Informational Entropy: a Failure Tolerance and Reliability Surrogate for Water Distribution Networks

Tiku T. Tanyimboh [1,2]

Abstract Evolutionary algorithms are used widely in optimization studies on water distribution networks. The optimization algorithms use simulation models that analyse the networks under various operating conditions. The solution process typically involves cost minimization along with reliability constraints that ensure reasonably satisfactory performance under abnormal operating conditions also. Flow entropy has been employed previously as a surrogate reliability measure. While a body of work exists for a single operating condition under steady state conditions, the effectiveness of flow entropy for systems with multiple operating conditions has received very little attention. This paper describes a multi-objective genetic algorithm that maximizes the flow entropy under multiple operating conditions for any given network. The new methodology proposed is consistent with the maximum entropy formalism that requires active consideration of all the relevant information. Furthermore, an alternative but equivalent flow entropy model that emphasizes the relative uniformity of the nodal demands is described. The flow entropy of water distribution networks under multiple operating conditions is discussed with reference to the joint entropy of multiple probability spaces, which provides the theoretical foundation for the optimization methodology proposed. Besides the rationale, results are included that show that the most robust or failure-tolerant solutions are achieved by maximizing the sum of the entropies.

Keywords Water distribution systems · Resilience · Discrete penalty-free constrained evolutionary optimization · Maximum entropy formalism · Uncertainty · Hydraulic reliability

---

✉ Tiku T. Tanyimboh
   tikutanyimboh@hotmail.co.uk

[1] Department of Civil and Environmental Engineering, University of Strathclyde, Glasgow, UK

[2] School of Civil and Environmental Engineering, University of the Witwatersrand, Johannesburg, South Africa

1 Introduction

Water distribution systems are part of the critical economic networks on which modern-day societies depend, and it is widely accepted that there is considerable uncertainty associated with their planning, design and operation. The uncertainty arises from many factors, for example, the deterioration in the structural integrity and hydraulic capacity, unpredictable demands for fire-fighting and random fluctuations in the demands in addition to the underlying temporal and spatial variations, to name but a few. Hence, it is widely accepted that, ideally, explicit criteria for hydraulic capacity reliability and failure tolerance should be included in the design specifications for water distribution systems.

However, it is often difficult to set such performance criteria. Moreover, quantified reliability measures for water distribution systems are particularly difficult to define and evaluate (Wagner et al. 1988). It is even more challenging to incorporate reliability in the procedures used to optimize the design of water distribution systems. Consequently, some surrogate reliability measures such as flow entropy and resilience index have been adopted as they are computationally less demanding (Templeman 1982; Yates et al. 1984).

Flow entropy has been suggested as a surrogate measure of reliability and failure tolerance for water distribution systems. Flow entropy has the advantages that it is relatively easy to calculate and the data it requires are minimal. Also, it lends itself to direct incorporation into formal optimization procedures. Flow entropy is an extension of Shannon's statistical entropy that is a measure of the amount of uncertainty that a probability distribution represents (Shannon 1948).

In the context of water distribution systems, uncertainty may arise due to a variety of reasons including incomplete or imprecise data, future demand requirements and the network's structural integrity and hydraulic capacity that cannot be described or predicted with sufficient accuracy due to incomplete knowledge or understanding of the relevant factors or processes. Furthermore, the uncertainty due to randomness is unavoidable. Examples include the pipe and other component failures that may be related to extraneous factors also, and random fluctuations in demand.

Many different approaches have been developed including stochastic and reliability-based methods that address aleatoric uncertainty due to randomness (Gupta and Bhave 1996; Xu and Goulter 1999a; Laguna et al. 2000; Tolson et al. 2004; Afshar et al. 2005; Kapelan et al. 2005; Giustolisi et al. 2009; Saleh and Tanyimboh 2013) and fuzzy sets for epistemic uncertainty due to incomplete or imprecise information (Xu and Goulter 1999b; Revelli and Ridolfi 2002; Bhave and Gupta 2004; Vamvakeridou-Lyroudia et al. 2005; Gupta and Bhave 2007; Branisavljević et al. 2009; Fu and Kapelan 2011; Shibu and Reddy 2011; Spiliotis and Tsakiris 2012; Sivakumar et al. 2015).

Also, scenario-based robust optimization approaches that address the uncertainty associated with the various scenarios seek solutions that provide the best compromise with respect to the scenarios considered (Watkins and McKinney 1997; Afonso and Cunha 2007; Cunha and Sousa 2010; Napolitano et al. 2016).

Other surrogate measures in the literature include: Tsallis entropy (Singh and Oh 2015), diameter-sensitive flow entropy (Liu et al. 2016), resilience indices (Todini 2000; Prasad and Park 2004; Jayaram and Srinivasan 2008; Liu et al. 2016) and surplus power factor (Vaabel et al. 2006). The results to date seem to indicate that flow entropy (Tanyimboh and Templeman 1993a) yields the most consistent results (Gheisi and Naser 2015; Tanyimboh et al. 2016; Liu et al. 2016). Recent reviews and comparisons include Liu et al. (2014, 2016), Tanyimboh et al.

(2016), Gheisi and Naser (2015), Atkinson et al. (2014) and Greco et al. (2012). In particular, Gheisi and Naser (2015) emphasized the importance of failure tolerance while Tanyimboh et al. (2016) highlighted the need for more consistency in future comparisons.

Flow entropy was tested on different networks over the years and the function proposed in Tanyimboh and Templeman (1993a) was consistently satisfactory. Nevertheless, most of the work focused on networks with a single operating condition based on steady-state simulation that assumes that nodal demands are constant. It is common practice to use the maximum daily demand and steady-state modelling in designing water distribution networks.

In reality, however, the demands vary with the time of the day and other loading conditions including fire flows that have to be satisfied by the network (Cunha and Sousa 2010; Simpson et al. 1994; Walski et al. 1987; Tanyimboh and Seyoum 2016). Alperovits and Shamir (1977) suggested that, in addition to the maximum daily demand and fire flows, the minimum demand periods should be considered also. Furthermore, Prasad (2010) has shown that even if a network satisfies the peak demands it does not follow that other operating conditions will be satisfied as a result. Prasad (2010) also demonstrated that designs that consider multiple operating conditions are more reliable than those based on a single operating condition. Similarly, maximum entropy designs achieve higher levels of reliability and redundancy by being maximally noncommittal to other operating conditions besides those that are designed for explicitly.

Evolutionary optimization algorithms that rely on simulation models are very popular in many disciplines including water resources (Mora-Melia et al. 2015; Borah and Bhattacharjya 2016; Chang et al. 2016; Masoumi et al. 2016; Shokoohi et al. 2016; Steffelbauer and Fuchs-Hanusch 2016). Simulation models make the inclusion of multiple operating conditions in optimization models relatively straightforward. On the other hand, the flow entropy in systems with multiple operating conditions has not been investigated hitherto.

This paper describes a multi-objective genetic algorithm that maximizes the flow entropy for multiple operating conditions for any given network (Czajkowska 2016). The new methodology proposed is consistent with the maximum entropy formalism (Jaynes 1957) that requires active consideration of "whatever is known". An alternative flow entropy function is presented that considers the relative uniformity of the nodal demands, as opposed to the relative uniformity of the source supplies, and its equivalence to the existing function is established. Also, the flow entropy function for multiple operating conditions is discussed with reference to the joint entropy of a combination of probability schemes to provide the theoretical foundation for the optimization model proposed.

2 The Flow Entropy Function

Shannon (1948) developed the informational entropy function as a statistical measure of the amount of uncertainty that a probability distribution represents. Shannon's entropy function for a single probability space is

$$S = -\sum_{i=1}^{n} p_i \ln p_i \qquad (1)$$

in which S is the entropy; p_i is the probability of the ith outcome; and n is the number of outcomes. Tanyimboh and Templeman (1993a, b) developed the framework that enabled the

pipe flow rates in a water distribution network to be cast as probabilities. Given the pipe flow rates, the entropy function is

$$S = S_0 + \sum_{i=1}^{nn} P_i S_i \tag{2}$$

where S is the entropy; S_0 is the entropy that accounts for the relative contributions of the supply nodes; S_i is the entropy at node i; $P_i \equiv T_i/T$ is the fraction of the total flow through the network that reaches node i; T_i is the total flow that reaches node i; T is the sum of the nodal demands; nn is the number of nodes in the network. The entropy for the relative contributions of the supply nodes is

$$S_0 = -\sum_{i \in I} \frac{Q_{0i}}{T} \ln\left(\frac{Q_{0i}}{T}\right) \tag{3}$$

where Q_{0i} is the inflow at supply node i; I represents the set of supply nodes. Similarly, the entropy at demand node i is

$$S_i = -\frac{Q_{i0}}{T_i} \ln\left(\frac{Q_{i0}}{T_i}\right) \quad -\sum_{ij \in ND_i} \frac{Q_{ij}}{T_i} \ln\left(\frac{Q_{ij}}{T_i}\right); \qquad i = 1, \quad \ldots, \quad nn \tag{4}$$

where Q_{i0} is the demand at node i; Q_{ij} is the flow rate in pipe ij with node i and j as the upstream and downstream nodes, respectively; the set ND_i represents the pipe flows from node i. The entropy reflects the number of paths supplying the various demand nodes and the relative uniformity of the flows in the respective supply paths. The greater the uniformity in the path flows and the number of supply paths, the greater the entropy (Tanyimboh 1993; Yassin-Kassab et al. 1999). Awumah et al. (1990, 1991) proposed several alternatives, but they did not account for the interdependencies in the various processes involved in a satisfactory way (Tanyimboh 1993).

   The entropy also depends on the flow directions in the pipes. The problem of identifying the optimal or most suitable set of flow directions is a complex nonlinear optimization problem (Saleh and Tanyimboh 2014, 2016). Furthermore, previous results suggest that efficient sets of flow directions in networks generally result in high maximum entropy values (Tanyimboh and Setiadi 2008). To clarify further, each feasible set of flow directions has a maximum entropy value and thus a network may have many different maximum entropy values due to the multiplicity of flow directions.

   The first term S_0 in Eq. (2) relates to the relative contributions of the supply nodes. For systems with multiple sources, S_0 may change continuously. For example, the network shown in Fig. 1 has five variable head supply nodes and three demand categories with demand patterns that change with time (Tanyimboh and Seyoum 2016). In addition to these practical considerations, the S_0 term arises because the underlying model comprises multiple inter-related probability spaces and the flow entropy function stems from the conditional entropy concept (Tanyimboh 1993; Khinchin 1953, 1957).

   It may appear, at first glance, that there is an imbalance between the supply nodes and the demand node in Eq. (2). However, a water distribution network combines and/or separates the flows in the various paths in the network repeatedly at successive nodes and junctions. A closer examination reveals that Eq. (2) involves only the flow separation
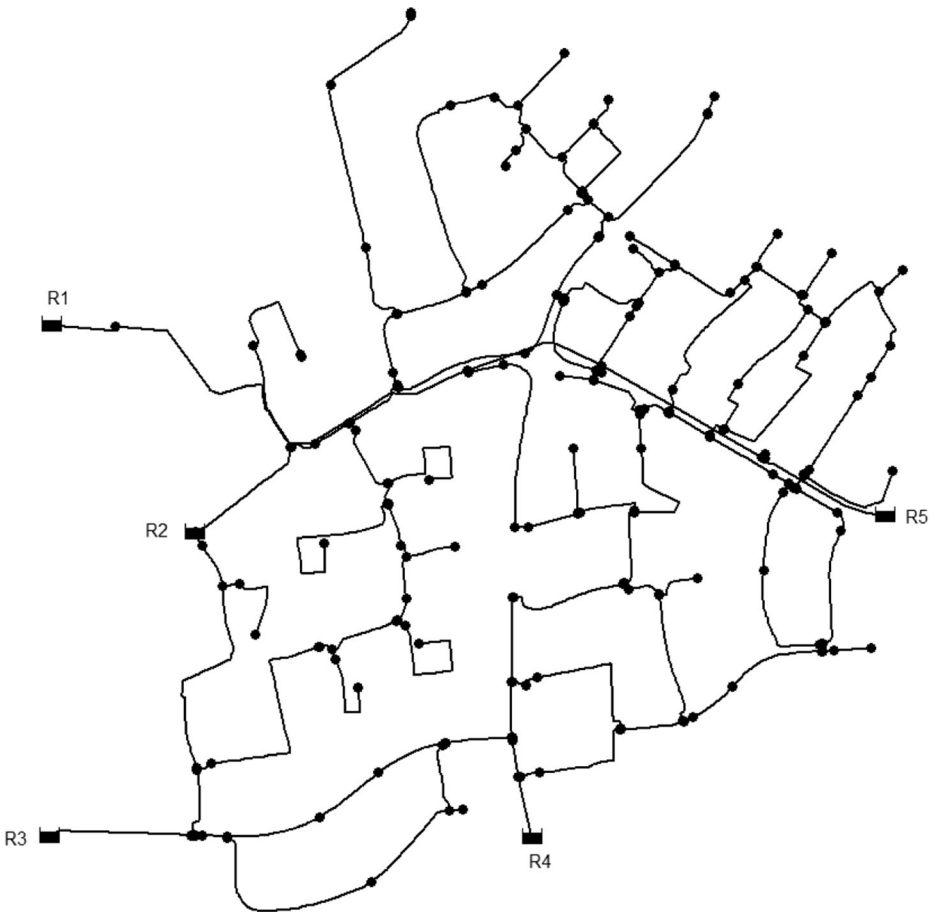
Fig. 1 A water distribution network with five variable head supply nodes (R1 to R5)

points in the network. In fact, there is an analogous function for the flow collection points that may be summarised briefly as follows (Tanyimboh 1993: 106–111).

$$S' = S'_0 + \sum_{i=1}^{nn} P'_i S'_i \tag{5}$$

$$S'_0 = -\sum_{i \in D} \frac{Q_{i0}}{T} \ln\left(\frac{Q_{i0}}{T}\right) \tag{6}$$

$$S'_i = -\frac{Q_{0i}}{T_i} \ln\left(\frac{Q_{0i}}{T_i}\right) - \sum_{ij \in NU_i} \frac{Q_{ij}}{T_i} \ln\left(\frac{Q_{ij}}{T_i}\right); \qquad i = 1, \quad \ldots, \quad nn \tag{7}$$

where S' is the entropy of the network and S'_0 is the entropy or uncertainty that arises due to the distribution, or relative contributions, of the nodal demands. S'_i is the entropy at node i, based on the inflows at node i. The inflows include any source supplies as the first term in Eq. (7)

shows. NU_i refers to the set of pipes upstream of node i that supply node i directly. T_i is the total flow that reaches node i and T is the sum of the demands. $P_i' \equiv P_i = T_i/T$ is the fraction of the total flow through the network that reaches node i.

As one might expect, $S' \equiv S$. In other words, the total amount of uncertainty associated with the path flow separation and collection processes is the same. However, in general $S_i' \neq S_i$ and $S_0' \neq S_0$.

3 Maximum Entropy Flows in Networks

Flow entropy is often used as a practical and computationally efficient surrogate measure for reliability and redundancy. It has been shown previously that on average reliability and redundancy increase as flow entropy increases in water distribution networks (Gheisi and Naser 2015). However, flow entropy is a relative measure for which there is no absolute scale.

Hence, it is frequently necessary to ascertain the maximum entropy value for any network under consideration. It its simplest form, the problem involves maximizing the entropy subject to the continuity of flow at the nodes and junctions, given the topology of the network and direction of flow in the pipes as follows.

$$\text{Maximize}_{\forall Q_{ij}} \quad S = S_0 + \sum_{i=1}^{nn} P_i S_i \tag{8}$$

Subject to:

$$\sum_{j \in N_i} Q_{ij} + Q_{0i} = 0; \qquad \forall i \in I \tag{9}$$

$$\sum_{j \in N_i} Q_{ij} + Q_{i0} = 0; \qquad \forall i \in D \tag{10}$$

$$Q_{ij} \geq 0; \qquad \forall ij \tag{11}$$

Q_{0i} and Q_{i0} are supplies and demands at the supply and demand nodes, respectively. The sets I and D represent the supply and demand nodes, respectively. The pipe flow rates Q_{ij} are the decision variables; the set N_i comprises the upstream or downstream nodes of the pipes connected to node i.

The model in Eqs. (8) to (11) is nonlinear and convex; i.e. it has a unique maximum entropy value for any given topology and set of flow directions (Tanyimboh 1993: 128). The computational solution of the optimization problem requires numerical nonlinear optimization. Efficient algorithms that avoid numerical optimization have been developed that are quick and non-iterative (Tanyimboh and Templeman 1993b; Walters et al. 1995; Yassin-Kassab et al. 1999; Ang and Jowitt 2005a, b). More specifically, Yassin-Kassab et al. (1999) described the fundamental property of maximum entropy flows that underpins the α-method they developed. Whereas Yassin-Kassab et al. (1999) carried out numerical experiments, Ang and Jowitt (2005a, b) provided a rigorous mathematical proof that they then used to develop path entropy methods.

The α and path entropy methods relate to a single operating condition, in other words, one set of demands under steady state conditions. They also require the flow directions in the pipes to be specified in advance. In the case of water distribution systems, in general the flow directions may not be available. For example, at the planning or design stages the topology and/or pipe diameters may not be available initially.

The global maximum entropy approach formulated recently by Saleh and Tanyimboh (2014, 2016) shows how the topology and/or flow directions along with the pipe diameters and flow rates may be optimized simultaneously. This is a key contribution in the development of flow entropy as it addresses a complex issue in practical applications.

4 Flow Entropy under Multiple Operating Conditions

In practice water distribution systems do not operate under steady state conditions. There are other loading conditions besides the daily demand variations e.g. fire-fighting flows. Every operating condition has an entropy value and an appropriate methodology for maximizing the entropy is required. Maximizing the entropy for every operating condition may result in too many objectives. Optimization problems with many objectives are extremely difficult to solve; as the proportion of solutions that are nondominated increases very rapidly and disproportionately as the objectives increase in number (Yuan et al. 2016).

An important property of informational entropy is that the joint entropy of a combination of independent probability schemes is the sum of their entropies (Tanyimboh 1993: 73–77, Shannon 1948). Furthermore, Jaynes (1957) stated that "in making inference on the basis of partial information, we must use that probability distribution that has maximum entropy subject to whatever is known". In other words, maximizing the entropy results in the least biased solution. Thus the formal definition of the *joint entropy* and the requirement to include *whatever is known* suggest that the flow entropy under multiple operating conditions is the sum of the entropies. An illustrative example in Section 5 includes results that appear to support this hypothesis.

5 Network Design under Multiple Operating Conditions

Results in the literature (e.g. Gheisi and Naser 2015, etc.) show that flow entropy is an effective surrogate measure of hydraulic reliability and failure tolerance. However, most of the research to date has focussed on a single operating condition and it is common practice to use the maximum daily demand based on steady state analysis. In reality the nodal demands follow a diurnal pattern and other loading patterns have to be satisfied also. Alperovits and Shamir (1977) suggested that the minimum daily demand be considered in addition to the maximum daily demand and fire-fighting flows.

A multi-objective evolutionary algorithm that handles multiple operating conditions seamlessly for any given network was developed and applied to three well-known networks in the literature. The results showed that solutions that considered multiple operating conditions outperformed solutions based on a single operating condition (Czajkowska 2016).

5.1 Formulation of the Optimization Model

The objectives are the initial construction cost that is to be minimized and flow entropy that is to be maximized subject to adequate flow and pressure at the demand nodes. The nodal mass balance and energy conservation constraints were satisfied externally using the EPANET 2 hydraulic solver (Rossman 2000). Pipe diameters were selected from a set of commercially available discrete pipe sizes. A solution is considered feasible if the residual pressures at all the

demand nodes are greater than or equal to the required pressures for the respective operating conditions.

The minimum node pressure constraints were transformed into an infeasibility objective that was minimized. In this way the constrained optimization problem was converted and solved as an unconstrained problem without introducing any constraint violation penalties as penalty-free genetic algorithms have achieved better results than other algorithms in the literature consistently (Saleh and Tanyimboh 2013; Siew et al. 2014, 2016).

The optimization problem may be summarized as follows.

$$\text{Minimize the total cost}: f_1 = \sum_{i=1}^{np} C_i(D_i, L_i) \tag{12}$$

$$\text{Minimize the largest node pressure deficit}: f_2 = Max\{max[0, (H_i^{des}-H_i)]; \forall i\} \tag{13}$$

$$\text{Maximize the flow entropy}: f_3 = S_{MOC} \tag{14}$$

where S_{MOC} is the flow entropy based on all the operating conditions considered and is to be established in this research; $C_i (D_i, L_i)$ is the cost associated with pipe i of diameter D_i and length L_i; np represents number of pipes; H_i is the available head at node i; and H_i^{des} is the required residual head at node i. The required residual head at a node is the head above which the demand is satisfied in full. The infeasibility function f_2 in Eq. (13) represents the largest node pressure deficit.

Three alternatives were considered for the entropy S_{MOC} as follows:

Case I.   To maximize the maximum entropy, the highest entropy value achieved was chosen, considering all the operating conditions. This option seeks feasible solutions with high entropy values in any operating condition, without attaching any weight to the rest of the entropy values. This option seems unduly optimistic and may be costly in certain circumstances.

Case II.   If the smallest entropy value achieved in any operating conditions is maximized, then it would alleviate the adverse effects of the worst-case failures. This option seems unduly pessimistic.

Case III.   By contrast, maximizing the sum of the entropies would consider all entropy values with an expectation that the resulting solutions would be the most robust or failure tolerant in all the operating conditions considered.

5.2 Solution Methodology

5.2.1 Description of the Optimization Algorithm

The non-dominated sorting genetic algorithm NSGA II (Deb et al. 2002) has been used widely by many researchers in various disciplines. It is an efficient evolutionary algorithm based on Pareto-dominance and global elitism. The general purpose binary coded NSGA II algorithm written in the C++ language was modified and coupled with the hydraulic simulation model

EPANET 2 (Rossman 2000). A subprogram that calculates the flow entropy for any given layout was developed, tested and incorporated in the optimization algorithm. The multiobjective genetic algorithm thus developed can handle multiple operating conditions for any network configuration (Czajkowska 2016).

It was observed that many infeasible solutions were present in the Pareto-optimal fronts achieved. This was expected as the algorithm did not prioritize any of the objective functions and, provided they were nondominated, even infeasible solutions with low cost and high entropy survived until the end of the optimization. This could be beneficial for the decision maker as solutions with a small shortfall in the residual pressure may be worth considering due to budgetary or other reasons.

Furthermore, it is widely known that evolutionary algorithms that deploy both feasible and infeasible solutions in the optimization outperform those that penalise infeasible solutions unduly (Woldesenbet et al. 2009; Saleh and Tanyimboh 2013; Siew et al. 2016). At the end of the optimization, after removing the infeasible solutions in the Pareto-optimal front, a program developed in the Perl language was used to select and sort all the *feasible* solutions, including those from all the preceding generations, based on Pareto-dominance considering entropy and cost.

5.2.2 Reliability and Failure Tolerance Evaluation

The nondominated *feasible* solutions were evaluated further by calculating the hydraulic reliability and pipe failure tolerance, using a pressure-driven analysis program (PRAAWDS) (Tanyimboh et al. 2003; Tanyimboh and Templeman 2010). The program is robust, computationally efficient, and has been tested and used extensively (Czajkowska 2016; Czajkowska and Tanyimboh 2013). An extension developed in the Perl language allows seamless pipe failure simulations for the reliability calculations (Czajkowska 2016).

For a water distribution system, the reliability may be considered a probabilistic measure of the ability to satisfy the required nodal demands at adequate pressure under normal and abnormal operating conditions. The pipe failure tolerance is a complementary measure that provides a probabilistic estimate of the demand that can be satisfied when one or more components are out of service. Accordingly, the hydraulic reliability and failure tolerance were calculated as in Tanyimboh and Templeman (2000).

5.3 Details of the Network Investigated

The network is shown in Fig. 2 and was based on Simpson et al. (1994). The peak demand and two fire flows were considered. Each fire flow comprised the peak demand plus a fire flow at one node, i.e. nodes 7 and 12, for fire flows 1 and 2, respectively (Table 1). Consequently, the minimum nodal head was higher for the peak demand than the fire flows (Walski et al. 1987; Farmani et al. 2006; Cunha and Sousa 2010). The properties of the pipes are shown in Table 2.

The original design problem (Simpson et al. 1994) was to determine the pipe diameters to upgrade and expand an existing network. For the purposes of the present research, all the pipe diameters were optimized. The network comprises 10 demand nodes, 14 pipes and 2 reservoirs. A Hazen-Williams roughness coefficient (Rossman 2000) of 120 was assumed for all the pipes (Table 2). The network is partially branched; demand node 12 has only one incident pipe. The water levels at supply nodes 1 and 5 were 365.76 m and 371.86 m, respectively.
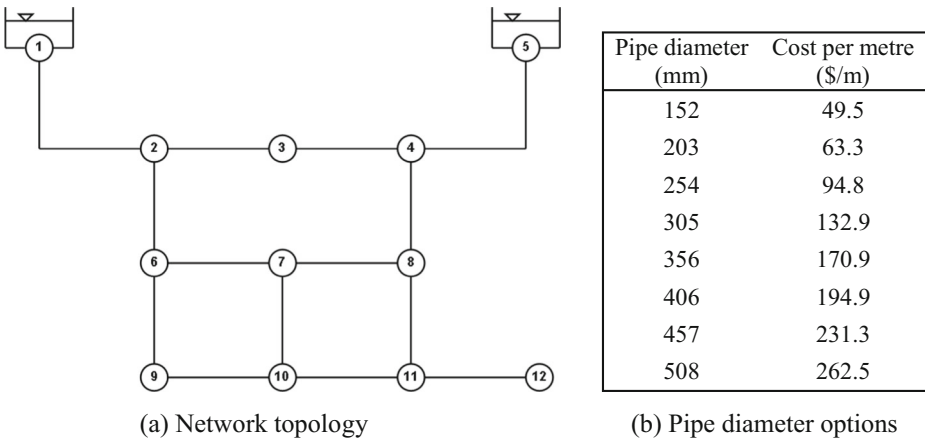
Pipe diameter (mm)	Cost per metre ($/m)
152	49.5
203	63.3
254	94.8
305	132.9
356	170.9
406	194.9
457	231.3
508	262.5

(a) Network topology (b) Pipe diameter options

Fig. 2 Pipe network topology and diameter options

With 8 pipe diameter options and 14 pipes to size, the solution space comprised a total of $8^{14} = 4.398 \times 10^{12}$ feasible and infeasible solutions. A 3-bit binary substring was used. Thus there were $2^3 = 8$ substrings, an exact match for the 8 diameter options with no redundant codes. A single-point crossover operator was used to produce two offspring from two parents. A bitwise mutation operator was used to change the bit from 0 to 1 or vice versa. For each entropy option in Case I to III, 30 GA runs were performed; i.e. 90 GA runs in total. The average CPU time for a single execution of the optimization algorithm was approximately 17 min on a PC (Intel Core 2 Duo @ 3.5GHz and RAM of 3GB).

Extensive testing and sensitivity analysis were carried out to determine suitable values for the parameters used in the optimization algorithm (Czajkowska 2016). The crossover probability was 1.0. The mutation probability was $1/42 = 0.0238$ based on a chromosome length of 42, i.e. a 2.38% chance that any single bit would mutate. The population size was 200. In each optimization run 200,000 function evaluations (i.e. 1000 generations) were allowed.

Table 1 Nodal demands and required residual heads for the network investigated

Node	Elevation (m)	Peak demand		Fire flow 1		Fire flow 2	
		Demand (l/s)	Required head (m)	Demand (l/s)	Required head (m)	Demand (l/s)	Required head (m)
2	320.04	12.62	28.18	12.62	14.09	12.62	14.09
3	326.14	12.62	17.61	12.62	14.09	12.62	14.09
4	332.23	0	17.61	0	14.09	0	14.09
6	298.70	18.93	35.22	18.93	14.09	18.93	14.09
7	295.66	18.93	35.22	82.03	10.57	18.93	14.09
8	292.61	18.93	35.22	18.93	14.09	18.93	14.09
9	289.56	12.62	35.22	12.62	14.09	12.62	14.09
10	289.56	18.93	35.22	18.93	14.09	18.93	14.09
11	292.61	18.93	35.22	18.93	14.09	18.93	14.09
12	289.56	12.62	35.22	12.62	14.09	50.48	10.57

The water levels at supply nodes 1 and 5 were 365.76 m and 371.86 m, respectively

Table 2 Pipe properties for the network investigated

Pipe number	Start node	End node	Length (m)	Hazen-Williams coefficient
1	1	2	4828	120
2	2	3	1609	120
3	3	4	1609	120
4	4	5	6437	120
5	2	6	1609	120
6	4	8	1609	120
7	6	7	1609	120
8	7	8	1609	120
9	6	9	1609	120
10	7	10	1609	120
11	8	11	1609	120
12	9	10	1609	120
13	10	11	1609	120
14	11	12	1609	120

5.4 Results and Discussion

It was observed that the ranges of the entropy values achieved by the various flow entropy options (Case I to III) were different. Also, the Pareto-optimal fronts had slightly different shapes. With three different operating conditions having three different sets of nodal demands and residual pressure requirements, this outcome was not entirely surprising and, indeed, the result seems to support the hypothesis that the entropy should include all the operating conditions, i.e. Case III.

It was observed that a large number of solutions for which the increase in cost was very high but the improvement in entropy was insignificant were at the upper ends of the entropy ranges. This is consistent with previous research in the literature on the trade-off between cost and reliability. In other words, after the improvements taper off, any further improvements become insignificant while the cost increases excessively (Vamvakeridou-Lyroudia et al. 2005, Czajkowska 2016, etc.). Therefore, cut-off points for entropy were set at 99% of the respective maximum entropy values. The resulting performance indicators are summarised in Table 3. The peak demands were used to evaluate the hydraulic reliability and failure tolerance. Unlike the fire flows that are extreme situations, the peak loading would be expected to occur frequently.

Table 3 Coefficient of determination for the relationships indicated

Properties and criteria	Total entropy	Maximum entropy	Minimum entropy
(a) Coefficient of determination			
Entropy vs. mean pipe diameter[a]	0.952	0.857	0.872
Cost vs. entropy nondominated solutions	0.651	0.599	0.007
Cost vs. reliability nondominated solutions	0.825	0.660	0.832
Cost vs. failure tolerance nondominated solutions	0.898	0.814	0.832
(b) Numbers of solutions in the categories stated			
Feasible solutions	1894	1364	2057
Cost vs. entropy nondominated (CEND) solutions	53	41	73
Cost vs. reliability nondominated (CRND) solutions	10	5	8
Cost vs. failure tolerance nondominated (CFTND) solutions	5	4	8

[a] The mean pipe diameter is a simplistic indicator; it does not consider the spatial aspects, etc.

Previous research has shown that both the mean and uniformity of the pipe diameters increase as the flow entropy increases. Larger and more uniform pipe diameters improve the hydraulic reliability by providing lower pipe failure rates and larger flow re-routing capacities (Czajkowska and Tanyimboh 2013; Tanyimboh and Setiadi 2008). The correlation between entropy and the mean pipe diameter was high for all the entropy options (Case I to III), with the total entropy having the highest value of 0.952.

Figure 3 shows the relationship between the hydraulic reliability and entropy for Case II that maximized the minimum entropy, for solutions that are nondominated based on the trade-off between cost and flow entropy (CEND), of which the solutions that are nondominated based on the trade-offs between cost and reliability (CRND) and between cost and failure tolerance (CFTND) have been identified also. Table 3 shows the corresponding coefficients of determination.

The minimum entropy maximization option (Case II) generated the largest number of feasible solutions (Table 3). However, for the cost-vs-entropy nondominated solutions (CEND), the coefficient of determination of 0.007 for this option (Case II) suggests there is practically no correlation. In fact negative correlation can be seen in Fig. 3 between entropy values of about 2.8 and 3.0. A general pattern of negative correlation can be seen also in the bottom right corner. Overall, the results seem to show that maximization of the minimum entropy (Case II) improves the reliability somewhat, but does not lead to the most effective solutions compared to Case I and III.

It may be noted, however, that the entropy values in Fig. 3 from around 3.0 and above seem to belong to a separate cluster with good correlation as can be seen in the graph. Indeed, it was observed that all the entropy maximization options had clusters of solutions located around the highest entropy values. This could be due to the NSGA II optimization algorithm or the optimization model itself (Eqs. 12–14). Despite using solutions up to 99% of the respective maximum entropy values, the majority of the solutions, with entropy values of more than about 3.0 were close to the highest entropy values. This may have influenced the reliability-entropy relationship achieved and may be an area for further research in the future. Also, reintroducing the nondominated feasible solutions from all the previous generations may have
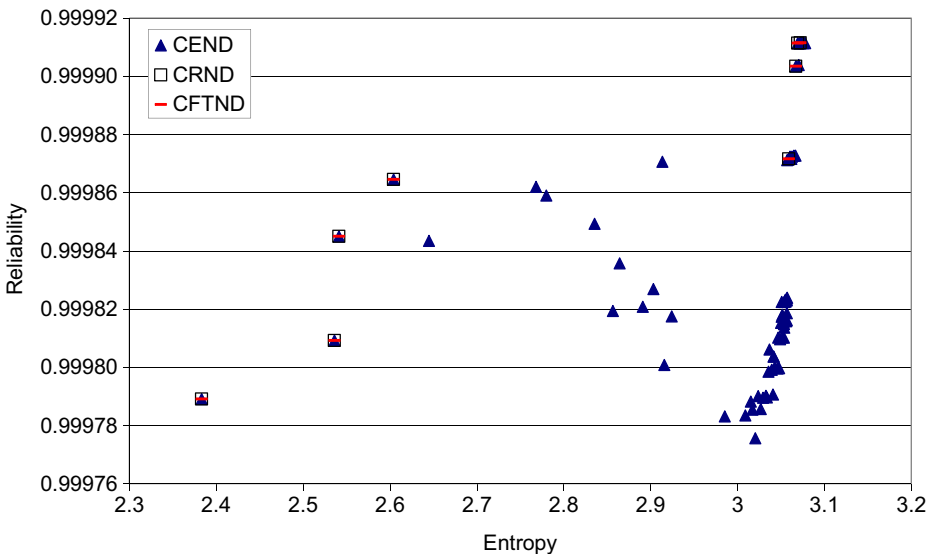

Fig. 3 Hydraulic reliability vs. entropy for the minimum entropy maximization option

altered the distribution of solutions in the final nondominated set. Additional investigations may be needed to verify this.

Furthermore, the observed preponderance of high entropy solutions is consistent with previous comments by Saleh and Tanyimboh (2014, 2016) who found that the high entropy solutions were over-represented if the entropy was maximized on its own as a separate objective function rather than within the hydraulic performance objective function they used. Hence, they developed a multi-directional approach to maximize the *global* and *local* maximum entropy values simultaneously. A more even distribution of the entropy values was achieved as a result. The *local* maximum entropy values are associated with the various alternative sets of feasible flow directions in the network.

Finally, Table 3 shows that maximization of the sum of the entropies achieved better results than maximization of the maximum entropy; it can be seen that the correlations were consistently stronger for the total entropy option (Case III). Maximization of the total entropy therefore seems to be the most robust approach among the three options investigated and reinforces the hypothesis put forward here, that it is essential to take the entropy values from all the operating conditions into consideration.

6 Conclusions

A penalty-free maximum entropy based multi-objective evolutionary optimization approach for water distribution networks under multiple operating conditions has been presented and evaluated. The novelty of this research in the context of entropy maximization is that the optimization works under many loading conditions for any given network. Sensitivity analysis was carried out to assess the robustness of the algorithm and to identify efficient input data values relating to the population size, mutation and crossover rates, and the allocation of redundant binary codes.

Maximization of the sum of the entropies produced the best results, i.e. the most robust or failure-tolerant solutions. The significant advantage of the total entropy option (Case III) is that it takes into consideration all the operating conditions. Jaynes (1957) stated that *the probability distribution that has maximum entropy subject to whatever is known* should be used as it is the least biased. Only the total entropy option is consistent with the maximum entropy principle (Jaynes 1957). Self-evidently, maximization of the maximum entropy or maximization of minimum entropy would place unjustified weight on one operating condition at the expense of all the others and thus would introduce bias.

Finally, it is worth noting that the majority of the solutions achieved were relatively close to the highest entropy values. This may have influenced the reliability-entropy relationship achieved and may be an area for further research in the future. Similarly, Saleh and Tanyimboh (2014, 2016) observed that the high-entropy solutions were over-represented if the entropy was maximized on its own as a separate objective. Hence the global maximum entropy approach they developed may be worth considering in the future as it provided a more even distribution of the entropy values.

Compliance with Ethical Standards

References

Afonso P, Cunha MC (2007) Robust optimal design of activated sludge bioreactors. J Environ Eng 133(1):44–52

Afshar MH, Akbari M, Marino MA (2005) Simultaneous layout and size optimization of water distribution networks: engineering approach. J Infrastructure Systems 11(4):221–223

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

Ang WK, Jowitt PW (2005a) Some new insights on informational entropy for water distribution networks. Eng Optim 37(3):277–289

Ang WK, Jowitt PW (2005b) Path entropy method for multiple-source water distribution networks. Eng Optim 37(7):705–715

Atkinson S, Farmani R, Memon FA, Butler D (2014) Reliability indicators for water distribution system design: comparison. J Water Resour Plan Manag 140(2):160–168

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

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

Bhave PR, Gupta R (2004) Optimal design of water distribution networks for fuzzy demands. Civ Eng Environ Syst 21(4):229–245. doi:10.1080/10286600412331314564

Borah T, Bhattacharjya RK (2016) Development of an improved pollution source identification model using numerical and ANN based simulation-optimization model. Water Resour Manag 30(14):5163–5176. doi:10.1007/s11269-016-1476-6

Branisavljević N, Prodanović P, Ivetić M (2009) Uncertainty reduction in water distribution network modelling using system inflow data. Urban Water J 6(1):69–79. doi:10.1080/15730620802600916

Chang J, Kan Y, Wang Y et al (2016) Conjunctive operation of reservoirs and ponds using a simulation-optimization model of irrigation systems. Water Resour Manage doi. doi:10.1007/s11269-016-1559-4

Cunha MC, Sousa J (2010) Robust design of water distribution network for proactive risk management. J Water Resour Plan Manag 136(2):227–236

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. Water Sci Technol Water Supply 13(5):1265–1271. doi:10.2166/ws.2013.119

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

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. Hydroinformatics 8(3):165–179

Fu G, Kapelan Z (2011) Fuzzy probabilistic design of water distribution networks. Water Resour Res 47: W05538. doi:10.1029/2010WR009739

Gheisi A, Naser G (2015) Multistate reliability of water-distribution systems: comparison of surrogate measures. J Water Resour Plan Manag. doi:10.1061/(ASCE)WR.1943-5452.0000529

Giustolisi O, Laucelli D, Colombo AF (2009) Deterministic versus aleatoric design of water distribution networks. J Water Resour Plan Manag 135(2):117–127. doi:10.1061/(ASCE)0733-9496(2009)135:2(117)

Greco R, Di Nardo A, Santonastaso G (2012) Resilience and entropy as indices of robustness of water distribution networks. J Hydroinf 14(3):761–771

Gupta R, Bhave PR (1996) Reliability-based design of water distribution systems. J Environ Eng 122(1): 51–54

Gupta R, Bhave PR (2007) Fuzzy parameters in pipe network analysis. Civ Eng Environ Syst 24(1):33–54

Jayaram N, Srinivasan K (2008) Performance-based optimal design and rehabilitation of water distribution networks using life cycle costing. Water Resour Res 44:W01417

Jaynes ET (1957) Information theory and statistical mechanics. Phys Rev 106: 620–630 and 108: 171–190

Kapelan ZS, Savic DA, Walters GA (2005) Multiobjective design of water distribution systems under uncertainty. Water Resour Res 41:W11407. doi:10.1029/2004WR003787

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

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

Laguna M, Lino P, Perez A, Quintanilla S, Vals V (2000) Minimizing weighted tardiness of jobs with stochastic interruptions in parallel machines. Eur J Oper Res 127(2):444–457

Liu H, Savić DA, Kapelan Z, Zhao M, Yuan Y, Zhao H (2014) A diameter -sensitive flow entropy method for reliability consideration in water distribution system design. Water Resour Res 50(7):5597–5610

Liu H, Savić DA, Kapelan Z, Creaco E, Yuan Y (2016) Reliability surrogate measures for water distribution system design: a comparative analysis. J Water Resour Plan Manag. doi:10.1061/(ASCE)WR.1943-5452.0000728

Masoumi M, Kashkooli BS, Monem MJ et al (2016) Multi-objective optimal design of on-demand pressurized irrigation networks. Water Resour Manag 30(14):5051–5063. doi:10.1007/s11269-016-1468-6

Mora-Melia D, Iglesias-Rey PL, Martinez-Solano FJ et al (2015) Efficiency of evolutionary algorithms in water network pipe sizing. Water Resour Manag 29:4817–4831

Napolitano J, Sechi GM, Zuddas P (2016) Scenario optimisation of pumping schedules in a complex water supply system considering a cost–risk balancing approach. Water Resour Manag 30(14):5231–5246. doi:10.1007/s11269-016-1482-8

Prasad TD (2010) Design of pumped water distribution networks with storage. J Water Resour Plan Manag 136(1):129–132

Prasad TD, Park NS (2004) Multiobjective genetic algorithms for design of water distribution networks. J Water Resour Plan Manag 130(1):73–82

Revelli R, Ridolfi L (2002) Fuzzy approach for analysis of pipe networks. J Hydraul Eng 128(1):93–101. doi:10.1061/(ASCE)07339429(2002)128:1(93)

Rossman LA (2000) EPANET 2 User's Manual. Water Supply and Water Resources Division, National Risk Management Research Laboratory, Cincinnati, OH45268

Saleh SHA, Tanyimboh TT (2013) Coupled topology and pipe size optimization of water distribution systems. Water Resour Manag. doi:10.1007/s11269-013-0439-4

Saleh SHA, Tanyimboh TT (2014) Optimal design of water distribution systems based on entropy and topology. Water Resour Manag 28(11):3555–3575. doi:10.1007/s11269-014-0687-y

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

Shannon CE (1948) A mathematical theory of communication. Bell System Technical J 27(3):379–428

Shibu A, Reddy J (2011) Uncertainty analysis of water distribution networks by fuzzy-cross entropy method. World Acad Sci Eng Technol 59:724–731

Shokoohi M, Tabesh M, Nazif S et al (2016) Water quality based multi-objective optimal design of water distribution systems. Water Resour Manag. doi:10.1007/s11269-016-1512-6

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. doi:10.1007/s11269-013-0488-8

Siew C, Tanyimboh TT, Seyoum AG (2016) Penalty-free multi-objective evolutionary approach to optimization of Anytown water distribution network. Water Resour Manag 30(11):3671–3688. doi:10.1007/s11269-016-1371-1

Simpson AR, Dandy GC, Murphy LJ (1994) Genetic algorithms compared to other techniques for pipe optimization. J Water Resour Plan Manag 120(4):423–443

Singh VP, Oh J (2015) A Tsallis entropy-based redundancy measure for water distribution networks. Phys A 421: 360–376

Sivakumar P, Prasad RK, Chandramouli S (2015) Uncertainty analysis of looped water distribution networks using linked EPANET-GA method. Water Resour Manag. doi:10.1007/s11269-015-1165-x

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

Steffelbauer DB, Fuchs-Hanusch D (2016) Efficient sensor placement for leak localization considering uncertainties. Water Resour Manag 30(14):5517–5553. doi:10.1007/s11269-016-1504-6

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

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

Tanyimboh TT, Templeman AB (1993b) Maximum entropy flows for single-source networks. Eng Optim 22(1): 49–63

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

Tanyimboh TT, Setiadi Y (2008) Sensitivity analysis of entropy-constrained designs of water distribution systems. Eng Optim 40(5):439–457

Tanyimboh TT, Templeman AB (2010) Seamless pressure-deficient water distribution system model. ICE J Water Management 163(8):389–396. doi:10.1680/wama.900013

Tanyimboh TT, Seyoum AG (2016) Multiobjective evolutionary optimization of water distribution systems: exploiting diversity with infeasible solutions. J. Environ Manag 183:133–141

Tanyimboh TT, Tahar B, Templeman AB (2003) Pressure-driven modelling of water distribution systems. Water Sci Technol Water Supply 3(1–2):255–261

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

Templeman AB (1982) Discussion of "optimization of looped water distribution systems" ASCE J. Environmental Engineering Division 108(3):599–602

Todini E (2000) Looped water distribution networks design using a resilience index based heuristic approach. Urban Water 2(3):115–122

Tolson BA, Maier HR, Simpson AR, Lence BJ (2004) Genetic algorithms for reliability-based optimization of water distribution systems. J Water Resour Plan Manag 130(1):63–72. doi:10.1061/(ASCE)0733-9496(2004)130:1(63)

Vaabel J, Ainola L, Koppel T (2006) Hydraulic power analysis for determination of characteristics of a water distribution system. 8th Annual Water Distribution Systems Analysis Symposium, ASCE, Reston

Vamvakeridou-Lyroudia LS, Waters GA, Savic DA (2005) Fuzzy multiobjective optimization of water distribution networks. J Water Resour Plan Manag 131(6):467–476

Wagner JM, Shamir U, Marks DH (1988) Water distribution reliability: analytical methods. J Water Resour Plan Manag 114(3):253–275

Walski TM, Brill ED, Gessler J, Goulter IC, Jeppson RM, Lansey K, Lee HL, Liebman JC, Mays L, Morgan DR, Ormsbee L (1987) Battle of the network models: epilogue. J Water Resour Plan Manag 113(2):191–203

Walters GA, Templeman AB, Tanyimboh TT (1995) Discussion on: maximum entropy flows in networks. Eng Optim 25:155–163

Watkins DW, McKinney DC (1997) Finding robust solutions to water resources problems. J Water Resour Plan Manag 123(1):49–58

Woldesenbet YG, Yen GG, Tessema BG (2009) Constraint handling in multiobjective evolutionary optimization. IEEE Trans Evol Comput 13(3):514–525

Xu C, Goulter IC (1999a) Reliability based optimal design of water distribution networks. J Water Resour Plan Manag 125(6):352–362

Xu C, Goulter IC (1999b) Optimal design of water distribution networks using fuzzy optimization. Civ Eng Environ Syst 16(4):243–266. doi:10.1080/02630259908970266

Yassin-Kassab A, Templeman AB, Tanyimboh TT (1999) Calculating maximum entropy flows in multi-source, multi-demand networks. Eng Optim 31(6):695–729

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

Yuan Y, Xu H, Wang B, Yao X (2016) A new dominance relation-based evolutionary algorithm for many-objective optimization. IEEE Trans Evol Comput 20(1):16–37