

Reliability Assessment of Water Distribution Systems with Statistical Entropy and other Surrogate Measures

Tiku T. Tanyimboh*, Marika T. Tietavainen** and Salah Saleh*

* Department of Civil Engineering, University of Strathclyde, Glasgow. John Anderson Building, 107 Rottenrow, Glasgow G4 0NG, UK. Email: tiku.tanyimboh@strath.ac.uk, salah.saleh@strath.ac.uk

** Mott MacDonald Limited, 1 Atlantic Quay, Broomielaw, Glasgow G2 8JB, UK.
E-mail: marika.tietavainen@mottmac.com

Abstract

There is ever increasing commercial and regulatory pressure to minimise the cost of water distribution networks even as the demand for them keeps on growing. But cost minimizing is only one of the demands placed on the network design. Satisfactory networks are required to operate above a minimum level even if they experience failure of components. Reliable hydraulic performance can be achieved if sufficient redundancy is built in the network. This has given rise to various water distribution system optimization methods including genetic algorithms and other evolutionary computing methods. Evolutionary computing approaches frequently assess the suitability of enormous numbers of potential solutions for which the calculation of accurate reliability measures could be computationally prohibitive. Therefore, surrogate reliability measures are frequently used to ease the computational burden. The aim of this paper is to assess the correlation of surrogate reliability measures in relation to more accurate measures. The surrogate measures studied are statistical entropy, network resilience, resilience index and modified resilience index. The networks were simulated with prototype software PRAAWDS that produces more realistic results for pressure-deficient water distribution systems. Statistical entropy outperformed resilience index in this study. The results also demonstrate there is a strong correlation between entropy and failure tolerance.

Keywords: Failure tolerance; pressure-deficient water distribution networks; pressure-dependent modelling; redundancy; resilience index; statistical entropy; reliability

INTRODUCTION

One of the major considerations in constructing water distribution networks is the capital cost of the project. To minimise cost of the network modellers endeavour to achieve a balance of smallest possible pipe sizes, tanks and pumps whilst still providing adequate amount of water to meet the demand. However, Walski (2001) questions the benefits of network optimization to achieve minimum cost at the expense of reduced system capacity, and consequently, reliability. He argues that due to uncertainty of future demands and loss in potential project net benefits following cost minimization, water distribution network (WDN) optimization based on cost alone is not viable. Hence another important factor in WDN design is reliability. To guarantee undisrupted water supply even during abnormal conditions, such as fire fighting or network component failure, a WDN has to have some redundancy built in it. But cost and reliability of a water distribution system are at least partly mutually conflicting constraints on design. Studies show that high cost is not always an indicator of high reliability, which means network optimization leading to minimum cost and maximum benefit and efficiency is desirable.

For water distribution networks, however, accurate reliability and failure tolerance or redundancy measures are complicated to calculate because of the considerable computational effort required, and surrogate based measures have been proposed to be used instead. This research concentrated on a selection of such measures including resilience index and statistical entropy. Compared to other surrogate measures the advantages of using statistical

entropy to assess network reliability are that only pipe flows are required to carry out the analysis and the calculations are easy and quick to perform. On the other hand, if the resilience index is used, a full hydraulic analysis of the network has to be done, requiring such data as nodal pressure heads, pipe diameters, lengths and roughness coefficients, node elevations, etc. The intuitive appeal of the resilience index and its relative simplicity in comparison to accurate hydraulic reliability measures has resulted in its widespread use in recent years (e.g. Reca et al. 2008), but no study on its robustness as a reliability indicator exists as yet. The aim of this paper is to assess the effectiveness of surrogate reliability measures in relation to more rigorous and accurate measures. The surrogate measures studied are statistical entropy, network resilience, resilience index and modified resilience index.

METHODS

This paper is based on the hypothetical network shown in Figure 1. Minimum cost designs subject to a maximum entropy constraint were generated for 65 layouts by Tanyimboh and Sheahan (2002) and Tanyimboh and Setiadi (2008a, b) using sequential quadratic programming. Different sets of flow directions were used for some of the layouts thus giving a total of 137 minimum cost maximum entropy (MCME) designs. Continuous pipe diameters in the range of 100mm to 600mm were used. The piezometric head at the source is 100m and the nodes have elevations of 0m. The minimum heads to fully satisfy demands are specified to be 30m. Zero nodal flow takes place if the nodal head drops to 0m. The pipes are 1000m long and the roughness coefficient used is Hazen-Williams at 130. The demands in litres per second are shown in Figure 1. A range of performance indicators were calculated and assessed in the present study for the above-mentioned MCME designs as described briefly below. The hydraulic simulations were carried out using prototype software for pressure-driven analysis called PRAAWDS (see e.g. Tanyimboh et al. 2003).

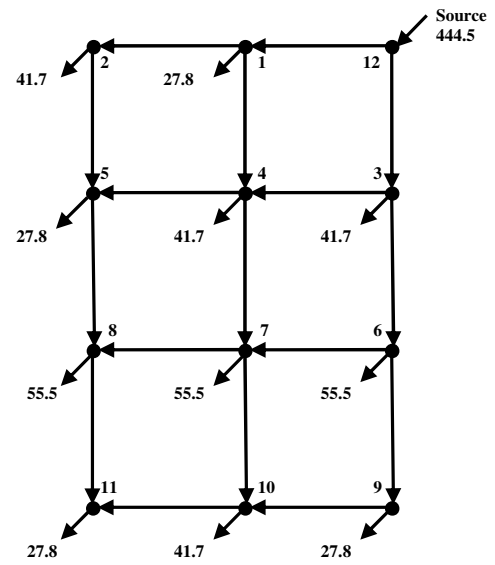


Figure 1 Network layout

Hydraulic Reliability

While there is not one absolute definition for hydraulic reliability in broad terms it can be described as the ability of the water distribution network to satisfy the nodal demands under both normal operating conditions as well as when one or more components of the system experience failure. The reliability equation used in this paper is (Tanyimboh and Templeman 2000)

$$R = \frac{1}{T} \left(p(0)T(0) + \sum_{m=1}^M p(m)T(m) + \sum_{m=1}^{M-1} \sum_{n=m+1}^M p(m,n)T(m,n) + \dots \right) + \frac{1}{2} \left(1 - p(0) - \sum_{m=1}^M p(m) - \sum_{m=1}^{M-1} \sum_{n=m+1}^M p(m,n) - \dots \right) \quad (1)$$

in which R = hydraulic reliability; M = number of links (pipes, pumps and valves); $P(0) = a_1 a_2 a_3 \dots a_M$ = probability that all links are in service; a_m = probability that link m is in service

at any given moment; $p(m) = p(0)(u_m/a_m)$ = probability that only link m is not in service; $u_m = 1 - a_m$ = probability that link m is not in service; $p(m,n) = p(0)(u_m/a_m)(u_n/a_n)$ = probability that only links m and n are not in service; $T(0)$, $T(m)$ and $T(m,n)$ are, respectively, the total flows supplied with all links in service, only link m out of service and only links m and n out of service; and T is the sum of the nodal demands. Pipe availability = a_m = ratio of ‘mean time between failures’ to ‘mean time between failures plus mean failure duration’ can be calculated with various formulae. In this study the formulae proposed by Cullinane et al. (1992) and Khomsi et al. (1996) were used. Due to the limitations of space, only the results based on the Cullinane et al. formula are included here. It should be noted, however, that essentially the same findings were obtained with the Khomsi et al. formulation. Two further pipe availability formulations were considered, namely Fujiwara and Tung (1991) and Su et al. (1987) but they proved unsuitable for the present study as their applicability would appear to be limited to small pipe sizes of up to 300mm and 200mm respectively. The first part of Eq. 1 indicates how much of the demand the system satisfies on average. In general, inclusion of all the possible combinations of multiple component failures in Eq. 1 is impracticable. The second part of Eq. 1 corrects the consequent underestimation inherent in the first part. Since the ratio u_m/a_m is in general small, cases including more than two links out of service are not normally computed. For this study, only the case of a single link out of service at any given moment was simulated.

Hydraulic redundancy, or failure tolerance, is a measure of the proportion of demands satisfied when some of the components in the network are unavailable. It excludes periods when all the links are available, therefore representing spare capacity in the network. The equation for failure tolerance is (Tanyimboh and Templeman 1998, Tanyimboh et al. 2001)

$$FT = \frac{R - p(0)T(0)/T}{1 - p(0)} \quad (2)$$

where FT is the failure tolerance. Tanyimboh et al. (2001) note that failure tolerance may be a better measure of disruption of supply for a network in case of failure than the hydraulic reliability for very reliable systems.

Resilience Index

Todini (2000) introduced the concept of resilience index. The total power in the network is the sum of the power dissipated in the pipes and the power that is delivered to the nodes and the resilience index is a measure of surplus power available to be dissipated in the network internally in case of a failure. Todini (2000) defined the resilience index as

$$RI = \frac{\sum_{i=1}^{n_n} Q_i^{req} (H_i - H_i^{req})}{\sum_{k=1}^{n_r} Q_k H_k + \sum_{j=1}^{n_{pu}} P_j / \gamma - \sum_{i=1}^{n_n} Q_i^{req} H_i^{req}} \quad (3)$$

where RI = resilience index; H_i = head at demand node i ; H_i^{req} = demand node head above which the demand is satisfied in full; Q_i^{req} = nodal demand; γ = specific weight of water; Q_k and H_k = reservoir k supply and head, respectively; P_j = power introduced to the network by pump j ; n_{pu} = number of pumps; n_n = number of demand nodes; and n_r = number of reservoirs.

Jayaram and Srinivasan (2008) questioned appropriateness of the resilience index if used to measure performance in multiple source networks. They pointed out that networks with high

surplus power can also have high power input and since the power input term is denominator in the equation it potentially results in low resilience index even with networks with plenty of surplus power. They proposed an alternative indicator known as modified resilience index that measures the surplus power as a percentage of the power required at the nodes, i.e.

$$MRI = \frac{\sum_{j=1}^{n_n} Q_j^{req} (H_j - H_j^{req})}{\sum_{j=1}^{n_n} Q_j^{req} H_j^{req}} \times 100 \quad (4)$$

in which MRI = modified resilience index. In the present study involving a single-source network the correlation between Eqs. 3 and 4 was almost perfect; MRI essentially rescaled RI and as such did not shed any additional light. Therefore, due to the limitations of space, results for the MRI are not included here.

Network resilience, NR , is a reliability measure developed by Prasad and Park (2004), which extends Todini's concept by combining the effects of both surplus power and reliable loops. Prasad and Park (2004) stated that reliable loops can be ensured if the pipes connected to a node are not widely varying in diameter. They defined the pipe diameter uniformity, for each node, as the ratio of the average of the diameters of the pipes incident at a node to the maximum diameter at that node. The network resilience is thus defined as

$$NR = \frac{\sum_{i=1}^{n_n} C_i Q_i^{req} (H_i - H_i^{req})}{\sum_{k=1}^{n_r} Q_k H_k + \sum_{j=1}^{n_{pu}} P_j / \gamma - \sum_{i=1}^{n_n} Q_i^{req} H_i^{req}} \quad (5)$$

in which C_j is the pipe diameter uniformity for node j .

Statistical Entropy

Entropy as a measure of uncertainty was introduced by Shannon (1948) and applies to simple exhaustive probability schemes in which all the outcomes are independent and the sum of the probabilities is unity. Tanyimboh and Templeman (1993a-c) developed the relevant probabilistic framework for WDNs. Their entropy function is

$$S = S_0 + \sum_{i=1}^{n_n} P_i S_i \quad (6)$$

in which S is the entropy for the WDN as a whole; S_0 is the entropy of source supplies; S_i is the entropy of node i ; $P_i = T_i / T$ is the fraction of the total flow through the network which reaches node i ; T_i is the total flow that reaches node i ; T is the sum of the nodal demands; and n_n is the number of demand nodes. S_0 and S_i are defined as follows.

$$S_0 = - \sum_{i \in I} \frac{Q_{0i}}{T} \ln \left(\frac{Q_{0i}}{T} \right) \quad (7)$$

where Q_{0i} is the supply at source node i ; and I is the set of source nodes.

$$S_i = - \frac{Q_i^{req}}{T_i} \ln \left(\frac{Q_i^{req}}{T_i} \right) - \sum_{ij \in ND_i} \frac{Q_{ij}}{T_i} \ln \left(\frac{Q_{ij}}{T_i} \right) \quad i = 1, \dots, n_n \quad (8)$$

in which Q_{ij} is the pipe flow from node i to node j ; and ND_i is the set of all pipe flows from node i . The derivation and detailed characterizations of S and the other parameters in Eqs. 6-8 can be found in Tanyimboh and Templeman (1993a-c). Recent reviews of entropy applications in WDNs and its numerous advantages can be found in Tanyimboh and Setiadi (2008a, b). It is worth restating that evaluation of Eqs. 6-8 is quick and easy as it requires only the pipe flow rates and nothing else. The MCME optimization process explained previously yielded the minimum cost and maximum entropy for each layout and/or set of flow directions.

RESULTS AND DISCUSSION

Figure 2 clearly demonstrates the positive correlation between entropy and hydraulic reliability, which is in line with previous studies. Figure 3 shows strong positive correlation between entropy and failure tolerance. The result of Figure 3 has not been demonstrated hitherto. The motivation for designing WDNs to carry maximum entropy flows is that MCME designs are by definition maximally noncommittal to any information not explicitly incorporated in the design. They are consequently bias free and better able to cope with a broad range of operating conditions than conventional designs. It is indeed very encouraging that Figure 3 would appear to bear this out in practice. Figure 4 demonstrates a strong positive correlation between reliability and failure tolerance. However, the amount of scatter is such that two WDSs with similar reliability values can have significantly different failure tolerances. Thus, as observed elsewhere in the literature, both measures are best deployed together in a complementary fashion.

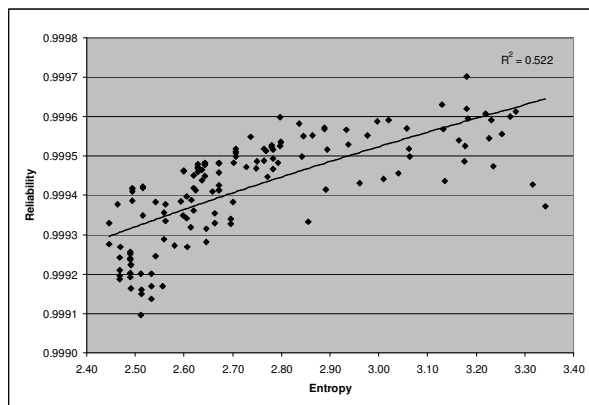


Figure 2 Reliability vs. entropy

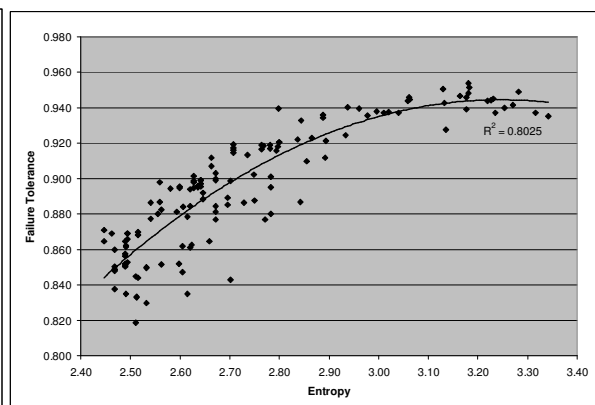


Figure 3 Failure tolerance vs. entropy

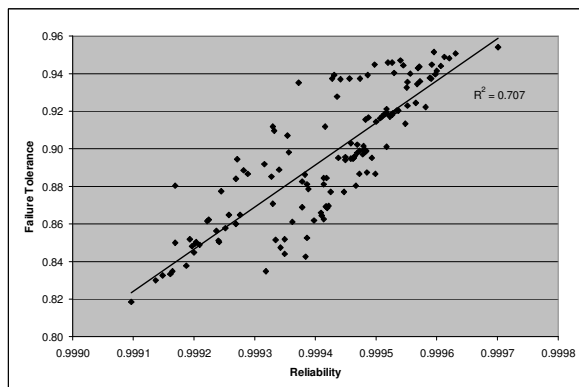


Figure 4 Failure tolerance vs. reliability

The other surrogate reliability measures produced much less correlation and showed a lot more scatter (Figure 5 and 6). Regarding RI and MRI , the power input to the network for all of the 137 designs is constant and therefore the two expressions effectively measure the same proportionality of network surplus power and minimum power required. Since the power input is not changing the RI and MRI lead to mutually consistent results. For assessment of the hypothetical networks of this study the modified resilience index adds no value.

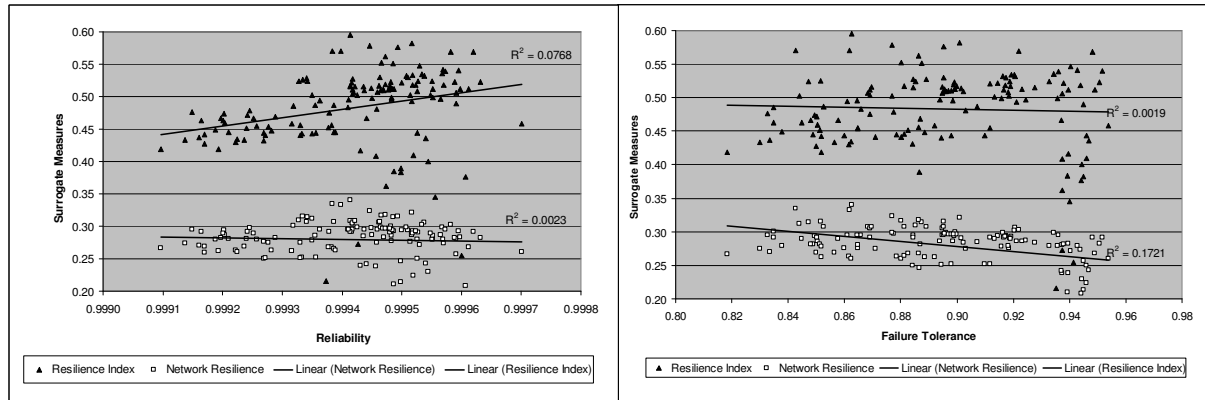


Figure 5 Surrogate measures vs. reliability Figure 6 Surrogate measures vs. failure tolerance

In order for the maximum entropy values achieved to have meaningful practical use they need to be repeatable in different network configurations. Similar entropy values should indicate similar network hydraulic performance regardless of the network layout. Previous studies (Tanyimboh and Sheahan 2002; Tanyimboh and Setiadi 2008b) showed results that seem to confirm low variation in the reliability values for WDN designs based on equal maximum entropy values (derived from different layouts or flow directions). In a similar manner, the present study investigated the similarity of the capital costs within equal maximum entropy groups (EMEGs) as a possible indicator of any similarity in the distribution of pipe diameters between the equal maximum entropy (EME) designs. In Figure 7, EMEGs were assessed in terms of the coefficient of variation of the cost (CVC).

There are 77 designs that fall into 29 EMEGs consisting of a minimum of two and maximum of six designs. The remaining 60 designs formed the non-EMEGs group used as one of the comparators. The CVC was calculated for various categories as shown in Figure 7. Self-evidently Figure 7 shows a single CVC value in each case except for the EMEGs. The aim is to provide a simple and quick visual comparison. The potential range was obtained by including only the smallest and the highest cost values in the calculation based on the fact that some of the EMEGs have only two members. The potential range of CVC for the entire set of 137 designs is expectedly much higher than any of the other categories. The weighted average of all EMEGs was based on the number of designs in each group and the CVC of the group. CVC for all designs belonging to EMEGs is smaller than that of all 137 designs. The non-EMEGs had CVC higher than CVC of all of the 137 designs. It is clear that the smallest CVC takes place in the EMEGs weighted average indicating that groups with higher numbers of designs have overall lower variation of cost, and once all designs are weighted they take the average down. This suggests that designs within EMEGs have high similarity of cost. However, there are EMEGs that have very high CVC, even higher than CVC for all of the 137 designs. But as mentioned, the weighted average of all designs in turn has a very small CVC, which indicates that to balance the occasional high peaks, some EMEGs have nearly exactly the same cost for their designs. These results seem logical and support the hypothesis that entropy can be used to assess the performance of different networks in a reliable manner.

When the EMEGs were studied in more detail, it emerged that of the groups that have lower CVC values than the EMEGs weighted mean, only two out of 29 had designs that do not belong to the 65 *original* designs in Tanyimboh and Sheahan (2002) and the majority of EMEGs with high CVC consist of designs that belong to the 72 *additional* designs in Tanyimboh and Setiadi (2008b). This is most probably due to the less direct and consequently

long flow paths adopted in the 72 *additional* designs that resulted in larger pipe sizes to cater for long flow paths in some of these designs (Tanyimboh & Setiadi, 2008b). If these designs were to be excluded, the CVC for the EMEGs would be even lower. In the network design process any unnecessarily long flow paths would be discarded and therefore high CVC for EMEG could possibly be used to eliminate less than ideal designs. The peaks of CVC within the EMEGs include groups with two to six designs. However, even the groups with only two designs have considerably smaller variation of cost than the potential range of all designs.

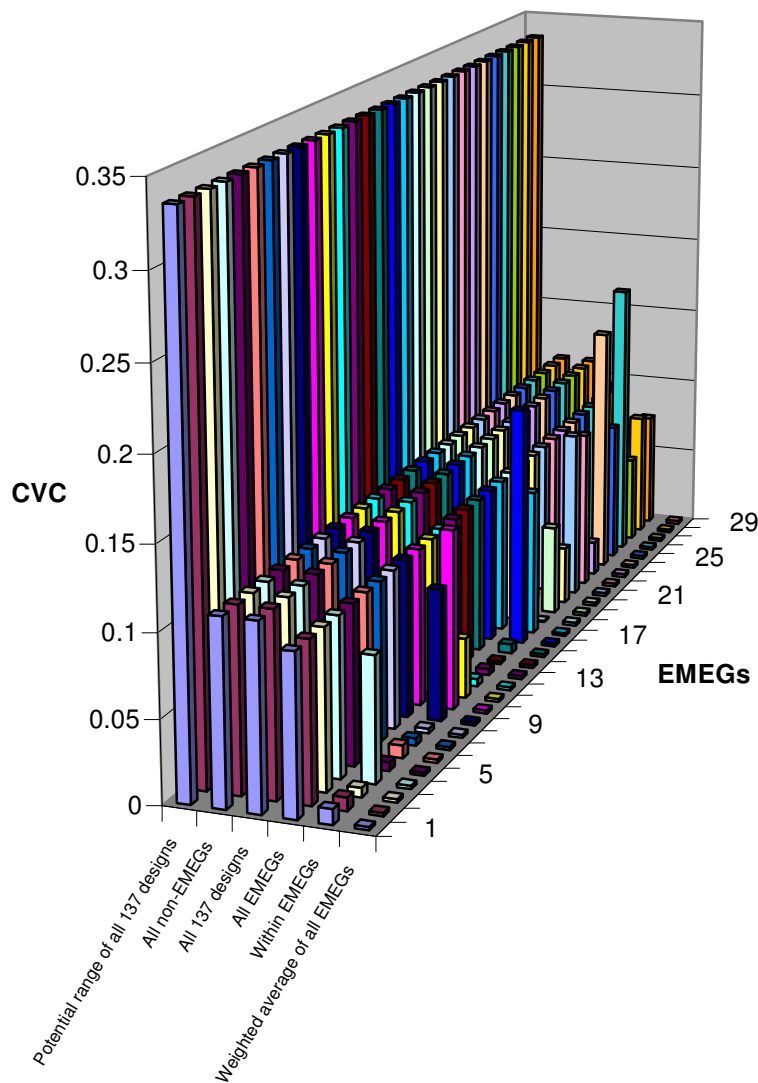


Figure 7 Coefficient of variation of cost

CONCLUSIONS

Entropy clearly outperformed the other surrogate reliability measures in terms of consistency of the results and, overall, hydraulic reliability, failure tolerance and entropy correlated well as a group. Previous studies demonstrated a strong positive correlation between entropy and hydraulic reliability. The present study has, in addition, demonstrated a strong positive correlation between entropy and failure tolerance. However, when the resilience index, network resilience and modified resilience index were plotted against reliability and failure tolerance, these measures often indicated counterintuitive results with decreasing values for increasing reliability and failure tolerance

values. Network resilience, resilience index and modified resilience index often gave confusing values, which needed to be analysed further by inspection of the individual designs. These findings seem to indicate that if the *RI*, *NR* and *MRI* are used, then the results have to be treated with some level of caution. In particular, it is unclear whether these measures can be used to assess pressure-deficient WDNs even for the purpose of comparing networks with similar less than fully satisfactory demand satisfaction ratios because the surplus power term (in Eqs. 3-5) is negative for nodes with less than fully satisfactory pressure. A possible refinement of this research involves the removal of the designs that are not Pareto optimal on a cost vs. entropy basis before assessing the strengths of the relationship between the various

reliability measures; and verification using larger networks is obviously indicated.

REFERENCES

- Cullinane M. J., Lansey K. E. and Mays L. W. 1992 Optimization-availability-based design of water distribution networks. *Journal of Hydraulic Engineering*. **118**(3), 420–441.
- Fujiwara O. and Tung H. 1991 Reliability improvement for water distribution networks through increasing pipe size. *Water Resour. Res.* **27**(7), 1395 – 1402.
- Jayaram N. and Srinivasan K. 2008 Performance-based optimal design and rehabilitation of water distribution networks using life cycle costing. *Water Resour. Res.* **44**, W01417.
- Khomsy D., Walters G. A., Thorley A. R. D. and Ouazar D. 1996 Reliability tester for water distribution networks. *J. Comp. Civil Eng.* **10**(1), 10-19.
- Prasad T. D. and Park N-S. 2004 Multiobjective genetic algorithms for design of water distribution networks. *J. Water Resources Planning and Management*. **130**(1), 73-82.
- Reca J., Martinez J., Banos R. and Gil C. 2008 Optimal design of gravity-fed looped water distribution networks considering the resilience index. *J. Water Resources Planning and Management*. **134**(3), 234-238.
- Shannon C. E. 1948 A mathematical theory of communication. *Bell System Technical Journal*. **27**, 379-423 and 623-656.
- Su Y., Mays L. W., Duan N. and Lansey K. E. 1987 Reliability based optimization model for water distribution systems. *J. Hydraulic Engineering*. **114**(12), 1539-1556.
- Tabesh M., Soltani J., Farmani R. and Savic D. 2009 Assessing pipe failure rate and mechanical reliability of water distribution networks using data-driven model. *Journal of Hydroinformatics*. **11**(1), 1-17.
- Tanyimboh T., Tabesh M. and Burrows R. 2001 Appraisal of source head methods for calculating reliability of water distribution networks. *J. Water Resources Planning and Management*. **127**(4), 206-213.
- Tanyimboh T. T., Tahar B. and Templeman A. B. 2003 Pressure-driven modelling of water distribution systems. *Water Science and Technology – Water Supply*. **3**(1-2), 255-262.
- Tanyimboh T. and Setiadi Y. 2008a Joint layout, pipe size and hydraulic reliability optimization of water distribution systems. *Engineering Optimization*. **40**(8), 729-747.
- Tanyimboh T. and Setiadi Y. 2008b Sensitivity analysis of entropy-constrained designs of water distribution systems. *Engineering Optimization*. **40**(5), 439-457.
- Tanyimboh T. and Sheahan C. 2002 A maximum entropy based approach to the layout optimization of water distribution systems. *Civ. Eng. and Env. Syst.* **19**(3), 223-253.
- Tanyimboh T. T. and Templeman A. B. 1993a Calculating maximum entropy flows in networks. *J. Operational Research Society*. **44**(4), 383-396.
- Tanyimboh T. T. and Templeman A. B. 1993b Maximum entropy flows for single-source networks. *Engineering Optimization*. **22**(1), 49-63.
- Tanyimboh T. T. and Templeman A. B. 1993c Optimum design of flexible water distribution networks. *Civil Engineering Systems*, **10**(3), 243-258.
- Tanyimboh T. T. and Templeman A. B. 1998 Calculating the reliability of single-source networks by the source head method. *Advances in Eng. Software*. **29**(7-9), 499-505.
- Tanyimboh T. T. and Templeman A. B. 2000 A quantified assessment of the relationship between the reliability and entropy of water distribution systems. *Engineering Optimization*. **33**(2), 179-199.
- Todini E. 2000 Looped water distribution networks design using a resilience index based heuristic approach. *Urban Water*. **2**(2), 115-122.
- Walski T. M. 2001 The wrong paradigm – why water distribution optimization doesn't work. *Journal of Water Resources Planning and Management*. **127**(4), 203-205.