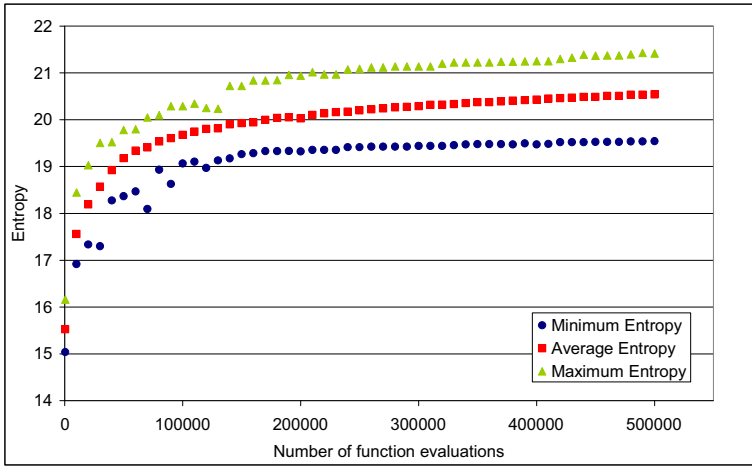
(a) Maximum total entropy (feasible solutions only)
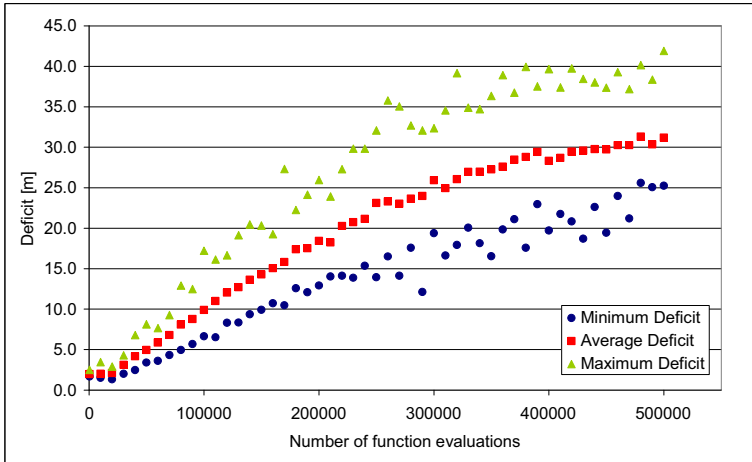
(b) Average deficit (infeasible solutions only)

Fig. 3 Convergence properties. The plots are based on 30 optimization runs

driven analysis are misleading in that the corresponding energy loss due to pipe friction exceeds the total energy of the system. Perhaps this apparent anomaly could be avoided in the future by adopting pressure-driven analysis for more realistic simulations of the infeasible solutions.

Infeasible solutions help drive the optimization, e.g. by sustaining the search near the feasibility boundaries and avoiding a purely interior search that may be suboptimal (Siew et al. 2014). Figure 3b shows the progress of the infeasible solutions. The mean deficit of the solutions increased as the optimization progressed, as the solutions with the smallest deficits (i.e. the solutions located close to the feasible region) improved and eventually became feasible. It was observed that, after decreasing rapidly initially, the number of feasible solutions increased gradually until the end of the optimization (Czajkowska 2016).
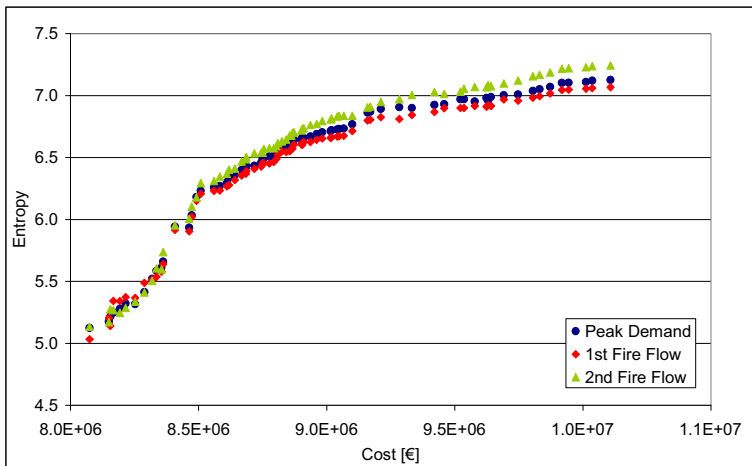
Fig. 4 Entropy values from total entropy maximization

4.3 Effectiveness of the Joint Entropy Formulation

Figure 4 shows the trade-off between flow entropy and cost for the non-dominated feasible solutions. There was a positive coefficient of correlation of 0.991 between the flow entropy and the mean diameter. Larger pipe diameters improve the hydraulic reliability by increasing the pipe flow capacities and lowering the pipe failure rates. This is consistent with previous results in the literature. There is evidence also, of strong positive correlation between the flow entropy and uniformity of the pipe diameters that improves flow re-distribution further.

It is obvious from Fig. 4 that no operating condition had the highest entropy value in all the solutions. For the relatively expensive solutions (€8.5 million and above) Fire Flows 2 and 1 had the highest and lowest entropy values, respectively. On the other hand, Fig. 4 shows that when considering the less expensive solutions (up to €8.5 million) there was no obvious trend. This provides empirical evidence that maximizing the total entropy is the most appropriate
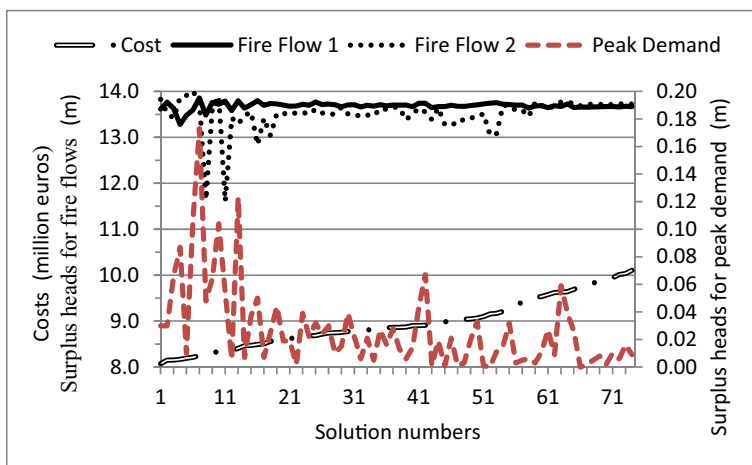

Fig. 5 Nondominated solutions from total entropy maximization

approach for multiple operating conditions. Also, the surplus heads for the fire flows in Fig. 5 reveal that none of the fire flows dominated the other in every solution.

Furthermore, maximizing the largest or smallest entropy value would consider only one operating condition out of three. The operating condition with the largest or smallest entropy could change in successive generations. This could potentially make the algorithm unstable or inefficient through lack of continuity. While the entropy values for different operating conditions were comparable in the present example, there may be significant differences in other situations.

Similar results were achieved on several additional networks in which the total, maximum and minimum entropy maximization approaches were investigated separately. Entropy maximization based on a single loading condition was investigated also. The results showed that multiple operating conditions improved the resilience (i.e. reliability and pipe failure tolerance) and maximizing the sum of the entropies gave the best solutions (Czajkowska 2016).

5 Conclusions

A methodology for flow entropy maximization in the design optimization of water distribution networks under multiple loading conditions was developed and assessed. The empirical results achieved demonstrated that the joint entropy of two or more independent loading conditions is the sum of the separate entropies. It was revealed also that no single loading condition was consistently dominant from the perspective of the flow entropy. The reason is that the critical loading conditions varied from one solution to the next and thus could not be ascertained beforehand. Maximizing the sum of the entropies was, therefore, the most logical approach. These observations are consistent with both the maximum entropy formalism (Jaynes 1957) and the formal definition of the joint entropy of independent probability schemes (Shannon 1948).

A large increase in the number of feasible solutions was achieved compared to previous investigations. It is possible, however, that in the network considered, the total number of feasible solutions is relatively small compared to the size of the solution space. It is also possible that the optimization performed mainly an exterior rather than an interior search within the feasible solution space. Thus some important challenges remain notably the effective incorporation and exploitation of infeasible solutions. Demand-driven hydraulic simulations were used in the optimization. Consequently, the flow entropy values of the infeasible solutions were misleading. Therefore, additional investigations are required. Alternative formulations of the optimization model based on pressure-driven simulation may be worth considering also.

References

Agathokleous A, Christodoulou C, Christodoulou SE (2017) Influence of intermittent water supply operations on the vulnerability of water distribution networks. J Hydroinf 19(6):838–852. https://doi.org/10.2166/hydro.2017.133

Alperovits E, Shamir U (1977) Design of optimal water distribution systems. Water Resour Res 13(6):885–900. https://doi.org/10.1029/WR013i006p00885

Awumah K, Goulter I, Bhatt SK (1990) Assessment of reliability in water distribution networks using entropy based measures. Stoch Hydrol Hydraul 4(4):309–320. https://doi.org/10.1007/BF01544084

Awumah K, Goulter I, Bhatt SK (1991) Entropy-based redundancy measures in water distribution network design. J. Hydraul Eng 117(5): 595–614

Creaco E, Pezzinga G (2015) Embedding linear programming in multi objective genetic algorithms for reducing the size of the search space with application to leakage minimization in water distribution networks. Environ Model Softw 69:308–318. https://doi.org/10.1016/j.envsoft.2014.10.013

Creaco E, Franchini E, Alvisi S (2010) Optimal placement of isolation valves in water distribution systems based on valve cost and weighted average demand shortfall. Water Resour Manag 24(15):4317–4338. https://doi.org/10.1007/s11269-010-9661-5

Creaco E, Franchini E, Alvisi S (2012) Evaluating water demand shortfalls in segment analysis. Water Resour Manag 26(8):2301–2321. https://doi.org/10.1007/s11269-012-0018-0

Cullinane MJ, Lansey KE, Mays LW (1992) Optimization-availability-based design of water distribution networks. J Hydraul Eng 118:420–441

Czajkowska AM (2016) Maximum entropy based evolutionary optimization of water distribution networks under multiple operating conditions and self-adaptive search space reduction method. PhD thesis, University of Strathclyde Glasgow, UK

Czajkowska AM, Tanyimboh TT (2013) Water distribution network optimization using maximum entropy under multiple loading patterns. Wa Sci Technol 13(5):1265–1271

Deb K, Pratap A, Agarwal S, Meyarivan T (2002) A fast and elitist multi-objective genetic algorithm: NSGA II. Trans Evol Comput 6(2):182–197. https://doi.org/10.1109/4235.996017

Díaz J, Montoya MC, Hernández S (2016) Efficient methodologies for reliability-based design optimization of composite panels. Adv Eng Softw 93:9–21. https://doi.org/10.1016/j.advengsoft.2015.12.001

Dorigo M, Caro GD, Gambardella LM (1999) Ant algorithms for discrete optimization. Art&Life 5(2):137–172. https://doi.org/10.1162/106454699568728

Dridi L, Parizeau M, Maihot A, Villeneuve J-P (2008) Using evolutionary optimization techniques for scheduling water pipe renewal considering a short planning horizon. Comput-aided civ Infrastruct Eng 23(8):625–635

Farmani R, Walters G, Savic D (2006) Evolutionary multi-objective optimization of the design and operation of water distribution network: total cost vs. reliability vs. water quality. J Hydroinf 8(3):165–179

Fonseca CM, Fleming PJ (1993) Genetic algorithms for multiobjective optimization: formulation, discussion and generalization. International conference on genetic algorithms. Vol. 5: 416-423, Los Altos, CA

Forrester AIJ, Sobester A, Keane AJ (2008) Engineering design via surrogate modelling. Wiley. https://doi.org/10.1002/9780470770801

Geem JZ, Kim W, Loganathan GV (2001) A new heuristic optimization algorithm: harmony search. SIMULATION 76(2):60–68. https://doi.org/10.1177/003754970107600201

Gheisi A, Naser G (2013) On the significance of maximum number of component failures in reliability analysis of water distribution systems. Urban Water 10(1):10–25. https://doi.org/10.1080/1573062X.2012.690433

Gheisi A, Naser G (2015) Multistate reliability of water-distribution systems: comparison of surrogate measures. J Water Res Pl-ASCE 141(10):04015018. https://doi.org/10.1061/(ASCE)WR.1943-5452.0000529

Glover F, Greenberg H (1989) New approaches for heuristic search: a bilateral linkage with artificial intelligence. Eur J Oper Res 39:19–130

Goldberg DE, Kuo CH (1987) Genetic algorithms in pipeline optimization. J Comput Civ Eng 2:128–141

Haghighi A, Samani HM, Samani ZM (2011) GA-ILP method for optimization of water distribution networks. Water Resour Manag 25(7):1791–1808. https://doi.org/10.1007/s11269-011-9775-4

Hajela P, Lin CY (1992) Genetic search strategies in multi-criterion optimal design. Structural. Optimization 4(2): 99–107

Harrison CG, Williams PR (2016) A systems approach to natural disaster resilience. Simul Model Pract Theory 65:11–31. https://doi.org/10.1016/j.simpat.2016.02.008

Herrera M, Abraham E, Stoianov I (2016) A graph-theoretic framework for assessing the resilience of sectorised water distribution networks. Water Resour Manag 30(5):1685–1699. https://doi.org/10.1007/s11269-016-1245-6

Holland JH (1975) Adaptation in natural and artificial systems: an introductory analysis with applications to biology, control, and artificial intelligence. University of Michigan Press

Ishibuchi H, Akedo N, Nojima Y (2015) Behavior of multiobjective evolutionary algorithms on many-objective knapsack problems. Trans Evol Comput 19(2):264–283. https://doi.org/10.1109/TEVC.2014.2315442

Jayaram N, Srinivasan K (2008) Performance-based optimal design and rehabilitation of water distribution networks using life cycle costing. Water Resour Res 44(1):W01417. https://doi.org/10.1029/2006WR005316

Jaynes ET (1957) Information theory and statistical mechanics. Physical review 106: 620-630 and 108: 171-190

Jeong HS, Abraham DM (2006) Operational response model for physically attacked water networks using NSGA-II. J Comput Civ Eng 20(5):328–338. https://doi.org/10.1061/(ASCE)0887-3801(2006)20:5(328)

Kennedy J, Eberhart RC (1995) Particle swarm optimization. IEEE International Conference on Neural Networks, Piscataway, 1942-1948

Khinchin AI (1953) The entropy concept in probability theory. Uspekhi Matematicheskikh Nauk 8(3):3–20

Khinchin AI (1957) Mathematical foundations of information theory. Dover, New York, pp 1–28

Knowles JD, Corne DW (2000) Approximating the non-dominated front using the Pareto archived evolution strategy. Evol Comput 8(2):149–172. https://doi.org/10.1162/106365600568167

Korteling B, Dessai S, Kapelan Z (2013) Using information-gap decision theory for water resources planning under severe uncertainty. Water Resour Manag 27(4):1149–1172. https://doi.org/10.1007/s11269-012-0164-4

Kursawe F (1990) A variant of evolution strategies for vector optimization. Parallel problem solving from. Nature 1:183–197

Michalewicz Z (1995) A survey of constraint handling techniques in evolutionary computation methods. McDonnell et al. (Eds.) evolutionary programming IV. MIT press, Massachusetts, pp 135-158

Munavalli GR, Kumar MS (2003) Optimal scheduling of multiple chlorine sources in water distribution systems. J Water Res Pl-ASCE 129(6):493–504. https://doi.org/10.1061/(ASCE)0733-9496(2003)129:6(493)

de Neufville R, Schaake J, Stafford J (1971) Systems analysis of water distribution networks. J. Sanitary engineering division ASCE 97(SA6): 825-842

Osyczka A, Kundu S (1995) A new method to solve generalized multicriteria optimization problems using the simple genetic algorithms. Struct Optimization 10(2):94–99

Prasad TD (2010) Design of pumped water distribution networks with storage. J Water Res Pl-ASCE 136(1):129–132. https://doi.org/10.1061/(ASCE)0733-9496(2010)136:1(129)

Prasad TD, Park NS (2004) Multiobjective genetic algorithms for design of water distribution networks. J Water Res Pl-ASCE 130(1):73–82. https://doi.org/10.1061/(ASCE)0733-9496(2004)130:1(73)

Preis A, Ostfeld A (2008) Multiobjective contaminant sensor network design for water distribution systems. J Water Res Pl-ASCE 134(4):366–377. https://doi.org/10.1061/(ASCE)0733-9496(2008)134:4(366)

Rao Z, Salomons E (2007) Development of a real-time, near-optimal control process for water-distribution networks. J Hydroinf 9(1):25–37. https://doi.org/10.2166/hydro.2006.015

Rossman LA (2000) EPANET 2 User's manual. Water supply and water resources division. National Risk Management Research Laboratory, US EPA, Cincinnati, OH

Saleh SH, Tanyimboh TT (2016) Multi-directional maximum-entropy approach to the evolutionary design optimization of water distribution systems. Water Resour Manag 30(6):1885–1901. https://doi.org/10.1007/s11269-016-1253-6

Schaffer JD (1985) Multiple objective optimization with vector evaluated genetic algorithms. 1st international conference on genetic algorithms. Pittsburgh, PA, USA, pp 93-100

Shannon C (1948) A mathematical theory of communication. Bell Systems Technical J 27(3):379–428. https://doi.org/10.1002/j.1538-7305.1948.tb01338.x

Siew C, Tanyimboh TT, Seyoum AG (2014) Assessment of penalty-free multiobjective evolutionary optimization approach for the design and rehabilitation of water distribution systems. Water Resour Manag 28(2):373–389. https://doi.org/10.1007/s11269-013-0488-8

Simpson AR, Dandy GC, Murphy LJ (1994) Genetic algorithms compared to other techniques for pipe optimization. J Water Res Pl-ASCE 120(4):423–443. https://doi.org/10.1061/(ASCE)0733-9496(1994)120:4(423)

Spiliotis M, Tsakiris G (2012) Water distribution network analysis under fuzzy demands. Civ Eng Environ Syst 29(2):107–122. https://doi.org/10.1080/10286608.2012.663359

Srinivas N, Deb K (1994) Multiobjective optimization using non-dominated sorting in genetic algorithms. Evol Comput 2(3):221–248. https://doi.org/10.1162/evco.1994.2.3.221

Storn R, Price K (1997) Differential evolution – a simple and efficient heuristic for global optimization over continuous spaces. J Glob Optim 11:341–359

Tanyimboh TT (1993) An entropy based approach to the optimum design of reliable water distribution networks. PhD Thesis, University of Liverpool, UK

Tanyimboh TT, Kalungi P (2008) Optimal long term design, rehabilitation and upgrading of water distribution systems. Water Sci Technol 39(4):249–255

Tanyimboh TT, Setiadi Y (2008) Sensitivity analysis of entropy constrained designs of water distribution systems. Eng Optim 40(5):439–457. https://doi.org/10.1080/03052150701804571

Tanyimboh TT, Templeman AB (1993a) Calculating maximum entropy flows in networks. J Oper Res Soc 44(4): 383–396

Tanyimboh TT, Templeman AB (1993b) Optimum design of flexible water distribution networks. Civ Eng Syst 10(3):243–258. https://doi.org/10.1080/02630259308970126

Tanyimboh TT, Templeman AB (2000) A quantified assessment of the relationship between the reliability and entropy of water distribution systems. Eng Optim 33(2):179–199. https://doi.org/10.1080/03052150008940916

Tanyimboh TT, Templeman AB (2010) Seamless pressure-deficient water distribution system model. J Water Manage 163(8):389–396

Tanyimboh T, Tabesh M, Burrows R (2001) An appraisal of source head methods for calculating the reliability of water distribution networks. J Water Res Pl-ASCE 127(4):206–213. https://doi.org/10.1061/(ASCE)0733-9496(2001)127:4(206)

Tanyimboh TT, Siew C, Saleh S, Czajkowska A (2016) Comparison of surrogate measures for the reliability and redundancy of water distribution systems. Water Resour Manag 30(10):3535–3552. https://doi.org/10.1007/s11269-016-1369-8

Todini E (2000) Looped water distribution networks using resilience index based on heuristic approach. Urban Water 2(3):115–122. https://doi.org/10.1016/S1462-0758(00)00049-2

Tsakiris G, Spiliotis M (2014) A Newton–Raphson analysis of urban water systems based on nodal head-driven outflow. Eur J Environ Civ Eng 18(8):882–896. https://doi.org/10.1080/19648189.2014.909746

Tsakiris G, Spiliotis M (2017) Uncertainty in the analysis of urban water supply and distribution systems. J Hydroinf 19(6):823–837. https://doi.org/10.2166/hydro.2017.134

Vairavamoorthy K, Ali M (2005) Pipe index vector: a method to improve genetic algorithm based pipe optimization. J Hydraul Eng 131(12):1117–1125

Vamvakeridou-Lyroudia LS, Walters GA, Savic DA (2005) Fuzzy multiobjective optimization of water distribution networks. J Water Res Pl-ASCE 131(6):467–476. https://doi.org/10.1061/(ASCE)0733-9496(2005)131:6(467)

Weickgenannt M, Kapelan Z, Blokker M, Savic D (2010) Risk-based sensor placement for contaminant detection in water distribution systems. J Water Res Pl-ASCE 136(6):629–636. https://doi.org/10.1061/(ASCE)WR.1943-5452.0000073

Wu ZY, Simpson AR (2001) Competent genetic-evolutionary optimization of water distribution systems. J Comput Civ Eng 15(2):89–101

Yan D, Werners SE, Huang HQ, Ludwig F (2016) Identifying and assessing robust water allocation plans for deltas under climate change. Water Resour Manag 30(14):5421–5435. https://doi.org/10.1007/s11269-016-1498-0

Yates DF, Templeman AB, Boffey TB (1984) The computational complexity of determining least capital cost designs for water supply networks. Eng Optim 7(2):143–155. https://doi.org/10.1080/03052158408960635

Zitzler E, Thiele L (1998) An evolutionary algorithm for multiobjective optimization: the strength Pareto approach. Technical report 43, Zurich, Switzerland. Computer engineering and networks laboratory (TIK), Swiss Federal Institute of technology

Zitzler E, Deb K, Thiele L (2000) Comparison of multiobjective evolutionary algorithms: empirical results. Evol Comput 8(2):173–195. https://doi.org/10.1162/106365600568202